Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

SCIENZE E TECNOLOGIE AGRARIE, AMBIENTALI E ALIMENTARI

Ciclo XXVII

Settore Concorsuale di afferenza: 07/G1

Settore Scientifico disciplinare: AGR/17

TITOLO TESI

ANALYSIS OF THE PIG GENOME FOR THE IDENTIFICATION OF GENOMIC REGIONS AFFECTING PRODUCTION TRAITS

Presentata da: Giuseppina Schiavo

Coordinatore Dottorato

Prof. Giovanni Dinelli

_____

Relatore

Prof. Luca Fontanesi

_____

Esame finale anno 2015

Contents

General Introduction

The Domestic Pig (*Sus scrofa*) originated from Wild Boar whose appearance in Eurasia has been estimated to go back to around 500,000 years. Since then, it has been spread in all over the world, with independent domestication events starting about 9,000 years ago, leading to breeds with highly different charachteristics from Europe to Asia (Groenen et al. 2012, Rubin et al. 2012). Depending on different market requirements, artificial selection pushed European commercial breeds toward two main different directions. In most part of Europe, customers requirements are oriented toward lean meat, and this aspects makes the "Heavy Pig" an almost completely Italian concept. Italian Large White pig selection started in the last decades, taking advantage of breeders and scientists experience concerning phenotypic traits and it dramatically improved during the last few years.

Italian Large White breed

The Italian meat industry is mainly based on the production of high quality tipical regional products and Protected Designation of Origin (PDO) products e.g Prosciutto di Parma and San Daniele. The dry-cured ham production is carried out with strict rules established by the ham Consortia. These rules go from the breeding strategies to the storage after slaughtering; the geographical origin of the animals is also controlled and regulated. In animals aimed at dry-cured ham production, most of the economic value is linked to the legs that will be processed and their fat content that will have an essential role in the aging. High quality ham does not contain additives and preservatives, and the aging is performed only by controlling the amount of added salt, humidity and temperature of the place in which they are processed (Reg. UE n. 1151/2012, Disciplinari Prosciutto di Parma DOP, Disciplinari Prosciutto San Daniele DOP).

The attentions that can lead to an optimal dry-cured ham start from the genetic selection. According to the national pig breeders association (ANAS) programs animals are selected to improve feeding efficiency, performance and carcass traits and meat quality traits. The piglets are brought to the

testing station at 30-45 days. During the growth and every two weeks, many parameters are measured such as Average Daily Gain (ADG) and feeding:gain ratio; after slaughtering, other characteristics are measured, including Back Fat Thickness (BFT). This is an indirect indicator of the amount of fat covering the legs. All these measurements are used for the calculation of Estimated Breeding Values (EBV) by means of multiple traits mixed models, namely BLUP animal model.

The national program of selection for Italian Large White pigs is aimed to maintain the traditional fat coverage of the carcass and improve meat quality for the transformation in tipical aged products. Furthermore, when compatible with the above goals, genetic selection points to increase fresh meat cuts and the adaptability of pigs to intensive livestock; the selection also aims to improve traits involved in reproduction.

The breeding program is based on sib-testing: it implies the calculation of EBV related on the measurement of production traits in triplets (2 females and 1 castrated male) from the same litter. They are slaughtered when they reach the weight of about 160 kg.

Due to these premises, it should be clear that the integration of genomic data with simple phenotypic aspects is fundamental for optimizing future breeding selection strategies.

Furthermore, the pig is also an interesting model for studying human traits in particular related to metabolic syndrome and obesity. Information that are developed in breeding programs could be a by-product also for this aspect.

The Pig Genome

The pig genome sequence has been assembled in 2012 thanks to the Swine Genome Sequencing Consortium (Groenen et al. 2012) and the latest version, Sscrofa10.2, is available since May 2012 at Ensembl database (Cunningham et. Al 2015). The assembly has been obtained using artificial bacterial chromosomes and whole-genome shotgun sequencing. It consists of a goldenpath of around 2.80 gigabases, arranged in 18 autosomal chromosomes, Y and X chromosomes and the

Mitochondrial DNA; beside the 20 total chromosomes, sequences of 4562 scaffolds still unplaced are available. The number of annotated genes is at present 25322, including those placed on scaffolds. Most of these genes have been automatically annotated as described in Ensembl pipeline (Pig annotation pipeline), while around 2000 come from the manually curated. At present, the database comprising SNPs and short indels identified in pig contains 28702985 and among them 28702828 are present in dbSNP.

SNP assay BeadChip Over 64000 SNPs can be assayed in one shot by means of high-density genomic array of Illumina Technology, the Porcine Illumina BeadChip 60k. Briefly, the technology consists of a series of beads carrying probes with DNA segments on their surface. The probes are generally 121bp long, carrying the SNP to be tested in the middle. After proper shearing and binding of DNA fragments to the beads, a replication step is carried on with fluorescently marked nucleotides. Only in presence of one form of the allele the replication event will occur and a fluorescent signal will be detected.

The choice of the 60k SNPs has been determined after several Next Generation Sequencing experiments combined with SNP discovery techniques (Ramos et al. 2009) from pools of 5 different breeds. The latest version, released in Jun 2012 contains 64232 declared SNPs.

The application of BeadChip in agriculture ranges from genomic selection to studies of diversity and to the detection of QTL regions. The fluorescent signal can be interpreted both in a qualitative way (namely the absence ore presence of the allele in exam), useful in case of Genome Wide Association Studies (GWAS) or to quantify the effective amount of genomic material with that exact sequence, as it happens for Copy Number Variation (CNV) studies.

Examples and application of genotype studies

Association studies have already identified some candidate genes associated with production traits in other pig breeds intended for ham production (Corominas et al. 2013, 2012) and analysis at

genome wide level, with Porcine BeadChip, successfully found markers associated with meat quality and composition (Ramajo-Caldas et al. 2012).

Concerning CNV studies, it is known that they are a big source of variability in mammals and that their presence can lead to drastic phenotype changes, as it is known for the KIT allele, whom Copy Number duplication causes the white color in pigs (Marklund et al. 1998). Several works detected CNV regions in different pig breeds (Ramayo-Caldas et al 2010, Wang et al. 2013), enlarging the landscape of pig variability.

It is interesting to observe how allele frequency changed in animals undergone to strong artificial selection processes; as in the case of Texel sheep (Boman et al. 2011), it is worth to note the side effects of breeding schemes.

Aim

The aim of this Thesis was to identify markers associated with production traits in the pig genome using different approaches. The peculiarity of the following experimental designs are the choice of samples, that were mainly based on Selective Genotyping approach: for some of the following described works, two sets of pigs with completely divergent EBVs for some production traits have been genotyped. The idea was to clearly separate animals depending on their phenotype, to treat them as case-control, and making the GWAS more effective than it is in a global population sample. Another subset of animals, consisting of pigs born within two decades, gives the possibility to have an eyesight on the temporal changes in the pig genome during a relatively short time frame that could produce selection signature.

In the first two chapters we have shown a GWAS study using SNPs, searching for association between genotypes and one of the main carcass characteristics, Back Fat Thickness and one of the most important production traits, Average Daily Gain. Since we were also interested in other markers affecting production traits, we applied a slightly different approach, and searched for CNV associated with BFT. The CNV identification has been described in chapter 3. To understand how selection shaped the allele frequency of some candidate genes and thousands of other SNPs, we compared their frequency in groups of pigs born in different years. For this dataset, in the fourth and fifth chapters, we have described how allele frequency changed during years. In chapter four, we have focused the attention on SNPs present in few candidate genes (*IGF2*, *MC4R*, *VRTN*, *PRKAG3* and *FTO*) that have great importance in meat production; in the fifth chapter we have described allele frequency differences at genome wide level. In the last chapter we have explored how Next Generation Sequencing can be optimized to discover new polymorphisms that can be associated to Back Fat Thickness, starting from pools of animals with divergent EBV for this traits.

CHAPTERS

1. A genome wide association study for backfat thickness in Italian Large White pigs highlights new regions affecting fat deposition including neuronal genes
2. A genome wide association study for average daily gain in Italian Large White pigs
3. Copy number variants in Italian Large White pigs detected using high-density single nucleotide polymorphisms and their association with back fat thickness
4. A retrospective analysis of major gene allele frequency changes during 20 years of selection in the Italian Large White pig breed
5. Genome wide allele frequency changes over twenty years of artificial directional selection in the Italian Large White pig breed
6. Reduced representation libraries from DNA pools analysed with next generation semiconductor based-sequencing to identify SNPs in extreme and divergent pigs for back fat thickness

CHAPTER 1

A genome wide association study for backfat thickness in Italian Large White pigs using a selective genotyping approach

Luca Fontanesi [1,2,§], Giuseppina Schiavo [1,2,*], Giuliano Galimberti [3,*], Daniela Giovanna Calò [3,*], Emilio Scotti [1], Pier Luigi Martelli [2,4], Luca Buttazzoni [5], Rita Casadio [2,4] and Vincenzo Russo [1]

[1] Department of Agro-Food Science and Technology, University of Bologna, Viale Fanin 46, 40127 Bologna Italy

[2] Centre of Genome Biology, University of Bologna, 40126 Bologna, Italy

[3] Department of Statistical Sciences "Paolo Fortunati", University of Bologna, Via delle Belle Arti 41, 40126 Bologna, Italy

[4] Biocomputing Group, Department of Experimental Evolutionary Biology, University of Bologna, Via San Giacomo, 40126 Bologna, Italy

[5] CRA, Centro di Ricerca per la Produzione delle Carni e il Miglioramento Genetico, Via Salaria 31, 00015 Monterotondo Scalo, Roma, Italy

* These authors contributed equally

Published in BMC Genomics. 2012 Nov 15;13:583. doi: 10.1186/1471-2164-13-583.

Abstract

Background

Carcass fatness is an important trait in most pig breeding programs. Following market requests, breeding plans for fresh pork consumption are usually designed to reduce carcass fat content and increase lean meat deposition. However, the Italian pig industry is mainly devoted to the production of Protected Designation of Origin dry cured hams: pigs are slaughtered at around 160 kg of live weight and the breeding goal aims at maintaining fat coverage, measured as backfat thickness to

avoid excessive desiccation of the hams. This objective has shaped the genetic pool of Italian heavy pig breeds for a few decades. In this study we applied a selective genotyping approach within a population of ~ 12,000 performance tested Italian Large White pigs. Within this population, we selectively genotyped 304 pigs with extreme and divergent backfat thickness estimated breeding value by the Illumina PorcineSNP60 BeadChip and performed a genome wide association study to identify loci associated to this trait.

Results

We identified 4 single nucleotide polymorphisms with P≤5.0E-07 and additional 119 ones with 5.0E-07<P≤5.0E-05. These markers were located throughout all chromosomes. The largest numbers were found on porcine chromosomes 6 and 9 (n=15), 4 (n=13), and 7 (n=12) while the most significant marker was located on chromosome 18. Twenty-two single nucleotide polymorphisms were in intronic regions of genes already recognized by the Pre-Ensembl Sscrofa10.2 assembly. Gene Ontology analysis indicated an enrichment of Gene Ontology terms associated with nervous system development and regulation in concordance with results of large genome wide association studies for human obesity.

Conclusions

Further investigations are needed to evaluate the effects of the identified single nucleotide polymorphisms associated with backfat thickness on other traits as a pre-requisite for practical applications in breeding programs. Reported results could improve our understanding of the biology of fat metabolism and deposition that could also be relevant for other mammalian species including humans, confirming the role of neuronal genes on obesity.

Background

Fat deposition is a key biological process that has important similarities between humans and pigs, potentially useful to elucidate mechanisms determining human obesity. This trait has practical and economical implications in pig breeding as it indirectly affects feeding efficiency and determines carcass value and consumers' acceptance of pork.

Following consumer demands, breeding goals for fresh pork generally aim at reducing carcass fatness and increasing lean meat content which has adversely affected pork quality (e.g. [1]).

The Italian pig breeding industry is mainly devoted to the production of high quality Protected Designation of Origin (PDO) dry cured hams for which pigs are raised until they reach about 160 kg live weight and appropriate fat coverage of the hams is required [2,3]. Therefore, breeding objectives aim at maintaining fat coverage measured as backfat thickness (BFT). This objective has shaped the genetic pool of Italian heavy pig breeds for a few decades.

To investigate molecular genetic aspects of fat deposition in these pigs, we have recently applied a systematic candidate gene approach and have identified tens of single nucleotide polymorphisms (SNPs) associated with BFT and/or intermuscular fat content in Italian Large White and Italian Duroc pigs [4-10]. For example, a list of more than 30 SNPs has been associated with BFT in Italian Large White, including SNPs already found by other authors in *IGF2*[11], *MC4R*[12], *TBC1D1*[8], *PPARG*[13] genes or newly identified in the *PCSK1*[14], *ACP2*, *CALR*, *JAK3*, and *NT5E*, among several other genes [10]. Moreover, many other SNPs in additional candidate genes have been shown to explain a proportion of genetic variability of fat deposition traits in pigs [4,15,16]. In addition, a large number of QTLs for a variety of fat deposition and related traits have been already reported and listed in the Pig QTL database [17,18].

Recently, with the development of a commercial high throughput SNP genotyping tool in pig (PorcineSNP60 BeadChip [19]), a number of genome wide association (GWA) studies have been

carried out in this species focusing on reproduction [20,21], boar taint [22,23], disease resistance[24], structural and body composition, including BFT [25].

We have herein applied a selective genotyping approach in the Italian Large White pig breed and genotyped extreme and divergent pigs for BFT estimated breeding value (EBV) by the Illumina PorcineSNP60 BeadChip (http://www.illumina.com webcite) tool to identify chromosome regions and markers associated with BFT.

Results and Discussion

SNP data

A total of 304 performance tested Italian Large White pigs were genotyped with the Illumina PorcineSNP60 BeadChip, interrogating 62,163 loci. One pig was excluded from further analysis due to a call rate below 90%. A call rate ≥0.90 was obtained for 58,680 SNPs (for 2,293 SNP, call rate was 0.0; 1,190 SNP had 0.0<call rate<0.90). About 15.8% (9,287 SNPs) of these potentially useful SNPs had a minor allele frequency <0.05 and were discarded. The remaining 49,393 SNPs were re-mapped on the Sscrofa10.2 genome assembly.

Genome wide association (GWA) results

Only individuals with extreme phenotypes were genotyped for association study. Several authors have shown that this approach allows to attain the same power with less genotyped individuals (e.g. [26-28]). A recent GWA study for human obesity showed that this design can obtain very similar results to previous studies on general body mass index performed on unselected cohorts of tens of thousands of subjects [29].

In our study, genotyped pigs had extreme and divergent EBV for BFT: 151 had the lowest (thinnest BFT) and 152 the highest (thickest BFT) EBV. These animals were two generation unrelated gilts taken from the performance test of the National selection program of the Italian Large White breed carried out by the National Pig Breeders Association (ANAS).

Figure 1 reports a Manhattan plot showing significant (P≤5.0E-07) and suggestively significant (5.0E-07<P≤5.0E-05) SNPs (Pnominal value thresholds for significant results were those indicated by the Wellcome Trust Case Control Consortium, WTCCC [30]). Using these values, 4 SNPs were significantly associated (Table 1) whereas 119 SNPs were suggestively associated with BFT (Additional file 1: Table S1).

The WTCCC criteria to reduce the number of false positive are rather conservative: the suggestive threshold for significance (P=5.0E-05) corresponds to a FDR of 0.02. Had we assumed a false discovery rate (FDR) of 0.05 (Pnominal value = 0.000412), a total of 410 SNPs would had been below this threshold and therefore considered at least "suggestively associated" (data not shown).

Single nucleotide polymorphisms with P≤5.0E-05 were located in all porcine autosomal chromosomes (SCC), and on SSCX, and 5 SNPs were in unassembled scaffolds of the Sscrofa10.2 genome version. Among the mapped SNPs, the largest number was on SSC6 and SSC9 (n = 15), SSC4 (n = 13), SSC7 (n = 12) and SSC1 (n = 11) (Additional file 1: Table S1). Twenty-two SNPs were in intronic regions of recognized genes in the Pre-Ensembl Sscrofa10.2 assembly. The closest gene for the remaining mapped SNPs (n = 96) was located in a range from 481 bp to 4.69 Mb (mean = 287.6 kb ± 580 kb, median = 88 kb).

The most significant SNP (ALGA0098168; P=3.07E-11) was on SSC18 (Table 1). This SNP was localized in intron 3 of the phosphodiesterase 1C, calmodulin-dependent 70kDa (*PDE1C*) gene. PDEC1, highly expressed in brain and heart, is involved in the regulation of the cellular level of adenosine 3′,5′-cyclic monophosphate (cAMP) and guanosine 3′,5′-cyclic monophosphate (cGMP) that play critical roles in signal transduction [31]. The second most significant SNP (M1GA0010276; P=1.45E-08) was localized on SSC7 at about 3.6 kb from the cysteine-rich secretory protein 1 (*CRISP1*) gene whose known function in reproduction processes is not directly linked to any fat or energy related biological function. The third top SNP (ALGA0109557; P=3.81E-07) was mapped on SSC15 at about 460 kb from the signal transducer and activator of transcription 4 (*STAT4*) gene. STAT4 is a member of the STAT family of transcription factors that

transduces interleukin and type 1 interferon cytokine signals in T cells and monocytes, leading to important immunological functions. Reduction of STAT4 activation has been proposed to control obesity-induced inflammation [32]. The fourth most significant marker (ALGA0069549; P=3.87E-07) was located on SSC13 at about 20 kb from the stabilin 1 (*STAB1*) gene. Another close marker (ALGA0109216; at position 38330168 of SSC13; Additional file 1: Table S1) was suggestively significant (P=1.01E-06). The SSC13 region bracketed by these two SNPs includes the *STAB1*-nischarin (*NISCH*) gene interval that in human has been shown to be associated with waist-hip ratio (a measure of body fat distribution) [33].

Several other genes close or within the additional suggestively significant SNPs (Additional file 1: Table S1) have been already involved in obesity related biological mechanisms. Among this list it is worth mentioning: ATP-binding cassette, sub-family B (MDR/TAP), member 1 (*ABCB1*) gene (SSC9; ALGA0109564, P=9.01E-07) whose altered function contributes to steatosis and obesity in mice [34] and a polymorphism in this gene has been associated with obesity risk in Japanese subjects [35]; galanin receptor 3 (*GALR3*) gene (SSC5; M1GA0007458, P=1.25E-06) that is upregulated in adipose tissues of mice fed a high fat diet [36], and whose function is to bind galanin, a neuropeptide that regulates food intake, neurogenesis, memory, and gut secretion; olfactory receptor genes (two genes on SSC9, *OR52N2* and *OR56A3*) have been associated with eating behaviour and adiposity in humans [37]; Parkinson protein 2 (*PARK2*) gene on SSC1 (ALGA0108518, P=5.48E-06) that is regulated in a lipid-dependent manner and modulates systemic fat uptake via ubiquitin ligase-dependent effects [38]; phosphodiesterase 4B, cAMP-specific (*PDE4B*) on SSC6 (ALGA0109354, P=5.95E-06) that has been already shown to be associated with BFT in pigs as well as with obesity in humans [39]; vacuolar protein sorting 13 homolog B (yeast) (*VPS13B*) on SSC4 (ALGA0024658, P=3.00E-05) that causes Cohen syndrome, characterized by truncal obesity [40]; iroquois homeobox 3 (*IRX3*) gene on SSC6 (M1GA0008432, P=4.66E-05), that is involved in the stress response after fat loss [41] and could be linked to obesity

and type 2 diabetes through its pancreatic function [42]. Interestingly the second closest gene to this latter SSC6 SNP was *FTO*, that is well known to affect human obesity (i.e. [43]).

Even though the annotation of the pig genome available at present in Pre-Ensembl should be considered preliminary, we further evaluated the potential functional role of regions around associated or suggestively associated SNPs with BFT in our pig population (Additional file 1: Table S1). For this evaluation we used Gene Ontology (GO) information of their corresponding closest genes. Table 2 reports GO terms enriched in this dataset. Interestingly, most of the statistically significant GO terms were related to nervous system development and regulation. This indication might support and extend the role of the nervous system in the genetic predisposition of fat accumulation in mammals, as in part reported in large GWA studies in humans [44] and, subsequently, in pigs [39]. Among the genes listed in these neuronal GO categories (Table 2), few have been already reported to be indirectly associated or involved in obesity related traits. Apart from those already described above (*IRX2* and *PARK2*), it is interesting to mention the delta-like 1 (Drosophila) gene (*DLL1*) as this gene is located in a quite large region (~1 Mb) associated to type 1 diabetes on human chromosome 6 [45]. For several other genes involved in neuronal processes, at present, there is no direct reported link with obesity or fat metabolism. It would be important to further explore their role in affecting the investigated phenotype as a possible strategy to identify new pathways and mechanisms affecting fat deposition. For example, it could be possible to speculatively suggest a relationship between dysbindin (DTNBP1), involved in the modulation of glutamatergic neurotransmission in the brain, schizophrenia and obesity [46].

Comparison with other studies in pigs

We compared our GWA results with results obtained in our previous candidate gene studies for BFT in pigs [4-10] and those obtained by other GWA [25] and QTL mapping studies. In our previous studies [5,10], the *IGF2* intron3-g.3072G>A mutation [11] was the most significant marker ($P < 1.00E-50$ by selective genotyping [10]). As the *IGF2* gene is not assembled in the Sscrofa10.2 genome version, it was not possible to obtain a direct comparison with results obtained

for SNPs mapped on SSC2 included in the Illumina PorcineSNP60 BeadChip. However, no SNP in the region where *IGF2* is likely to be found (0–10 Mb) reached the significance level of P<5.0E-05 (Additional file 1: Table S1). Only one SNP (ASGA0008884, position 9139348; P=2.12E-04) was included in the list of markers with PFDR<0.05. Several other SSC2 SNPs were suggestively significant (Additional file 1: Table S1) indicating that they might pick up other regions affecting fatness as already reported by QTL studies (e.g. [47,48]) or candidate gene studies [5,10,49,50].

The second most significant marker of our previous candidate gene investigation was the *MC4R*p.Asp298Asn substitution [10]. In the current GWA study, no significant or suggestively significant SNPs were located in the SSC1 region around the *MC4R* gene, even if a few markers had P<1.0E-3 (data not shown). The GWA study by Fan et al. [25], conducted on gilts of a commercial breeding stock, showed that markers around *MC4R* were significantly associated with 10th rib and last rib backfat. These slight differences in terms of level of significance of the markers between the two studies might be due to different *MC4R* haplotype structures in the two pig populations (Fontanesi et al. submitted) or to different positions in the pig body where BFT measurements were taken. However, in general, few results we obtained confirmed those previously obtained by Fan et al.[25] in their GWA study on BFT. This could be due to different experimental designs, incomplete power in the two studies, and/or to differences between the investigated populations. Other results we previously obtained in candidate gene studies (i.e. [10]) could be confirmed if we relaxed the significance threshold up to FDR <0.05 (data not shown).

QTLs for fat deposition traits can be found over all pig chromosomes. Many different studies have repeatedly reported the presence of complex QTL patterns for fat related traits in SSC1, SSC2, SSC4, SSC6 and SSC7 [18]. In the present GWA study, SSC4, SSC6, SSC7, and also SSC9 resulted to be rich in significant or suggestively significant markers (SSC4: expected proportion = 0.068, observed = 0.110; SSC6: expected = 0.059, observed = 0.127; SSC7: expected = 0.063, observed = 0.102; SSC9: expected = 0.061, observed = 0.127). These results seem to indicate these chromosomes to support an important proportion of genetic variability for BFT in the Italian Large

White breed. In particular, two markers below the suggestive significance threshold were located both on *IGSF3* or close to *PKN2* on SSC4 and a few close blocks of SNPs with P<5.0E-05 (from about 65.1 - 65.4, 70.6 - 72.5, and 100.7 - 101.8 Mb) were located on SSC6 (Additional file1: Table S1). As mentioned above, *FTO* is close to the marker at position 28215213 on SSC6. Single marker analysis using a few *FTO* SNPs in our previous large association study with BFT in Italian Large White pigs did not produce significant results [10]. However, subsequent haplotype analysis at this locus tended to confirm *FTO* as an important locus affecting fat deposition also in this pig breed [51].

Conclusions

This study is the first genome wide association analysis for BFT in Italian heavy pigs. The targeted trait is of paramount importance for the Italian pig breeding industry that is devoted to the production of high quality dry-cured hams for which fat coverage is a key factor during the processing and curing steps [2,3]. The genetic dissection of BFT could open new perspectives to improve selection efficiency. In this study we applied a selective genotyping approach within the Italian Large White pig population to reduce the cost of genotyping without losing much power [26-31]. We took advantage of the large number of pigs that have been performance tested and genetically evaluated under the National selection program for this breed. The association analysis that compared SNP genotype frequencies between low BFT-EBV vs. high BFT-EBV groups identified 123 SNPs with P<5.0E-5 that were more densely represented in a few chromosomes known to harbor important QTLs for fat deposition traits. The quite large number of markers below this threshold (spread in different chromosome regions) might indirectly support the fact that many genes, each with a small-medium contribution, are involved in determining BFT, according to the classical definition of a quantitative trait.

Several significant or suggestively significant SNPs were close to genes whose function might be directly or indirectly related to energy metabolism and fat deposition. Many other cannot be easily linked to the targeted trait and might provide, if confirmed in following up studies, new evidence on

this matter. Even if the annotation available in Pre-Ensembl for Sscrofa10.2 is preliminary, GO enrichment analysis indicated that neuronal genes might affect fat deposition in pig confirming and enlarging previous indications reported in humans [44].

Summarizing, as more information is becoming available in pigs on biological aspects of fat metabolism and deposition, it is more and more clear that this species could represent an attractive biomedical model for human obesity and associated diseases. Data here reported could give an insight over genetic mechanisms of fat metabolism and deposition that could be helpful in understanding also biology aspects of human obesity.

Methods

Animals and phenotypic traits

All animals used in this study were kept according to Italian and European legislation for pig production and all procedures described were in compliance with national and European Union regulations for animal care and slaughtering.

The national selection program of the Italian Large White breed is based on triplets of siblings from the same litter (two females and one castrated male) that are individually performance tested at the Central Test Station of the National Pig Breeder Association (ANAS) for the genetic evaluation of a boar from the same litter (sib-testing). Performance evaluation starts when the pigs are 30 to 45 days of age and it ends when the animals reach 155 ± 5 kg live weight. The nutritive level is *quasi ad libitum*, meaning that about 60% of the pigs are able to ingest the entire supplied ration. At the end of test, animals are transported to a commercial abattoir where they are slaughtered following standard procedures [52]. Then, backfat thickness is measured on the carcasses at the level of *Musculus gluteus medius*.

The association study was conducted following a selective genotyping approach (e.g. [26-31]). In this study we genotyped two extreme and divergent groups of Italian Large White gilts of these

triplets (one female per triplet), performance tested in the period 1996–2007. Two-generation unrelated females (i.e. gilts with different and unique parents) were chosen according to their EBV for BFT (152 with most negative and 152 with most positive EBV) within a performance tested population of ~12000 pigs (details of EBV calculation are reported below). The two extreme groups were chosen ranking the animals according to their BFT EBV and then taking only the first unrelated gilts in the list (with the most positive or the most negative BFT EBV). BFT EBV used to choose the animals were recalculated for the whole performance tested population in 2007. Average BFT EBV in the negative and positive selected groups of pigs were −9.8 ± 1.6 mm and +6.6 ± 2.3 mm, respectively. Genotyped pigs were a subset of the 560 two-generation unrelated pigs used in our previous candidate gene association study [10].

Genotyping

Genomic DNA was extracted from dried-blood by standard protocols. Based on quality control, all animals were used for genotyping using the PorcineSNP60 BeadChip [19] developed by Illumina according to manufacturer's protocol [53].

Data analyses

Estimated breeding values for BFT were calculated in the population using a BLUP-Multiple Trait-Animal Model that included the fixed effect of sex (considering the triplets of pigs from the same litter), batch on trial, inbreeding coefficient of the animal, interaction of sex by age at slaughtering, date of slaughtering and random effect of litter and animal. Three criteria were used to filter animals and SNP before association analysis: call rate >0.9 both at the 1) animal and 2) SNP level, and 3) MAF>0.05. Animals and SNPs that passed these filters were taken for association analysis treating the two groups as cases and controls. Full pedigree information available was used to obtain a kinship matrix. In order to correct for possible family-based stratification (see Additional file 2: Figure S1), the EIGENSTRAT method [54] was applied including the kinship matrix, and association tests were performed. All analyses were performed in R [55], using an option of the

package GenABEL [56] for computing the test-statistics according to the EIGENSTRAT method, and the package kinship [57] for building the pedigree kinship matrix.

For *n* animals, the first *K*<*n* principal components, $\boldsymbol{c}1, \ldots, \boldsymbol{c}K$, of the kinship matrix among the animals were used as axes of genetic variation. Let *gij* and *pj* be the genotype at SNP ***i*** (*gij* = 0, 1 or 2) and the phenotype of animal ***j***, respectively, a PC-based adjustment was performed on genotypes and phenotypes according to the following formulas:

$$g_{ij}^{*} = g_{ij} - \beta_{1i} c_{1j} - .. - \beta_{Ki} c_{Kj} \quad (1)$$

$$p_{j}^{*} = p_{j} - \gamma_{1} c_{1j} - .. - \gamma_{K} c_{Kj} \quad (2)$$

where *ckj* is the score of the *k*-th component on animal *j*, *βki* and *γk* are the partial regression coefficients for predicting the ***i***-th genotype and the phenotype, respectively, on the basis of the ***k***-th component (with $k = 1, ., K$).

The association test-statistic is computed as $(n - K - 1) ri2$, where

$$r_{i}^{2} = \frac{\left(\sum_{j=1}^{n} p_{j}^{*} g_{ij}^{*}\right)^{2}}{\sum_{j=1}^{n} \left(p_{j}^{*}\right)^{2} \sum_{j=1}^{n} \left(g_{ij}^{*}\right)^{2}} \quad (3)$$

is the squared correlation coefficient between the ***i***-th PC-adjusted genotype and PC-adjusted phenotype. As noted by Price et al. [54], this statistic is a generalization of the Armitage trend statistic for discrete genotypes and phenotypes.

Wellcome Trust Case Control Consortium significance thresholds, whose definition depends on the prior odds and power, were used in this study [30]. In addition, correction for multiple testing was achieved by using a False Discovery Rate approach [58]. For each chromosome, the expected proportion of SNPs with P<5.0E-5 was computed under the assumption of uniform distribution

from the informative SNPs over the chromosome. This proportion was compared to the proportion of significant or suggestively significant markers actually observed on the same chromosome.

Bioinformatics analyses

Mapping of the PorcineSNP60 BeadChip SNPs was obtained by using BWA [59] on the Sscrofa9.2 and Sscrofa10.2 genome assemblies as previously described [10] and confirmed using the BLAT analysis available at the Ensembl (http://www.ensembl.org/Sus_scrofa/Info/Index webcite) and Pre-Ensembl (http://pre.ensembl.org/Sus_scrofa/Info/Index webcite) databases (February 2012). Coordinates for the Sscrofa10 genome preliminary version (September 2010) were downloaded from the Animal Genome repository web sitehttp://www.animalgenome.org/repository/ webcite. Identification of the closest genes to SNPs with P<5.0E-05 was obtained using Pre-Ensembl annotation of Sscrofa10.2 genome version and verified using Ensembl Sscrofa9.2 genome version (February 2012). Starting from the corresponding protein sequences retrieved from these databases, the corresponding gene symbols were extracted from NCBI Gene section (http://www.ncbi.nlm.nih.gov/ webcite) and/or Uniprot (http://www.uniprot.org/ webcite) databases (February 2012). Gene annotation was verified by BLAST analysis (http://blast.ncbi.nlm.nih.gov/ webcite). Gene Ontology analysis was carried out using DAVID Bioinformatics Resources 6.7 (http://david.abcc.ncifcrf.gov/ webcite[60]).

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

LF conceived and coordinated the study, analysed data and drafted the manuscript. GS, GG, DGC, and PLM performed statistical and bioinformatics analyses. ES carried out laboratory activities. LB, RC and VR coordinated and conceived the study. All authors reviewed and contributed to draft the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We thank ANAS for providing samples and data. This work was supported by University of Bologna RFO and FAGenomicH projects and by Italian MiPAAF INNOVAGEN funds.

References

1. Lonergan SM, Huff-Lonergan E, Rowe LJ, Kuhlers DL, Jungst SB: **Selection for lean growth efficiency in Duroc pigs influences pork quality.**
J Anim Sci 2001, **79:**2075-2085.
2. Russo V, Nanni Costa L: **Suitability of pig meat for salting and the production of quality processed products.**
Pig News Inform 1995, **16:**7N-26N.
3. Bosi P, Russo V: **The production of the heavy pig for high quality processed products.**
Ital J Anim Sci 2004, **3:**309-321.
4. Fontanesi L, Scotti E, Buttazzoni L, Davoli R, Russo V: **The porcine fat mass and obesity associated (FTO) gene is associated with fat deposition in Italian Duroc pigs.**
Anim Genet 2009, **40:**90-93.
5. Fontanesi L, Speroni C, Buttazzoni L, Scotti E, Dall'Olio S, Nanni Costa L, Davoli R, Russo V:**The insulin-like growth factor 2 (IGF2) gene intron3-g.3072G>A polymorphism is not the only Sus scrofa chromosome 2p mutation affecting meat production and carcass traits in pigs: evidence from the effects of a cathepsin D (CTSD) gene polymorphism.**
J Anim Sci 2010, **88:**2235-2245.
6. Fontanesi L, Colombo M, Scotti E, Buttazzoni L, Bertolini F, Dall'Olio S, Davoli R, Russo V:**The porcine tribbles homolog 3 (TRIB3) gene: identification of a missense mutation and association analysis with meat quality and production traits in Italian heavy pigs.**
Meat Sci 2010, **86:**806-813.

7. Fontanesi L, Scotti E, Speroni C, Buttazzoni L, Russo V: **A selective genotyping approach identifies single nucleotide polymorphisms in porcine chromosome 2 genes associated with production and carcass traits in Italian heavy pigs.**
Ital J Anim Sci 2011, **10:**e15.

8. Fontanesi L, Colombo M, Tognazzi L, Scotti E, Buttazzoni L, Dall'Olio S, Davoli R, Russo V:**The porcine TBC1D1 gene: mapping, SNP identification, and association study with meat, carcass and production traits in Italian heavy pigs.**
Mol Biol Rep 2011, **38:**1425-1431.

9. Fontanesi L, Speroni C, Buttazzoni L, Scotti E, Dall'Olio S, Davoli R, Russo V: **Association between polymorphisms in cathepsin and cystatin genes with meat production and carcass traits in Italian Duroc pigs: confirmation of the effects of a cathepsin L (CTSL) gene marker.**
Mol Biol Rep 2012, **39:**109-115.

10. Fontanesi L, Galimberti G, Calò DG, Fronza R, Martelli PL, Scotti E, Colombo M, Schiavo G, Casadio R, Buttazzoni L, Russo V: **Identification and association analysis of several hundred single nucleotide polymorphisms within candidate genes for back fat thickness in Italian Large White pigs using a selective genotyping approach.**
J Anim Sci 2012, **90:**2450-2464. Van Laere A-S, Nguyen M, Braunschweig M, Nezer C, Collette C, Moreau L, Archibald AL, Haley CS, Buys N, Tally M, Andersson G, Georges M, Andersson L: **A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig.**
Nature 2003, **425:**832-836.

11. Kim KS, Larsen N, Short T, Plastow G, Rothschild MF: **A missense variant of the porcine melanocortin-4 receptor (MC4R) gene is associated with fatness, growth, and feed intake traits.**
Mamm Genome 2000, **11:**131-135.

12. Fan B, Onteru SK, Mote BE, Serenius T, Stalder KJ, Rothschild MF: **Large-scale association study for structural soundness and leg locomotion traits in the pig.** ***Genet Sel Evol*** 2009, **41:**14.
13. Fontanesi L, Bertolini F, Scotti E, Trevisi P, Buttazzoni L, Dall'Olio S, Davoli R, Bosi P, Russo V: **Polymorphisms in an obesity related gene (PCSK1) are associated with fat deposition and production traits in Italian heavy pigs.** ***Animal*** 2012, **6:**1913-1924.
14. Fan B, Du ZQ, Rothschild MF: **The fat mass and obesity-associated (FTO) gene is associated with intramuscular fat content and growth rate in the pig.** ***Anim Biotechnol*** 2009, **20:**58-70.
15. Switonski M, Stachowiak M, Cieslak J, Bartz M, Grzes M: **Genetics of fat tissue accumulation in pigs: a comparative approach.** ***J Appl Genet*** 2010, **51:**153-168.
16. Rothschild MF, Hu Z-L, Jiang Z: **Advances in QTL mapping in pigs.** ***Int J Biol Sci*** 2007, **3:**192-197.
17. Hu Z-L, Reecy JM: **Animal QTLdb: Beyond a repository – A public platform for QTL comparisons and integration with diverse types of structural genomic information.** ***Mamm Genome*** 2007, **18:**1-4.
18. Ramos AM, Crooijmans RP, Affara NA, Amaral AJ, Archibald AL, Beever JE, Bendixen C, Churcher C, Clark R, Dehais P, Hansen MS, Hedegaard J, Hu ZL, Kerstens HH, Law AS, Megens HJ, Milan D, Nonneman DJ, Rohrer GA, Rothschild MF, Smith TP, Schnabel RD, Van Tassell CP, Taylor JF, Wiedmann RT, Schook LB, Groenen MA: **Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology.** ***PLoS One*** 2009, **4:**e6524.

19. Uimari P, Sironen A, Sevon-Aimonen ML: **Whole-genome SNP association analysis of reproduction traits in the Finnish Landrace pig breed.**

Genet Sel Evol 2011, **43:**42.

20. Onteru SK, Fan B, Du ZQ, Garrick DJ, Stalder KJ, Rothschild MF: **A whole-genome association study for pig reproductive traits.**

Anim Genet 2012, **43:**18-26.

21. Grindflek E, Lien S, Hamland H, Hansen MH, Kent M, van Son M, Meuwissen TH: **Large scale genome-wide association and LDLA mapping study identifies QTLs for boar taint and related sex steroids.**

BMC Genomics 2011, **12:**362.

22. Gregersen VR, Conley LN, Sorensen KK, Guldbrandtsen B, Velander IH, Bendixen C:**Genome-wide association scan and phased haplotype construction for quantitative trait loci affecting boar taint in three pig breeds.**

BMC Genomics 2012, **13:**22.

23. Boddicker N, Waide EH, Rowland RR, Lunney JK, Garrick DJ, Reecy JM, Dekkers JC:**Evidence for a major QTL associated with host response to Porcine Reproductive and Respiratory Syndrome virus challenge.**

J Anim Sci 2012.

24. Fan B, Onteru SK, Du ZQ, Garrick DJ, Stalder KJ, Rothschild MF: **Genome-wide association study identifies Loci for body composition and structural soundness traits in pigs.**

PLoS One 2011, **6:**e14726.

25. Darvasi A, Soller M: **Selective genotyping for determination of linkage between a marker locus and a quantitative trait locus.**

Theor Appl Genet 1992, **85:**353-359.

26. Van Gestel S, Houwing-Duistermaat JJ, Adolfsson R, van Duijn CM, Van Broeckhoven C:**Power of selective genotyping in genetic association analyses of quantitative traits.** ***Behav Genet*** 2000, **30:**141-146.

27. Zhang G, Nebert DW, Chakraborty R, Jin L: **Statistical power of association using the extreme discordant phenotype design.** ***Pharmacogenet Genomics*** 2006, **16:**401-413.

28. Paternoster L, Evans DM, Nohr EA, Holst C, Gaborieau V, Brennan P, Gjesing AP, Grarup N, Witte DR, Jørgensen T, Linneberg A, Lauritzen T, Sandbaek A, Hansen T, Pedersen O, Elliott KS, Kemp JP, St Pourcain B, McMahon G, Zelenika D, Hager J, Lathrop M, Timpson NJ, Smith GD, Sørensen TI: **Genome-wide population-based association study of extremely overweight young adults - the GOYA study.** ***PLoS One*** 2011, **6:**e24303.

29. Wellcome Trust Case Control Consortium: **Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls.** ***Nature*** 2007, **447:**661-683.

30. Lakics V, Karran EH, Boess FG: **Quantitative comparison of phosphodiesterase mRNA distribution in human brain and peripheral tissues.** ***Neuropharmacology*** 2010, **59:**367-374.

31. Chakrabarti SK, Wen Y, Dobrian AD, Cole BK, Ma Q, Pei H, Williams MD, Bevard MH, Vandenhoff GE, Keller SR, Gu J, Nadler JL: **Evidence for activation of inflammatory lipoxygenase pathways in visceral adipose tissue of obese Zucker rats.** ***Am J Physiol Endocrinol Metab*** 2011, **300:**E175-E187.

32. Heid IM, Jackson AU, Randall JC, Winkler TW, Qi L, Steinthorsdottir V, Thorleifsson G, Zillikens MC, Speliotes EK, Mägi R, Workalemahu T, White CC, Bouatia-Naji N, Harris TB, Berndt SI, Ingelsson E, Willer CJ, Weedon MN, Luan J, Vedantam S, Esko T, Kilpeläinen TO, Kutalik Z, Li S, Monda KL, Dixon AL, Holmes CC, Kaplan LM, Liang L,

Min JL, ***et al***.: **Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution.**

Nat Genet 2010, **42:**949-960.

33. Foucaud-Vignault M, Soayfane Z, Ménez C, Bertrand-Michel J, Martin PG, Guillou H, Collet X, Lespine A: **P-glycoprotein dysfunction contributes to hepatic steatosis and obesity in mice.**

PLoS One 2011, **6:**e23614.

34. Ichihara S, Yamada Y, Kato K, Hibino T, Yokoi K, Matsuo H, Kojima T, Watanabe S, Metoki N, Yoshida H, Satoh K, Aoyagi Y, Yasunaga A, Park H, Tanaka M, Nozawa Y: **Association of a polymorphism of ABCB1 with obesity in Japanese individuals.**

Genomics 2008, **91:**512-516.

35. Kim A, Park T: **Diet-induced obesity regulates the galanin-mediated signalling cascade in the adipose tissue of mice.**

Mol Nutr Food Res 2010, **54:**1361-1370.

36. Choquette AC, Bouchard L, Drapeau V, Lemieux S, Tremblay A, Bouchard C, Vohl MC, Pérusse L: **Association between olfactory receptor genes, eating behaviour traits and adiposity: results from the Quebec Family Study.**

Physiol Behav 2012, **105:**772-776.

37. Kim KY, Stevens MV, Akter MH, Rusk SE, Huang RJ, Cohen A, Noguchi A, Springer D, Bocharov AV, Eggerman TL, Suen DF, Youle RJ, Amar M, Remaley AT, Sack MN: **Parkin is a lipid-responsive regulator of fat uptake in mice and mutant human cells.**

J Clin Invest 2011, **121:**3701-3712.

38. Lee KT, Byun MJ, Kang KS, Park EW, Lee SH, Cho S, Kim H, Kim KW, Lee T, Park JE, Park W, Shin D, Park HS, Jeon JT, Choi BH, Jang GW, Choi SH, Kim DW, Lim D, Park HS, Park MR, Ott J, Schook LB, Kim TH, Kim H: **Neuronal genes for subcutaneous fat**

thickness in human and pig are identified by local genomic sequencing and combined SNP association study.

PLoS One 2011, **6:**e16356.

39. Balikova I, Lehesjoki AE, de Ravel TJ, Thienpont B, Chandler KE, Clayton-Smith J, Träskelin AL, Fryns JP, Vermeesch JR: **Deletions in the VPS13B (COH1) gene as a cause of Cohen syndrome.**

Hum Mutat 2009, **30:**E845-E854.

40. Dankel SN, Fadnes DJ, Stavrum AK, Stansberg C, Holdhus R, Hoang T, Veum VL, Christensen BJ, Våge V, Sagen JV, Steen VM, Mellgren G: **Switch from stress response to homeobox transcription factors in adipose tissue after profound fat loss.**

PLoS One 2010, **5:**e11033.

41. Ragvin A, Moro E, Fredman D, Navratilova P, Drivenes O, Engström PG, Alonso ME, de la Calle Mustienes E, Gómez Skarmeta JL, Tavares MJ, Casares F, Manzanares M, van Heyningen V, Molven A, Njølstad PR, Argenton F, Lenhard B, Becker TS: **Long-range gene regulation links genomic type 2 diabetes and obesity risk regions to HHEX, SOX4, and IRX3.**

Proc Natl Acad Sci USA 2010, **107:**775-780.

42. Dina C, Meyre D, Gallina S, Durand E, Körner A, Jacobson P, Carlsson LM, 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, Bougnères P, Kovacs P, Marre M, Balkau B, Cauchi S, Chèvre JC, Froguel P: **Variation in FTO contributes to childhood obesity and severe adult obesity.**

Nat Genet 2007, **39:**724-726.

43. Willer CJ, Speliotes EK, Loos RJ, Li S, Lindgren CM, Heid IM, Berndt SI, Elliott AL, Jackson AU, Lamina C, Lettre G, Lim N, Lyon HN, McCarroll SA, Papadakis K, Qi L, Randall JC, Roccasecca RM, Sanna S, Scheet P, Weedon MN, Wheeler E, Zhao JH, Jacobs

LC, Prokopenko I, Soranzo N, Tanaka T, Timpson NJ, Almgren P, Bennett A, Genetic Investigation of ANthropometric Traits Consortium, ***et al***.: **Six new loci associated with body mass index highlight a neuronal influence on body weight regulation.**

Nat Genet 2009, **41:**25-34.

44. Bradfield JP, Qu HQ, Wang K, Zhang H, Sleiman PM, Kim CE, Mentch FD, Qiu H, Glessner JT, Thomas KA, Frackelton EC, Chiavacci RM, Imielinski M, Monos DS, Pandey R, Bakay M, Grant SF, Polychronakos C, Hakonarson H: **A genome-wide meta-analysis of six type 1 diabetes cohorts identifies multiple associated loci.**

PLoS Genet 2011, **7:**e1002293.

45. Numakawa T, Yagasaki Y, Ishimoto T, Okada T, Suzuki T, Iwata N: **Evidence of novel neuronal functions of dysbindin, a susceptibility gene for schizophrenia.**

Hum Mol Genet 2004, **13:**2699-26708.

46. Lee SS, Chen Y, Moran C, Cepica S, Reiner G, Bartenschlager H, Moser G, Geldermann H:**Linkage and QTL mapping for Sus scrofa chromosome 2.**

J Anim Breed Genet 2003, **120**(Suppl. 1)**:**11-19.

47. Tortereau F, Gilbert H, Heuven HC, Bidanel JP, Groenen MA, Riquet J: **Number and mode of inheritance of QTL influencing backfat thickness on SSC2p in Sino-European pig pedigrees.**

Genet Sel Evol 2011, **43:**11.

48. Cepica S, Ovilo C, Masopust M, Knoll A, Fernandez A, Lopez A, Rohrer GA, Nonneman D:**Four genes located on a SSC2 meat quality QTL region are associated with different meat quality traits in Landrace × Chinese-European crossbred population.**

Anim Genet 2012, **43:**333-336.

49. Fontanesi L, Buttazzoni L, Scotti E, Russo V: **Confirmation of the association between a single nucleotide polymorphism in the porcine LDHA gene and average daily gain and correlated traits in Italian Large White pigs.**

Anim Genet 2012, **43:**649-650.

50. Fontanesi L, Russo V: **Nucleotide variability and haplotype heterogeneity at the porcine fat mass and obesity associated (FTO) gene.** ***Anim Genet***

51. Fontanesi L, Davoli R, Nanni Costa L, Beretti F, Scotti E, Tazzoli M, Tassone F, Colombo M, Buttazzoni L, Russo V: **Investigation of candidate genes for glycolytic potential of porcine skeletal muscle: association with meat quality and production traits in Italian Large White pigs.**

Meat Sci 2008, **80:**780-787.

52. Steemers FJ, Gunderson KL: **Whole genome genotyping technologies on the BeadArray platform.**

Biotechnol J 2007, **2:**41-49.

53. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D: **Principal components analysis corrects for stratification in genome-wide association studies.**

Nat Genet 2006, **38:**904-909.

54. R Development Core Team: R:

A language and environment for statistical computing.

http://www.R-project.org/

55. Aulchenko YS, Ripke S, Isaacs A, van Duijn CM: **GenABEL: an R library for genome-wide association analysis.**

Bioinformatics 2007, **23:**1294-1296.

56. Atkinson B, Therneau T:

kinship: mixed-effects Cox models, sparse matrices, and modeling data from large pedigrees.

R package version 1.1.0-23, 2009.

57. Benjamini Y, Hochberg Y: **Controlling the false discovery rate: a practical and powerful approach to multiple testing.**

J Royal Stat Soc Series B 1995, **57:**289-300.

58. Li H, Durbin R: **Fast and accurate short read alignment with Burrows-Wheeler transform.**

Bioinformatics 2009, **25:**1754-1760.

59. Huang DW, Sherman BT, Lempicki RA: **Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists.**

Nucleic Acids Res 2009, **37:**1-13.

Figure 1. Manhattan plot of genome wide association results for backfat thickness in Italian Large White pigs.

Red line: threshold for significant (P<5.0E-07). Blue line: threshold for suggestively significant results (5.0E-07<P<5.0E-05).

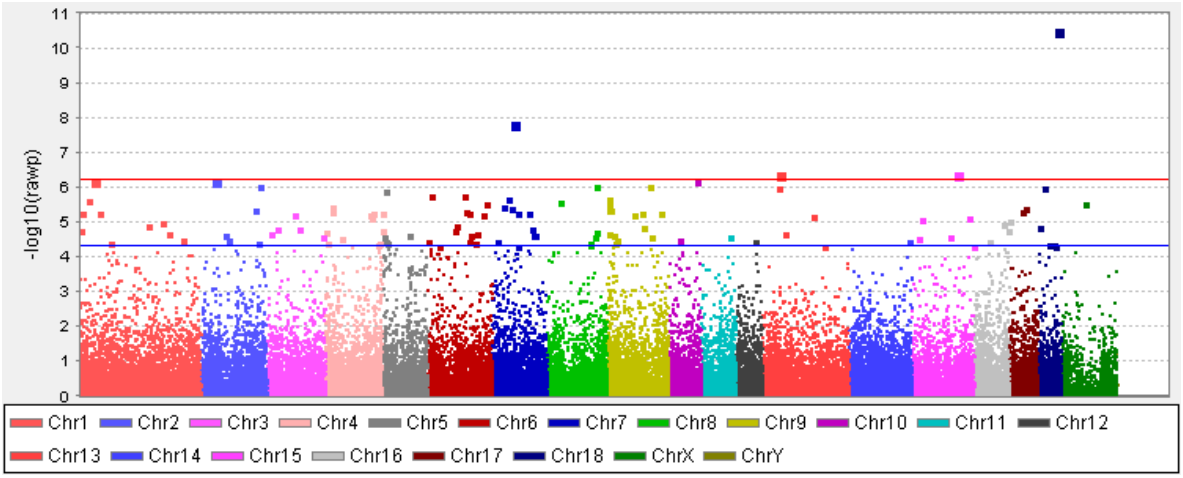


Table 1 - Significant SNPs (P<5.0E-07), their chromosome positions and their closest genes in Sscrofa10.2 (Pre-Ensembl)

SNP	CHR:position[1] (Sscrofa10.2)	CHR:position[1] (Sscrofa10)	CHR:position[1] (Sscrofa9.2)	P[1]	FDR[1]	SNP position/distance[1]	Closest gene coordinates	Closest gene protein in Pre-Ensembl	Gene Symbol
ALGA0098168	18:45408799	18:44540120	18:25516667	3.07E-11	1.51E-06	Intron 3	18:45404849-45567252	ENSSSCP00000017656	*PDE1C*
M1GA0010276	7:50272760	7:50024255	7:50440974	1.45E-8	3.58E-04	3657	7:50276417-50297466	ENSSSCP00000001892	*CRISP1*
ALGA0109557	15:107079255	15:102388547	-	3.81E-7	4.17E-3	4609542	15:102429318-102469713	ENSP00000351255	*STAT4*
ALGA0069549	13:37851945	13:37353675	13:23719928	3.87E-7	4.17E-3	20143	13:37872088-37876725	ENSSSCP00000012196	*STAB1*

[1]Chromosome and nucleotide position in the different genome versions

[2]P-raw value

[3]False Discovery Rate

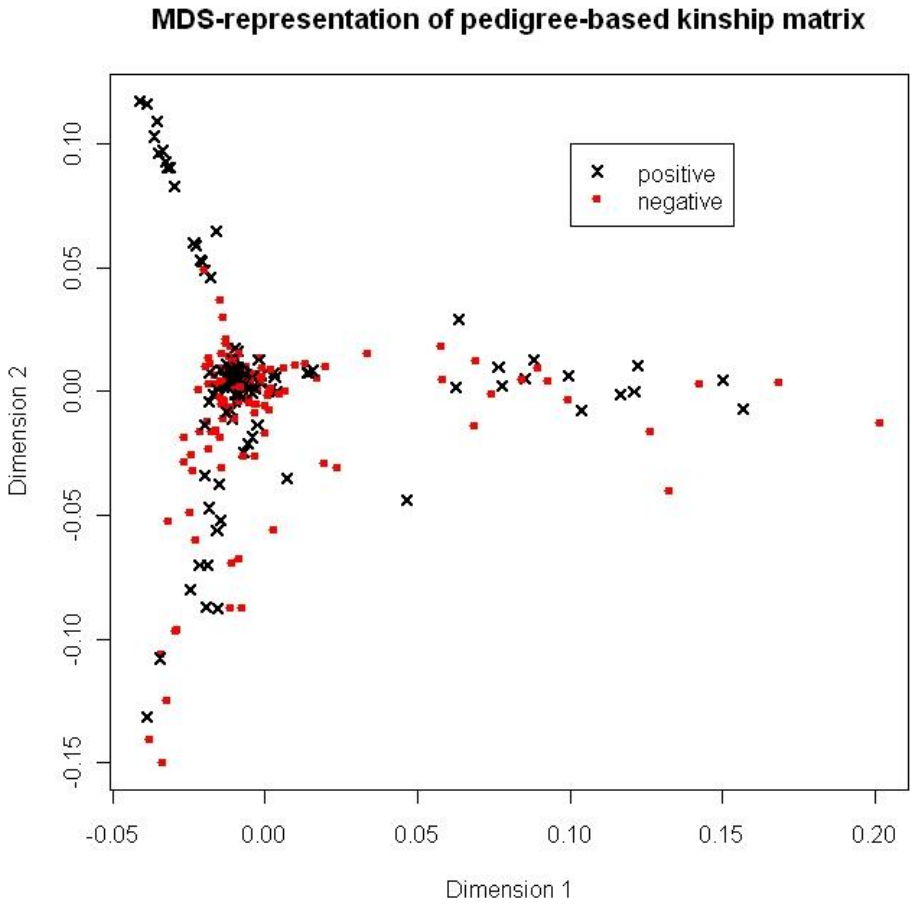
[4]Localization of the SNP in the corresponding gene or distance from the closest gene (in bp)

Table 2 - Results of the Gene Ontology (GO) analysis including closest genes to SNP with P<5.0E-05

Go sub-ontology	**GO term accession**	**GO term description**	**Number of involved genes**	**Involved genes**	**DAVID P-value**
Biological Process	GO:0050767	regulation of neurogenesis	5	*ACTR3, LINGO1, IRX3, XRCC2, DLL1*	0.015
Biological Process	GO:0006928	cell motion	8	*ACTR3, CXCR4, SPOCK1, SCNN1G, IL12B, APBB2, ELMO1, CTNNA2*	0.018
Biological Process	GO:0051960	regulation of nervous system development	5	*ACTR3, LINGO1, IRX3, XRCC2, DLL1*	0.024
Biological Process	GO:0060284	regulation of cell development	5	*ACTR3, LINGO1, IRX3, XRCC2, DLL1*	0.029
Biological Process	GO:0030182	neuron differentiation	7	*LINGO1, CXCR4, MTPN, APBB2, OLFM3, NTM, CTNNA2*	0.038
Biological Process	GO:0045664	regulation of neuron differentiation	4	*ACTR3, LINGO1, IRX3, DLL1*	0.040
Biological Process	GO:0048666	neuron development	6	*LINGO1, CXCR4, APBB2, OLFM3, NTM, CTNNA2*	0.044
Molecular Function	GO:0031420	alkali metal ion binding	5	*KCNK9, KCNT2, ATP1B3, SLC22A4, SCNN1G*	0.031
Molecular Function	GO:0000166	nucleotide binding	19	*RBM24, XRCC2, SUCLG2, PKN2, ABCB1, ACTR3, MAP3K5, KCNT2, HIPK1, ASCC3, PDE1C, CELF4, SLC22A4, CELF2, DPYD, RAB38, DOCK10, ARL4C, MOCS1*	0.038
Cellular Component	GO:0043005	neuron projection	7	*NUMA1, CXCR4, MTPN, PARK2, APBB2, DTNBP1, CTNNA2*	0.009
Cellular Component	GO:0031252	cell leading edge	4	*ACTR3, CXCR4, APBB2, CTNNA2*	0.037

Additional files

Figure S1:
Two-dimensional graphical representation of relatedness among animals based on a multidimensional scaling representation of pedigree-based kinship matrix. Different symbols are used to denote pigs with positive or negative backfat thickness estimated breeding values.



Tables S1: Suggestively significant SNPs (5.0E-07< P≤5.0E-05), their chromosome positions and their closest genes in Sscrofa10.2 (Pre-Ensembl). Notes are the same as those reported for Table1

CHR:POSITION (Sscrofa10)	CHR:POSITION (Sscrofa9.2)	P	FDR	SNP position/distance	Closest gene coordinates	Closest gene protein in Pre-ensembl	Gene Symbol
2:33084396	2:16685404	5.78E-7	0.004166020	24212	2:32547222-32721004	ENSSSCP00000014159	MPPED2
1:29091230	1:22692347	6.28E-7	0.004166020	44184	1:30747588-30883017	ENSSSCP00000004497	MAP3K5
0:0	0:0	6.69E-7	0.004166020	1636608	10:66363855-66585202	ENSSSCP00000011861	CELF2
0:0	0:0	9.01E-7	0.004166020	176391	8:122086037-122189190	ENSSSCP00000009755	LEF1
2:144447032	0:0	9.01E-7	0.004166020	168569	2:144721630-145255083	ENSP00000378401	SPOCK1
0:0	0:0	9.01E-7	0.004166020	Intron 24	9:102530763-102740603	ENSSSCP00000016317	ABCB1
18:14473279	0:0	9.70E-7	0.004166020	23921	18:14099822-14138480	ENSSSCP00000017509	MTPN
13:37839628	0:0	1.01E-6	0.004166020	42000	13:38213820-38288168	ENSP00000378235	SFMBT1
5:7170363	5:4993938	1.25E-6	0.004733860	Intron 2	5:7209683-7425562	ENSSSCP00000000125	GALR3
6:88826611	6:84170972	1.64E-6	0.005430940	70103	6:89419071-89434147	ENSP00000361834	COL9A2
6:6655662	6:787265	1.65E-6	0.005430940	Intron 7	6:7185321-7278279	ENSSSCP00000002903	CMIP
7:40806202	7:39630963	2.01E-6	0.005579550	Intron 1	7:40420291-40443740	ENSP00000362282	MOCS1
0:0	0:0	2.05E-6	0.005579550				
0:0	0:0	2.05E-6	0.005579550				
9:3876584	9:1636263	2.20E-6	0.005579550	2716	9:4381985-4382947	ENSSSCP00000015580	OR52N2
1:19076324	0:0	2.26E-6	0.005579550	396553	1:20543890-20705746	ENSP00000321826	STXBP5
8:34071238	0:0	2.59E-6	0.006084880	26261	8:33236813-33423156	ENSP00000427211	APBB2
6:140135176	6:38749591	3.01E-6	0.006241370	115725	6:140421919-140654461	ENSSSCP00000004142	FGGY
X:5842394	X:1957529	3.05E-6	0.006241370	319535	X:57674949-57680848	ENSSSCP00000013164	ZC4H2
9:4132852	9:1885017	3.33E-6	0.006241370	3530	9:4689819-4695785	ENSSSCP00000015601	TRIM6
7:27941247	7:27559898	3.55E-6	0.006241370	Intron 13	7:27771735-27791629	ENSSSCP00000001520	
4:1622309	0:0	3.70E-6	0.006241370	3814	4:16999359-17046727	ENSP00000287380	WDR67
0:0	0:0	3.70E-6	0.006241370				
17:38522336	17:30701716	4.06E-6	0.006241370	26280	17:38009364-38021616	ENSSSCP00000007647	
7:48384424	7:48680546	4.09E-6	0.006241370	70532	7:48466353-48497373	ENSSSCP00000001883	GPR110
9:6883156	9:3538170	4.31E-6	0.006241370	Intron 3	9:7354382-7382403	ENSSSCP00000015721	NUMA1
2:133309779	0:0	4.34E-6	0.006241370	635700	2:134893490-134898449	ENSSSCP00000015151	xxx
9:3801354	9:1561033	4.42E-6	0.006241370	9318	9:4300172-4301116	ENSSSCP00000015573	OR56A3

4:17175674	4:15308823	4.90E-6	0.006241370	473939	4:17236913-17252340	ENSP00000259512	DERL1
6:92512862	6:19249624	4.93E-6	0.006241370	463598	6:92352797-92436806	ENSSSCP00000003990	CCDC165
17:33362084	17:26260352	5.01E-6	0.006241370	87698	17:31797648-31806811	ENSSSCP00000007564	NAA20
1:45272366	0:0	5.26E-6	0.006241370	105862	1:47894988-48057459	ENSSSCP00000004583	MAN1A1
9:129569427	0:0	5.26E-6	0.006241370	377318	9:130733849-130819218	ENSSSCP00000016450	FAM5B
7:62574826	7:62259181	5.29E-6	0.006241370	21854	7:62377787-62379628	ENSSSCP00000002047	LINGO1
7:86475401	7:89262644	5.29E-6	0.006241370	115821	7:87588830-87598695	ENSSSCP00000002446	ARRDC4
1:5966847	0:0	5.48E-6	0.006241370	210444	1:7197052-8260240	ENSP00000355865	PARK2
4:135547432	4:124901359	5.57E-6	0.006241370	494069	4:139808249-139885772	ENSSSCP00000007385	PKN2
4:112488203	4:104404674	5.65E-6	0.006241370	8020	4:116713231-116742444	ENSSSCP00000007210	HIPK1
6:100615252	6:21686112	5.66E-6	0.006241370	Intron 3	6:100634610-100726475	ENSSSCP00000004012	CABLES1
9:85510638	0:0	5.69E-6	0.006241370	7196	9:85222169-85427470	ENSSSCP00000016269	COL28A1
3:66806262	0:0	5.72E-6	0.006241370	205347	3:67119262-67171441	ENSSSCP00000008813	CTNNA2
0:0	9:32928624	5.86E-6	0.006241370	70099	9:65008422-65418861	ENSSSCP00000016176	NTM
6:133887826	0:0	5.95E-6	0.006241370	185116	6:134871060-134913076	B3TNN4	PDE4B
4:103716327	4:96456312	6.07E-6	0.006241370	2911	4:107721372-107756757	ENSSSCP00000007091	GOLPH3L
13:122657370	13:71567378	6.98E-6	0.007033850	830315	13:124263010-124301092	ENSSSCP00000012530	TBL1XR1
4:110056273	4:101855616	7.14E-6	0.007054590	Intron 8	4:114119961-114213903	ENSSSCP00000007185	IGSF3
0:0	15:88200074	7.71E-6	0.007311480	Intron 7	15:139951470-140112107	ENSSSCP00000017201	DOCK10
4:110037486	4:101836829	7.74E-6	0.007311480	Intron 6	4:114119961-114213903	ENSSSCP00000007185	IGSF3
0:0	9:12097010	7.88E-6	0.007311480	5925	9:26044315-26060761	NP_001106917.1	CHORDC1
15:22820592	15:18937717	7.99E-6	0.007311480	401999	15:22725584-22786931	ENSP00000263238	ACTR3
16:84950193	16:50607024	9.09E-6	0.008165940	Intron 2	16:85607485-85996578	ENSSSCP00000018130	MRPL36
0:0	1:160101834	1.02E-5	0.009027780	51990	1:202935182-203018563	ENSSSCP00000005421	DDHD
16:68967125	16:40846554	1.16E-5	0.010080550	70256	16:69953974-69964216	ENSSSCP00000018057	IL12B
1:159037158	1:132208765	1.24E-5	0.010506940	388433	1:168992153-168993757	ENSSSCP00000005249	SOCS6
0:0	6:13199918	1.26E-5	0.010506940	92575	6:70488620-70538497	ENSP00000355060	ARHGEF10L
0:0	18:3519750	1.33E-5	0.010940450	159381	18:4937374-5046258	ENSSSCP00000017407	XRCC2
0:0	0:0	1.43E-5	0.011562550	654529	9:88556137-88562911	NP_001090937.1	NDUFA4
7:94085855	7:96295118	1.52E-5	0.012083440	31324	7:95175083-95190851	ENSSSCP00000002470	PLEKHG3
3:77487801	3:29947330	1.59E-5	0.012372110	345967	3:77679432-77683803	ENSSSCP00000008917	C1D
3:22987343	0:0	1.60E-5	0.012372110	3735	3:23445919-23476070	ENSSSCP00000008365	SCNN1G
6:63308009	6:12205604	1.67E-5	0.012457580	51348	6:65324504-65333801	ENSSSCP00000003702	LOC100523441
0:0	16:49935458	1.67E-5	0.012457580	308915	16:84598910-84600254	ENSSSCP00000018128	
1:507934	1:79763	1.74E-5	0.012457580	276135	1:281894-289492	ENSP00000355718	DLL1
0:0	1:187582635	1.74E-5	0.012457580	178986	1:238795293-238950756	ENSSSCP00000005584	PTPRD

6:62987043	6:12099404	1.74E-5	0.012457580	10559	6:65091662-65093997	ENSSSCP00000003694	
4:135647790	4:124999317	1.78E-5	0.012543920	592028	4:139808249-139885772	ENSSSCP00000007385	PKN2
4:2166585	4:1093405	1.84E-5	0.012784300	Intron 1	4:2362701-2385405	ENSSSCP00000006345	EIF2C2
8:120086170	8:55541534	1.95E-5	0.013407070	131407	8:121603244-121621973	ENSSSCP00000009750	AGXT2L1
1:205181858	1:171970180	2.05E-5	0.013661490	Intron 5	1:218678153-218721352	ENSSSCP00000005511	C9orf82
3:8609388	3:1421124	2.05E-5	0.013661490	49889	3:8461662-8832324	ENSP00000318234	EMID2
0:0	0:0	2.08E-5	0.013675650	17410	13:53693142-54030490	ENSP00000419325	SUCLG2
9:3443230	9:1405722	2.12E-5	0.013805320	32123	9:4067145-4077257	ENSSSCP00000015562	FAM160A2
6:118161553	6:28731790	2.16E-5	0.013849160	5811	6:119141320-119158448	ENSSSCP00000004062	CYP4A22
6:101630451	6:22306965	2.28E-5	0.014329510	Intron 13	6:101665088-101967594	ENSSSCP00000004019	OSBPL1A
2:71529148	2:33619289	2.34E-5	0.014329510	53137	2:60989514-61018267	ENSSSCP00000014730	AP1M1
5:62745013	0:0	2.36E-5	0.014329510	44102	5:64512420-64630474	ENSP00000394008	
9:16279686	9:8293530	2.39E-5	0.014329510	674924	9:16491108-16491935	ENSSSCP00000012445	ATP1B3
7:101399631	7:103012134	2.41E-5	0.014329510	7309	7:101881462-102041577	ENSSSCP00000002531	DPF3
7:103079877	7:105149510	2.41E-5	0.014329510	29403	7:103608527-103800694	ENSSSCP00000002568	FCF1
3:134753808	3:62456911	2.51E-5	0.014773180	745117	3:135803385-135804164	ENSSSCP00000009215	ID2
0:0	11:50323734	2.57E-5	0.014773180	1488702	11:69611870-69632132	ENSSSCP00000010134	GPG6
9:10682988	9:5175148	2.57E-5	0.014773180	Intron 1	9:106872890-107132218	ENSP00000303212	SEMA3E
8:112863525	0:0	2.66E-5	0.014935630	484540	8:115058036-115074694	ENSSSCP00000009727	C4ORF21
5:2614484	5:1709451	2.69E-5	0.014935630	53502	5:2668477-2802118	ENSSSCP00000000029	
15:87858348	15:96862419	2.69E-5	0.014935630	37532	15:91529100-91635557	ENSP00000249442	MTX2
15:16434289	15:9274593	2.87E-5	0.015766960	1016630	15:18125415-18128466	ENSSSCP00000016630	CXCR4
4:38407453	4:35314280	3.00E-5	0.016275000	Intron 11	4:40007864-40389295	ENSP00000351346	VPS13B
10:24904892	10:40777707	3.13E-5	0.016599120	Intron 20	10:23898728-24100299	ENSP00000294725	KCNT2
5:2363355	5:1569232	3.15E-5	0.016599120	Intron 6	5:2584675-2648147	ENSSSCP00000000028	
2:65875152	2:31528235	3.16E-5	0.016599120	481	2:66652122-66665807	ENSP00000407182	TNPO2
9:24085776	9:11844447	3.27E-5	0.017010480	361423	9:23798576-23862377	ENSP00000243662	RAB38
0:0	0:0	3.33E-5	0.017130220	60643	1:251421046-251461799	ENSP00000366243	TMEM2
7:13276847	7:13211824	3.45E-5	0.017401510	92990	7:13741795-13752147	ENSP00000368341	RBM24
6:101115892	6:21946783	3.45E-5	0.017401510	102479	6:101061150-101103291	ENSP00000323387	ANKRD29
7:11267787	7:10833754	3.50E-5	0.017469850	755607	7:12191634-12309233	ENSP00000341680	DTNBP1
12:44714600	12:21841563	3.55E-5	0.017521270	3278	12:44259718-44287870	ENSSSCP00000018786	PSMD11
5:7711829	5:5536068	3.61E-5	0.017526790	18843	5:8010270-8012726	ENSP00000300147	ELFN2
16:37690744	16:20163875	3.64E-5	0.017526790	22141	16:37561961-37610054	ENSSSCP00000017924	IL31RA
14:147276707	14:142677189	3.66E-5	0.017526790	92379	14:147439203-147531465	ENSSSCP00000011457	C10orf90-like
6:28750	0:0	3.70E-5	0.017526790	168932	6:182073-196415	ENSSSCP00000002843	LOC100522374

4:2830209	4:1622937	3.75E-5	0.017526790	58382	4:3049220-3122950	ENSSSCP00000006346	KCNK9
1:71880188	1:58249040	3.83E-5	0.017526790	Intron 1	1:76387794-76423833	ENSSSCP00000004703	ASCC3
2:139863716	2:71795836	3.83E-5	0.017526790	Intron 8	2:139999368-140042942	ENSSSCP00000015186	SLC22A4
4:126728020	4:117054271	3.83E-5	0.017526790	291381	4:131668429-132548759	ENSP00000359211	DPYD
9:20862575	9:9731446	3.90E-5	0.017675560				
8:103459010	8:106088938	3.97E-5	0.017675560	842570	8:103990222-104118478	ENSSSCP00000009689	INTU
0:0	5:7848427	3.97E-5	0.017675560	1171356	5:9957810-10078107	ENSSSCP00000000159	ISX
6:113289871	6:27097026	4.01E-5	0.017675560	146512	6:113847688-114192947	ENSP00000355089	CELF4
4:123556548	4:114273898	4.21E-5	0.018413120	94086	4:128209320-128249552	ENSSSCP00000007312	OLFM3
0:0	0:0	4.26E-5	0.018458120				
18:37872202	18:37323259	4.32E-5	0.018572150	2130540	18:40536822-40982997	ENSP00000312185	ELMO1
8:35824311	8:15983038	4.43E-5	0.018851840	21796	8:35416743-35531570	ENSP00000382670	GRXCR1
18:26693255	18:13385958	4.48E-5	0.018914990	Intron 17	18:26590290-26641146	ENSP00000377040	AASS
6:27015877	6:6573204	4.66E-5	0.019511120	44341	6:28168509-28170872	ENSP00000331608	IRX3
0:0	0:0	4.95E-5	0.019983320				
15:144204155	0:0	4.95E-5	0.019983320	88060	15:149242879-149243454	ENSP00000375057	ARL4C
18:42499600	0:0	4.95E-5	0.019983320	59573	18:43218007-43461582	ENSP00000297161	BMPER
13:146978045	13:106088600	4.95E-5	0.019983320	97317	13:150430086-150614290	ENSP00000377370	IGSF11
7:60438297	7:60170948	4.98E-5	0.019983320	12415	7:60158035-60195946	ENSSSCP0000000202	WDR93

CHAPTER 2

A genomewide association study for average daily gain in Italian Large White pigs

[1]L. Fontanesi*,†,[2], G. Schiavo*,[3], G. Galimberti‡,[3], D. G. Calò‡,[3] and V. Russo*

*Department of Agricultural and Food Sciences, Division of Animal Sciences, University of Bologna, Viale Fanin 46, 40127 Bologna, Italy

†Centre for Genome Biology, University of Bologna, 40126 Bologna, Italy

‡Department of Statistical Sciences "Paolo Fortunati", University of Bologna, Via delle Belle Arti 41, 40126 Bologna, Italy

Published in Journal of Animal Science. 2014 Apr;92(4):1385-94. doi: 10.2527/jas.2013-7059. Epub 2014 Feb 10.

Abstract

Average daily gain is an important target trait in pig breeding programs. In this study we performed a genomewide association study for ADG in Italian Large White pigs using a selective genotyping approach. Two extreme and divergent groups of Italian Large White pigs (number 190 + 190) were selected among a population of about 10,000 performance tested gilts (EBV for ADG in the 2 groups were –30 ± 14 g and 81 ± 12 g, respectively) and genotyped with the Illumina PorcineSNP60 BeadChip. Association analysis was performed treating the pigs of the 2 extreme groups as cases and controls after correction for family-based stratification. A total of 127 SNP resulted significantly associated with ADG (*P* nominal value [*P*raw] < 2.0 × 10–7, *P* < 0.01 Bonferroni corrected [*P*Bonferroni] < 0.01, false discovery rate < 7.76 × 10–5). Another 102 SNP were suggestively associated with the target trait (*P*raw between 2.0 × 10–7and 2.02 × 10–

6, *P*Bonferroni < 0.10, false discovery rate < 4.19 × 10–4). These SNP were located on all autosomes and on porcine chromosome (SSC) X. The largest number of SNP within this list was on SSC5 (*n* = 42), SSC7 (34), SSC6 (30), SSC4 (23), and SSC16 (16). These chromosomes were richer in significant or suggestively significant markers than expected (*P* < 0.001). A quite high number of these SNP (*n* = 23) were associated with backfat thickness in a previous genomewide association study performed in the same pig population, confirming the negative correlation between the 2 traits. Two or more SNP targeted the same gene: *IGSF3* and *HS2ST1* (SSC4), *OTOGL* (SSC5), *FTO* region (SSC6), and *MYLK4* and *MCUR1* (SSC7). Other regions that were associated with ADG in previous candidate gene studies (e.g., *MC4R* on SSC1, *IGF2* and *LDHA* on SSC2, *MUC4* on SSC13) 1) included markers with *P*raw < 0.01 that, however, did not pass the stringent threshold of significance adopted in this study or 2) could not be tested because not assigned to the Sscrofa10.2 genome version. Functional annotation of the significant regions using Gene Ontology suggested that many and complex processes at different levels are involved in affecting ADG, indicating the complexity of the genetic factors controlling this ultimate phenotype. The obtained results may contribute to understand the genetic mechanisms determining ADG that could open new perspectives to improve selection efficiency in this breed.

Introduction

Growth rate, measured at different growth stages, is an important objective in pig breeding programs as it is directly related to economic advantages. Therefore, measures of this phenotype, such as ADG, are usually included as target traits in selection programs in purebred and commercial pig lines.

Quantitative trait loci for growth performances and related traits have been reported on almost all porcine chromosomes (Hu et al., 2013), suggesting that growth efficiency is a complex trait determined by a large number of loci. In addition, candidate genes have been associated with ADG

in different pig populations, including Italian heavy pig breeds (e.g., Fontanesi et al., 2010b, 2011, 2012a,b).

Several of our previous association studies between DNA markers and production traits in these breeds were based on a selective genotyping strategy in which only the most extreme animals for the target trait, selected within a large performance tested population, were genotyped. This method provided a cost-effective and powerful experimental design (Darvasi and Soller, 1992) to identify gene associated with economically important traits in genetically evaluated pigs of nucleus herds (Fontanesi et al., 2009, 2012a,b,c,d).

The recent development of a high throughput commercial genotyping platform in pigs (Porcine-SNP60 Genotyping BeadChip, Illumina inc. San Diego, CA; Ramos et al., 2009) that can analyze more than 60,000 SNP throughout the pig genome now enables us to perform genomewide association studies (GWAS), improving efficiency in detecting genome regions affecting production traits. We already performed a GWAS in Italian Large White pigs for backfat thickness (BFT) and identified novel chromosome regions affecting fat deposition (Fontanesi et al., 2012d).

In this work we performed a GWAS for ADG in Italian Large White using a selective genotyping approach and identified SNP associated with this trait adding information about the genetic complexity affecting growth performances in pigs.

All animals used in this study were kept according to Italian and European legislation for pig production and all procedures described were in compliance with national and European Union regulations for animal care and slaughtering.

Animals

The association study was conducted following a selective genotyping approach, as already described (Fontanesi et al., 2012c,d). Briefly, 2 extreme and divergent groups of Italian Large White gilts, identified within a population of about 10,000 pigs performance tested in the period

1996 through 2009, were used in this study. These animals were included in the national selection program of the Italian Large White breed. This program is based on triplets of siblings from the same litter, 2 females and 1 castrated male that are individually performance tested at the Central Test Station of the National Pig Breeder Association for the genetic evaluation of a boar from the same litter (sib testing). This population is virtually free from the *RYR1* c.1843T allele (Fontanesi et al., 2008, 2012c). Performance evaluation starts when the pigs are 30 to 45 d of age and it ends when the animals reach 155 ± 5 kg live weight. The nutritive level is quasi ad libitum, meaning that about 60% of the pigs are able to ingest the entire supplied ration (Fontanesi et al., 2010b). During the performance test period, body weight of the pigs is measured every 15 d after fasting, and then daily gain is calculated using body weight regress on the repeated test day. At the end of test, animals are transported to a commercial abattoir where they are slaughtered following standard procedures. The extreme and divergent gilts were chosen according to their relatedness and their EBV for ADG: 1) all gilts, among a population of about 10,000 performance tested gilts, were ranked according to their EBV for ADG, 2) among the animals related at 2-generation levels, only the most extreme gilt (with most positive or most negative EBV for ADG) was selected, 3) this procedure selected 190 gilts with the most negative and 190 gilts with the most positive EBV not related at 2-generation levels, and 4) average EBV for ADG in the negative and positive selected groups of pigs were –30 ± 14 g (mean ± SD; minimum: –76 g; maximum: –9 g) and 81 ± 12 g (minimum: 69 g; maximum: 129 g), respectively.

Genotyping

Blood was collected from all performance tested animals and then dried. Dried blood of chosen gilts was used to extract genomic DNA applying standard protocols. After quality control, 375 animals were used for genotyping using the Illumina PorcineSNP60 BeadChip (Ramos et al., 2009) according to the manufacturer's protocol.

Data Analyses

Estimated breeding values for ADG were calculated in the whole performance tested population in 2010 using a BLUP-Multiple Trait-Animal Model. The model included the fixed effect of sex (considering the triplets of pigs from the same litter), batch on trial, inbreeding coefficient of the animal, interaction of sex × age at slaughtering, and date of slaughtering and random effect of litter and animal. The following criteria were used to filter animals and SNP before association analysis: call rate > 0.9 (both at the animal and SNP level) and minor allele frequency (MAF) > 0.05. Association analysis was performed treating the pigs of the 2 extreme groups as cases and controls, using the 2 groups of animals with divergent ADG EBV. To detect and correct for possible genetic substructure in the experimental design adopted (Fontanesi et al., 2012d), association tests were performed according to the method for single marker association proposed by Price et al. (2006). For the ***n*** animals involved in the study, the ***n*** × ***n*** kinship matrix K was estimated starting from available pedigree information. Classical multidimensional scaling was applied on 0.5 – K (which acts as a pairwise distance matrix) to identify a number ***D*** << ***n*** of first axes describing as much genetic difference among animals as possible. Let $\boldsymbol{c}_1,\ldots,\boldsymbol{cD}$ denote these ***D*** axes of genetic variation. Adjustment for possible family-based stratification was performed by regressing the genotype at the ***i***th SNP and the phenotype onto the ***D*** continuous axes and taking the corresponding regression residuals as corrected genotypes and phenotypes, respectively. Namely, let ***gij*** and ***pj*** be the genotype at SNP ***i*** (***gij*** = 0, 1 or 2) and the phenotype of animal ***j***, respectively. The adjustment was performed on genotypes and phenotype according to the following formulas:

$$g_{ij}^{*} = g_{ij} - \hat{\beta}_{1i}c_{1j} - .... - \hat{\beta}_{Di}c_{Dj},$$

and

$$p_{j}^{*} = p_{j} - \hat{\gamma}_{1}c_{1j} - .... - \hat{\gamma}_{D}c_{Dj},$$

in which *cdj* is the score of animal *j* along the *d*th axis of genetic variation and $\hat{\beta}_{di}$ and $\hat{\gamma}_{d}$ are the corresponding estimated partial regression coefficients. These coefficients were obtained using multiple regression models for predicting the *i*th genotype and the phenotype, respectively, on the basis of the *D* axes.

The association test statistic is computed as $(n-D-1)r_i^2$, in which

$$r_i^2 = \left(\sum_{j=1}^{n} p_j^* g_{ij}^* \right)^2 / \sum_{j=1}^{n} \left(p_j^* \right)^2 \sum_{j=1}^{n} \left(g_{ij}^* \right)^2$$

is the squared correlation coefficient between the ***i***th adjusted genotype and the adjusted phenotype. As remarked by Price et al. (2006), this statistic is a generalization of the Armitage trend statistic usually adopted for categorical genotypes and phenotypes.

Supplementary Fig. 1 shows enrichment of low *P* nominal values beyond what would be expected under a uniform distribution.

Single nucleotide polymorphisms with *P* nominal value (*P*raw) < 2.0×10^{-7}(***P*** < 0.01 Bonferroni corrected [*P*Bonferroni]) were considered significantly associated with ADG. The corresponding false discovery rate (FDR) was equal to 7.76×10^{-5} (Benjamini and Hochberg, 1995). Single nucleotide polymorphisms with P_{raw} values between 2.0×10^{-7} and 2.02×10^{-6}($P_{Bonferroni}$ = 0.10, FDR < 4.19×10^{-4}) were considered suggestively associated with the target trait. For each chromosome, the expected proportion of SNP with $P_{Bonferroni}$ < 0.10 was computed under the assumption of uniform distribution from the informative SNP over the chromosome. This proportion was compared to the proportion of significant or suggestively significant markers actually observed on the same chromosome. Proportion of phenotype variance explained by each significant SNP was not calculated as the selective genotyping design would produce a biased estimation.

All analyses were performed in R (R Development Core Team, 2013). Package kinship2 (Therneau et al., 2012) was used to compute the pedigree-based kinship matrix; package GenABEL (Aulchenko et al., 2007) was used to perform association tests.

Bioinformatics Analyses

Mapping of the PorcineSNP60 BeadChip SNP was obtained on the Sscrofa10.2 genome assembly as previously described (Fontanesi et al., 2012c,d). Significant unassigned SNP in the Sscrofa10.2 were mapped on the Sscrofa9.2 genome version. Identification of the closest genes to SNP with $\boldsymbol{P}_{\text{Bonferroni}} < 0.10$ was obtained using Ensembl annotation of Sscrofa10.2 genome version (July 2013) and Biomart (www.ensembl.org/biomart/martview/). For subsequent analyses, a window of 100 kb in 5′ and 100 kb in 3′ of the SNP in this list was used to retrieve additional genes close to the significant or suggestively significant markers. This window of 0.2 Mb can be considered a conservative approach that can be easily extended using coordinate systems reported in this study. Starting from the corresponding protein sequences retrieved from these databases, the corresponding gene symbols were extracted from National Center for Biotechnology Information gene section (www.ncbi.nlm.nih.gov/gene/) and/or Uniprot (www.uniprot.org/) databases (July 2013). Gene annotation was verified by basic local alignment search tool (BLAST) analysis (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Gene Ontology (GO) analysis was performed using DAVID Bioinformatics Resources 6.7 (http://david.abcc.ncifcrf.gov/; Huang et al., 2009) using information about the closest gene.

Results and Discussion

Genotyping Data

Of the 375 genotyped animals, 5 were excluded from further analysis because their call rate was <0.90. After filtering the 62,163 SNP of the Illumina PorcineSNP60 BeadChip (750 SNP had call rate <0.90 and 12,000 SNP had MAF <0.05), a total of 49,413 SNP were used for subsequent

analyses. These SNP were remapped on the Sscrofa10.2 genome version: 42,885 were assigned to assembled porcine chromosomes in only 1 position, 6,528 were assigned to unassembled scaffolds, and 2,938 were not assigned (or were assigned to more than 1 position and were not considered as uniquely mapped).

Genome Scan Results

Figure 1 reports a Manhattan plot showing significant and suggestively significant SNP. A total of 127 SNP resulted significantly associated with ADG (Table 1; Supplementary Table 1).

Another 102 SNP resulted suggestively associated with the target trait (Supplementary Table 1). Among the 229 SNP associated or suggestively associated with ADG, only 8 were not assigned to any chromosome and 4 were placed in unassigned scaffolds of the Sscrofa10.2 genome version. Mapped SNP were located on all autosomes and on porcine chromosome (SSC) X. The largest number of the SNP within this list was on SSC5 (number of SNP with $P_{\text{Bonferroni}} < 0.10 = 42$ and number of SNP with $P_{\text{Bonferroni}} < 0.01 = 31$), SSC7 (number of SNP considering the 2 thresholds were 34 and 13), SSC6 (30 and 20 SNP, respectively), SSC4 (23 and 14 SNP, respectively) and SSC16 (16 and 13 SNP, respectively; Supplementary Table 1). These chromosomes were richer in significant or suggestively significant markers than expected ($P < 0.001$; SSC4: expected proportion = 0.067, observed = 0.111; SSC5: expected = 0.045, observed = 0.203; SSC6: expected = 0.060, observed = 0.145; SSC7: expected = 0.063, observed = 0.164; and SSC16: expected = 0.035, observed = 0.077). A large number of QTL for ADG and growth performances have been already reported on these chromosomes. For example, on July 2013 (release 20) the PigQTLdb (www.animalgenome.org/cgi-bin/QTLdb/SS/index) reports 45, 9, 25, 36, and 3 QTL for ADG on SSC4, SSC5, SSC6, SSC7, and SSC16, respectively. *Sus scrofa* chromosome 4, SSC6, and SSC7 were among the richest chromosomes of significant SNP in our previous GWAS for BFT in Italian Large White pigs (Fontanesi et al., 2012d), suggesting that several regions identified in our previous GWAS and the current study might contain QTL with pleiotropic effects on both traits. Twenty-

three of the 229 SNP identified for ADG in the current study were also previously reported to be significant or suggestively significant for BFT in the same pig population (Fontanesi et al., 2012d; Table 1; Supplementary Table 1). Moreover, several other SNP associated with ADG in this study are close to the markers associated with BFT studied before, and the direction of the effects was opposite (Fontanesi et al., 2012d). This could be expected as EBV for ADG and BFT in Italian Large White pigs are negatively correlated ($\boldsymbol{r}^2 = -0.44$; Fontanesi et al., 2013).

The most significant SNP (ALGA0030787, $P_{raw} = 3.19 \times 10^{-11}$) was not mapped in Sscrofa10.2 even if it was assigned to SSC5 (position 11032453) in Sscrofa9.2. The second and third most significant SNP (ALGA0004718, $P_{raw} = 1.04 \times 10^{-10}$, and ALGA0004837, $P_{raw} = 1.17 \times 10^{-10}$) were localized on SSC1. Other highly significant SNP ($P_{raw} < 1.00 \times 10^{-9}$) were identified on SSC4 (M1GA0006302, M1GA0006343, and M1GA0006613), SSC6 (ALGA0035254), and SSC16 (M1GA0021128; Table 1).

Several chromosome regions included 3 or more SNP ($P_{Bonferroni} < 0.10$) separately to each other by less than 1.5 Mb (Supplementary Table 1). In particular, 2 regions with these features were identified on SSC4 (111.51–114.17 and 141.17–143.22 Mb), 5 on SSC5 (2.61–3.09, 8.03–9.29, 65.43–67.62, 72.59–74.37, and 105.35–109.06 Mb), 4 on SSC6 (26.20–30.90, 50.50–50.85, 91.14–92.90, and 100.66–101.77 Mb), 2 on SSC7 (7.54–11.61 and 128.67–130.84 Mb), and 1 on SSC16 (80.84–82.68 Mb), with significant SNP in the middle or close to these regions (including blocks with 2 closely spaced SNP) that might reflect the presence of different haploblocks (L. Fontanesi, personal communication).

Other regions with 2 closely spaced (less than 1.5 Mb) SNP with $P_{Bonferroni} < 0.10$ were localized on SSC1 and SSC13 (Supplementary Table 1).

Functional Annotation of Associated SNP

Twenty-eight SNP with $P_{Bonferroni} < 0.10$ were in intragenic regions of recognized genes in the Ensembl Sscrofa10.2 assembly (Table 1; Supplementary Table 1). For the remaining mapped

significant or suggestively significant SNP (number = 189), the distances from their closest genes ranged from 149 bp to 1.53 Mb (mean = 324.97 kb ± 22.33 kb, median = 22.48 kb).

Two or more SNP targeted the same gene. For example, 2 SNP (M1GA0006299, position 11,4151,582, and M1GA0006302, position 114,170,369) were located within or very close to the *immunoglobulin superfamily, member 3* (*IGSF3*) gene (Table 1) that is included in 1 of the regions with several significant or suggestively significant SNP on SSC4. This gene seems involved in immune cell regulation even if its function is not well characterized yet (Clark et al., 2001). Both markers were also associated with BFT in our previous study (Fontanesi et al., 2012d). Another 2 SNP on SSC4, included in another significant group of SNP of this chromosome (M1GA0006854, position 141,458,072, and M1GA0006869, position 141,552,372), were close and intragenic to the *heparan sulfate 2-O-sulfotransferase 1* (*HS2ST1*) gene, respectively. The *HS2ST1* gene encodes a member of the heparan sulfate biosynthetic enzyme family that transfers sulfate to the 2 position of the iduronic acid residue of heparan sulfate. This enzyme seems important in the signaling pathways involved in kidney formation and immunological functions (Muramatsu, 2000). Three SNP (M1GA0008164, position 106,046,784, DRGA0006447, position 107,303,857, and DRGA0006450, position 107,536,603), in 1 of the significant regions of SSC5, were upstream and downstream to the *otogelin-like* (*OTOGL*) gene. Mutations in this gene, which are mainly expressed in the inner ear of vertebrates during embryonic development, cause recessive deafness (Yariz et al., 2012). Its potential role on growth related metabolism or functions needs to be further investigated.

One of the most significant regions on SSC6 (26.20–30.90 Mb) might include the *fat mass and obesity associated* (*FTO*) gene that is associated with fat deposition traits in Italian Duroc, Italian Large White, and heavy pig commercial hybrids and feed conversion rate in Italian Large White (Fontanesi et al., 2009, 2010a; Fontanesi and Russo, 2013). To be precise, *FTO* position is available only on Sscrofa10.0 (27,697,754–28,086,339) as this gene is not assembled in Sscrofa10.2, but comparative mapping may confirm that its position on Sscrofa10.2 should be within the indicated

region of SSC6 in the latest assembly (data not shown). This region also includes a marker associated with BFT (M1GA0008432; Fontanesi et al., 2012d).

The *LOC100157526* (also identified as *MYLK4- putative myosin light chain kinase 3-like*) and the *mitochondrial calcium uniporter regulator 1*(*MCUR1* or *CCDC90A*) gene, both located on SSC7, were each identified with 1 upstream and 1 downstream close marker (Supplementary Table 1). As far as we know, *MYLK4* is not functionally characterized in any species, yet. *MCUR1* encodes a component of mitochondrial Ca^{2+} uptake that regulates cellular metabolism (Mallilankaraman et al., 2012).

Highly significant SNP were close to additional genes. The most significant SNP on SSC1 (ALGA0004718, associated with BFT, and ALGA0004837) were close to a novel pseudogene and the *5-hydroxytryptamine*(*serotonin*) *receptor 1B, G protein-coupled* (*HTR1B*) gene. *HTR1B* is highly expressed in the brain and is associated with several behavior and neurological related functional roles. In dairy cattle, it is involved in the homeostatic regulation of lactation (Collier et al., 2012). M1GA0006343, one of the most significant SNP (also associated with BFT in our previous GWAS; Fontanesi et al., 2012d), located on SSC4, was close to the *olfactomedin-like 3* (*OLFML3*) gene that in pig may affect prenatal skeletal muscle development (Zhao et al., 2012). Another highly significant SNP of SSC4 (M1GA0006613) is close to the *Rho GTPase activating protein 29* (*ARHGAP29*) gene that is involved in cell spreading and endothelial barrier function, important in chronic inflammation, atherosclerosis, and vascular leakage (Post et al., 2013). A highly significant SNP on SSC6 (ALGA0035254) is close to the *kin of IRRE like 2* (*Drosophila*; *KIRREL2*) gene that encodes a cell adhesion molecule regulating neural activity-dependent formation of precise axonal projections in the main olfactory system (Serizawa et al., 2006). The highly significant SNP identified on SSC16 (M1GA0021128) was close to the putative *tetratricopeptide repeat protein 1-like* gene (LOC100519063) whose function is not characterized yet.

Several other genes have been tagged by the remaining SNP (Table 1; Supplementary Table 1). Therefore, to have a global picture of the potential functional role of regions around associated or suggestively associated SNP with ADG in our pig population, we used GO information of their corresponding closest genes and reported GO terms enriched in this dataset (Supplementary Tables 2 and 3). Thirty-five GO terms (Supplementary Table 2) and 31 annotation clusters (Supplementary Table 3) were retrieved. Several GO terms were significantly enriched if we considered a *P* nominal value: the 4 most significant terms ($P_{raw} < 0.02$) were 0005509 (calcium ion binding), 0045934 (negative regulation of nucleobase, nucleoside, nucleotide, and nucleic acid metabolic process), 0051172 (negative regulation of nitrogen compound metabolic process), and 0051172 (negative regulation of nitrogen compound metabolic process), which might indicate a direct role of genes involved in several metabolic processes. However, none of the terms were significant after Bonferroni correction. This might suggest that, as expected, many different processes at different levels are involved in affecting this complex phenotype that expresses growth efficiency.

Comparison with other studies in pigs

A few other GWAS for ADG or correlated production traits have been performed in other pig populations/breeds. Becker et al. (2013)performed a GWAS in a relatively small population of Swiss Large White boars for a large number of EBV for different traits but no significant markers have been reported for ADG. Sahana et al. (2013) performed a GWAS for feed efficiency in a Duroc population using 2 statistical approaches and identified a total of 79 and 44 significant SNP, respectively. The most significant markers were located on SSC4, SSC7, SSC8, and SSC14. None of the significant SNP that Sahana et al. (2013)reported in their study was significant in our GWAS for ADG even if several markers they identified were close (<0.5 Mb) to significant markers we reported on different chromosomes (e.g., SSC4, SSC5, SSC7, SSC16, and SSC17). Not overlapping results between Sahana et al. (2013) and our GWAS could be due to the different populations used and by the fact that the considered traits, even if correlated, are not the same. Another GWAS on residual feed intake (RFI) and other related traits (including ADG) was performed in purebred

Yorkshires of 2 selection lines for RFI (high and low) using different approaches (Onteru et al., 2013). Significant SNP for RFI were identified on SSC3, SSC5, SSC6, SSC7, SSC13, and SSC14. Significant regions for ADG were reported in 15 SSC for a total of 44 chromosome positions. A few of these positions (SSC1, 167.00–168.00 Mb; SSC10, 15.00–16.00 Mb; SSC13, 36.00–37.00 Mb; and SSC16, 59.00–60.00 Mb) were very close to or included in the SNP list identified in this study. However, in general, results obtained by Onteru et al. (2013) poorly overlapped our results. This could be due to different experimental designs, incomplete power in the 2 studies, and/or differences between the investigated populations as already discussed comparing GWAS results for BFT in Italian Large White and other studies for the same trait (Fontanesi et al., 2012b).

The most significant region for ADG identified by Onteru et al. (2013) was on SSC1 and included the *MC4R* gene. A missense mutation in this gene (p.Asp298Asn) has been associated with several production traits including ADG in different pig populations (Kim et al., 2000) as well as in Italian Large White (Fontanesi et al., 2013). However, markers of the PorcineSNP60 BeadChip in the *MC4R* region were not significant in our GWAS for ADG, even if a few SNP had $P < 0.001$. This might indicate that, despite their effects, polymorphisms in the *MC4R* gene are not the most important markers to explain the variability of the target trait: their *P*-value could not pass the stringent threshold for significance we adopted in GWAS (Bonferroni corrected) that in single marker tests for a candidate gene is usually less stringent. Similar results for the *MC4R* region of SSC1 were obtained in our previous GWAS for BFT (Fontanesi et al., 2012d).

Other studies we performed in Italian Large White pigs showed a very strong effect of the *IGF2* intron3-g.3072G > A mutation (Van Laere et al., 2003) on ADG (Fontanesi et al., 2010b,c). Unfortunately, this gene is not assembled in Sscrofa10.2 and it was impossible to obtain a direct comparison with results obtained for SNP mapped on SSC2, which might be close to *IGF2*, included in the Illumina PorcineSNP60 BeadChip. We recently investigated another gene on SSC2 (*LDHA*) that was associated with ADG in the Italian Large White breed (Fontanesi et al., 2012b). This gene is localized at position 43898277 to 43909456 in the Sscrofa10.2 genome version.

Markers in this region were not significant after Bonferroni correction but several had $P_{raw} < 0.01$; for example, ASGA0010122 (position 43911957) had a $P_{raw} = 0.0023$. A similar situation can be seen for the *MUC4* g.8227C > G polymorphism that we recently investigated. This gene is located on SSC13 (position 143786443–143842402) that we recently investigated. The g.8227C > G SNP, associated with susceptibility to enterotoxigenic *Escherichia coli* K88 strains (locus *F4bcR*), was antagonistically associated with ADG in Italian Large White and in Italian Landrace (Fontanesi et al., 2012a). In the current GWAS for ADG a marker close to this gene (ALGA0072062, position 143866440) had a $P_{raw} = 2.24 \times 10^{-5}$ that, however, could not pass the threshold for significance.

It seems that chromosome regions with moderate effects could not be detected in our GWAS for a few reasons: 1) high stringency of the significant threshold needed to overcome the problem of multiple testing, 2) linkage disequilibrium structure of the investigated population that could not be captured completely by the Illumina PorcineSNP60 BeadChip(L. Fontanesi, personal communication), and 3) the incomplete power of the experimental design, despite the adopted selective genotyping strategy tended to maximize it (Darvasi and Soller, 1992).

Implications

In this study, the genomewide association between DNA markers and ADG was analyzed in the Italian heavy pig breed for the first time. The investigated trait is included in the selection index for the Italian Large White breed. The obtained results may contribute to understand the genetic mechanisms affecting ADG opening potential new perspectives to improve selection efficiency in this breed.

The study was designed to take advantage from the large number of pigs that are performance tested and genetically evaluated within the national selection program for this breed using a selective genotyping approach. Only extreme and divergent gilts for ADG EBV were genotyped to reduce the genotyping cost without losing much power (Darvasi and Soller, 1992;Van Gestel et al., 2000; Zhang et al., 2006). On the whole 229 SNP spread on all autosomes and on SSCX were significant (number = 127) or suggestively significant (number = 102). This large number of

identified SNP might indicate that, according to the classical definition of a quantitative trait, a large number of genes, each with a small or medium effect, contributes to explain the genetic variability of ADG. This study also missed detecting some chromosome regions that might have a low/moderate effect on the target trait or other important regions probably due to features of the genotyping tool and assembled genome available. It is interesting to point out that about 1/10 (23/229) of SNP identified in this study were also associated with BFT in our previous GWAS (Fontanesi et al., 2012d), according to the high negative correlation between the 2 EBV in the Italian Large White population. These results might indirectly provide evidence on the correctness of the statistical approaches and the efficiency of the experimental designs we used in the 2 GWAS. Finally, the large number of genes and biological processes that should be involved in defining ADG indicates the complexity of the genetic factors affecting this ultimate phenotype. To better understand the biological mechanisms determining growth efficiency in pigs it will be important to dissect this phenotype into several intermediate and internal phenotypes.

Footnotes

We thank Associazione Nazionale Allevatori Suini (ANAS) for providing samples and data. We are grateful to Dr. Luca Buttazzoni (CRA, Rome) for advices and data and Dr. Emilio Scotti (University of Bologna) for technical assistance. This work was funded by the Italian MiPAAF (INNOVAGEN project).

Received August 21, 2013.

Accepted January 17, 2014.

LITERATURE CITED

↲ Aulchenko Y. S., Ripke S., Isaacs A., van Duijn C. M.. 2007. GenABEL: An R library for genome-wide association analysis. Bioinformatics 23:1294–1296. Abstract/FREE Full Text

↲ Becker D., Wimmers K., Luther H., Hofer A., Leeb T.. 2013. A genome-wide association study to detect QTL for commercially important traits in Swiss Large White boars. PLoS ONE 8:e55951. A-LinkCrossRefMedlineGoogle Scholar

↲ Benjamini Y., Hochberg Y.. 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. J. R. Stat. Soc., B 57:289–300. A-LinkGoogle Scholar

↲ Clark K. L., Zeng Z., Langford A. L., Bowen S. M., Todd S. C.. 2001. PGRL is a major CD81-associated protein on lymphocytes and distinguishes a new family of cell surface proteins. J. Immunol. 167:5115–5121. Abstract/FREE Full Text

↲ Collier R. J., Hernandez L. L., Horseman N. D.. 2012. Serotonin as a homeostatic regulator of lactation. Domest. Anim. Endocrinol. 43:161–170. A-LinkCrossRefMedlineGoogle Scholar

↲ Darvasi A., Soller M.. 1992. Selective genotyping for determination of linkage between a marker locus and a quantitative trait locus. Theor. Appl. Genet. 85:353–359. A-LinkMedlineGoogle Scholar

↲ Fontanesi L., Bertolini F., Dall'Olio S., Buttazzoni L., Gallo M., Russo V.. 2012a. Analysis of association between the MUC4 g.8227C > G polymorphism and production traits in Italian heavy pigs using a selective genotyping approach. Anim. Biotechnol. 23:147–155. A-LinkCrossRefMedlineGoogle Scholar

↲ Fontanesi L., Buttazzoni L., Galimberti G., Calò D. G., Scotti E., Russo V.. 2013. Association between melanocortin 4 receptor (MC4R) gene haplotypes and carcass and production traits in Italian Large White pigs evaluated with a selective genotyping approach. Livest. Sci. 157:48–56. A-LinkCrossRefGoogle Scholar

↲ Fontanesi L., Buttazzoni L., Scotti E., Russo V.. 2012b. Confirmation of the association between a single nucleotide polymorphism in the porcine LDHA gene and average daily gain and correlated traits in Italian Large White pigs. Anim. Genet. 43:649–650. A-LinkCrossRefMedlineGoogle Scholar

↲ Fontanesi L., Davoli R., Nanni Costa L., Beretti F., Scotti E., Tazzoli M., Tassone F., Colombo M., Buttazzoni L., Russo V.. 2008. Investigation of candidate genes for glycolytic potential of porcine skeletal muscle: Association with meat quality and production traits in Italian Large White pigs. Meat Sci. 80:780–787. A-LinkCrossRefMedlineGoogle Scholar

↲ Fontanesi L., Galimberti G., Calò D. G., Fronza R., Martelli P. L., Scotti E., Colombo M., Schiavo G., Casadio R., Buttazzoni L., Russo V.. 2012c. Identification and association analysis of several hundred single nucleotide polymorphisms within candidate genes for back fat thickness in Italian Large White pigs using a selective genotyping approach. J. Anim. Sci. 90:2450–2464. Abstract/FREE Full Text

↲ Fontanesi L., Russo V.. 2013. Nucleotide variability and haplotype heterogeneity at the porcine fat mass and obesity associated (FTO) gene. Anim. Genet. 44:96–100. A-LinkCrossRefMedlineGoogle Scholar

↲ Fontanesi L., Schiavo G., Galimberti G., Calò D. G., Scotti E., Martelli P. L., Buttazzoni L., Casadio R., Russo V.. 2012d. A genome wide association study for backfat thickness in Italian Large White pigs highlights new regions affecting fat deposition including neuronal genes. BMC Genomics 13:583. A-LinkCrossRefMedlineGoogle Scholar

↲ Fontanesi L., Scotti E., Buttazzoni L., Dall'Olio S., Bagnato A., Lo Fiego D. P., Davoli R., Russo V.. 2010a. Confirmed association between a single nucleotide polymorphism in the FTO gene and obesity-related traits in heavy pigs. Mol. Biol. Rep. 37:461–466. A-LinkCrossRefMedlineGoogle Scholar

↲ Fontanesi L., Scotti E., Buttazzoni L., Davoli R., Russo V.. 2009. The porcine fat mass and obesity associated (FTO) gene is associated with fat deposition in Italian Duroc pigs. Anim. Genet. 40:90–93. A-LinkCrossRefMedlineGoogle Scholar

↲ Fontanesi L., Scotti E., Speroni C., Buttazzoni L., Russo V.. 2011. A selective genotyping approach identifies single nucleotide polymorphisms in porcine chromosome 2 genes associated

with production and carcass traits in Italian heavy pigs. Ital. J. Anim. Sci. 10:e15. A-LinkGoogle Scholar

↲ Fontanesi L., Speroni C., Buttazzoni L., Scotti E., Dall'Olio S., Nanni Costa L., Davoli R., Russo V.. 2010b. The insulin-like growth factor 2 (IGF2) gene intron3-g.3072G > A polymorphism is not the only Sus scrofa chromosome 2p mutation affecting meat production and carcass traits in pigs: Evidence from the effects of a cathepsin D (CTSD) gene polymorphism. J. Anim. Sci. 88:2235–2245. Abstract/FREE Full Text

↲ Hu Z.-L., Park C. A., Wu X.-L., Reecy J. M.. 2013. Animal QTLdb: An improved database tool for livestock animal QTL/association data dissemination in the post-genome era. Nucleic Acids Res. 41:D871–D879. Abstract/FREE Full Text

↲ Huang D. W., Sherman B. T., Lempicki R. A.. 2009. Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 37:1–13. Abstract/FREE Full Text

↲ Kim K. S., Larsen N., Short T., Plastow G., Rothschild M. F.. 2000. A missense variant of the porcine melanocortin-4 receptor (MC4R) gene is associated with fatness, growth, and feed intake traits. Mamm. Genome 11:131–135. A-LinkGoogle Scholar

↲ Mallilankaraman K., Cárdenas C., Doonan P. J., Chandramoorthy H. C., Irrinki K. M., Golenár T., Csordás G., Madireddi P., Yang J., Müller M., Miller R., Kolesar J. E., Molgó J., Kaufman B., Hajnóczky G., Foskett J. K., Madesh M.. 2012. MCUR1 is an essential component of mitochondrial Ca2+ uptake that regulates cellular metabolism. Nat. Cell Biol. 14:1336–1343. A-LinkCrossRefMedlineGoogle Scholar

↲ Muramatsu T. 2000. Essential roles of carbohydrate signals in development, immune response and tissue functions, as revealed by gene targeting. J. Biochem. 127:171–176. Abstract/FREE Full Text

↵ Onteru S. K., Gorbach D. M., Young J. M., Garrick D. J., Dekkers J. C., Rothschild M. F.. 2013. Whole genome association studies of residual feed intake and related traits in the pig. PLoS ONE 8:e61756. A-LinkCrossRefMedlineGoogle Scholar

↵ Post A., Pannekoek W. J., Ross S. H., Verlaan I., Brouwer P. M., Bos J. L.. 2013. Rasip1 mediates Rap1 regulation of Rho in endothelial barrier function through ArhGAP29. Proc. Natl. Acad. Sci. USA 110:11427–11432. Abstract/FREE Full Text

↵ Price A. L., Patterson N. J., Plenge R. M., Weinblatt M. E., Shadick N. A., Reich D.. 2006. Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet. 38:904–909. A-LinkCrossRefMedlineGoogle Scholar

↵ R Development Core Team. 2013. R: A language and environment for statistical computing. www.R-project.org/ (accessed 9 Nov. 2012).

↵ Ramos A. M., Crooijmans R. P., Affara N. A., Amaral A. J., Archibald A. L., Beever J. E., Bendixen C., Churcher C., Clark R., Dehais P., Hansen M. S., Hedegaard J., Hu Kerstens Z. L., Law A. S., Megens H. J., Milan D., Nonneman D. J., Rohrer G. A., Rothschild M. F., Smith T. P., Schnabel R. D., Van Tassell C. P., Taylor J. F., Wiedmann R. T., Schook L. B., Groenen M. A.. 2009. Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology. PLoS ONE 4:e6524. A-LinkCrossRefMedlineGoogle Scholar

↵ Sahana G., Kadlecová V., Hornshøj H., Nielsen B., Christensen O. F.. 2013. A genome-wide association scan in pig identifies novel regions associated with feed efficiency trait. J. Anim. Sci. 91:1041–1050. Abstract/FREE Full Text

↵ Serizawa S., Miyamichi K., Takeuchi H., Yamagishi Y., Suzuki M., Sakano H.. 2006. A neuronal identity code for the odorant receptor-specific and activity-dependent axon sorting. Cell 127:1057–1069. A-LinkCrossRefMedlineGoogle Scholar

↲ Therneau T., Atkinson E., Sinnwell J., Matsumoto M., Schaid D., McDonnell S.. 2012. kinship2: Pedigree functions. R package version 1.3.7. http://cran.r-project.org/web/packages/kinship2/index.html (accessed 9 Nov. 2012).

↲ Van Gestel S., Houwing-Duistermaat J. J., Adolfsson R., van Duijn C. M., Van Broeckhoven C.. 2000. Power of selective genotyping in genetic association analyses of quantitative traits. Behav. Genet. 30:141–146. A-LinkCrossRefMedlineGoogle Scholar

↲ 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 425:832–836. A-LinkCrossRefMedlineGoogle Scholar

↲ Yariz K. O., Duman D., Seco C. Z., Dallman J., Huang M., Peters T. A., Sirmaci A., Lu N., Schraders M., Skromne I., Oostrik J., Diaz-Horta O., Young J. I., Tokgoz-Yilmaz S., Konukseven O., Shahin H., Hetterschijt L., Kanaan M., Oonk A. M., Edwards Y. J., Li H., Atalay S., Blanton S., Desmidt A. A., Liu X. Z., Pennings R. J., Lu Z., Chen Z. Y., Kremer H., Tekin M.. 2012. Mutations in OTOGL, encoding the inner ear protein otogelin-like, cause moderate sensorineural hearing loss. Am. J. Hum. Genet. 91:872–882. A-LinkCrossRefMedlineGoogle Scholar

↲ Zhang G., Nebert D. W., Chakraborty R., Jin L.. 2006. Statistical power of association using the extreme discordant phenotype design. Pharmacogenet. Genomics 16:401–413. A-LinkCrossRefMedlineGoogle Scholar

↲ Zhao S., Zhang J., Hou X., Zan L., Wang N., Tang Z., Li K.. 2012. OLFML3 expression is decreased during prenatal muscle development and regulated by microRNA-155 in pigs. Int. J. Biol. Sci. 8:459–469. A-LinkMedline

Figure 1. Manhattan plot of SNP in the genome wide association study with average daily gain in Italian Large White pigs. Red line: $P_{Bonferroni} = 0.01$; blue line: $P_{Bonferroni} = 0.10$.

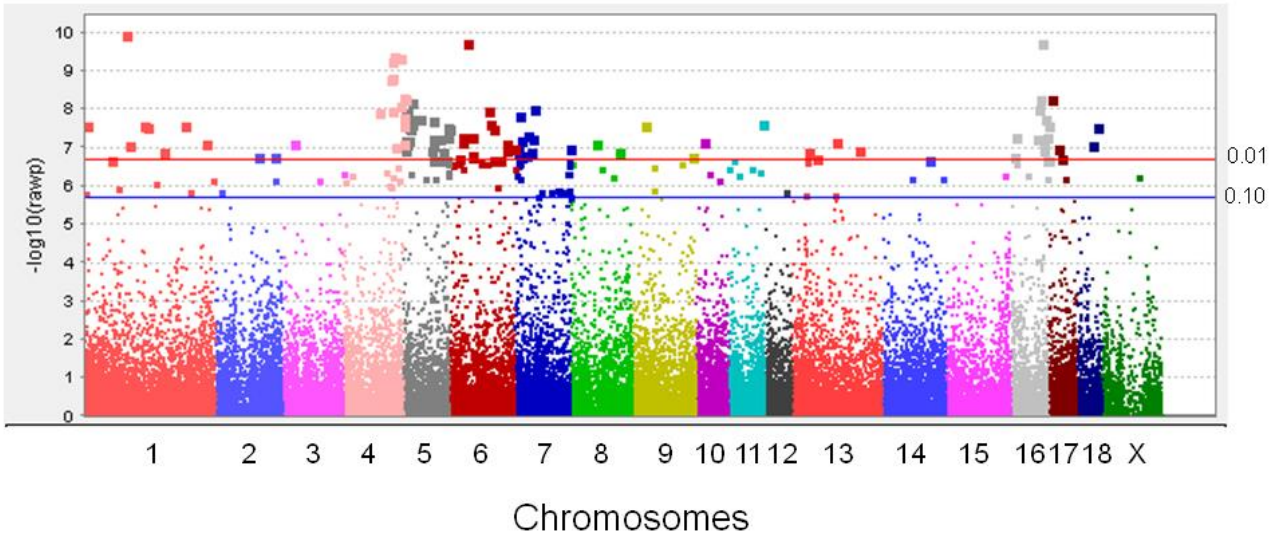

-log10(rawp)
0
1
2
3
4
5
6
7
8
9
10
0.01
0.10
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
X
Chromosomes

Table 1. Significant SNP associated with average daily gain in Italian Large White pigs with information about the closest gene as reported in Sscrofa10.2.

SNP [1]	SSC	Position	P_{raw}	$P_{Bonferroni}$	Gene Symbol [2]	Distance (bp) [2]	SNP [1]	SSC	Position [3]	P_{raw}	$P_{Bonferroni}$	Gene Symbol [2]	Distance (bp) [2]
ALGA0000014*	1	56,5627	2.30E-08	1.14E-03	*CH242-271M12.1*	72,271	M1GA0008525	6	47,265,406	4.76E-08	2.35E-03	*BCL3*	97,353
ALGA0003521	1	60,455,614	1.82E-07	9.01E-03	*DDX43*	1,041,253	M1GA0008519	6	47,321,053	4.76E-08	2.35E-03	*PRR2*	Intragenic
ALGA0004718*	1	94,737,894	1.04E-10	5.15E-06	Novel pseudogene	36,318	M1GA0008536	6	50,495,796	1.50E-07	7.43E-03	*RRAS*	82,042
ALGA0004837	1	99,300,150	1.17E-10	5.78E-06	*HTR1B*	165,442	M1GA0008537	6	50,638,891	1.45E-07	7.15E-03	*AP2A1*	85,602
ALGA0004958	1	102,500,451	7.79E-08	3.85E-03	*MB21D1*	1,035,398	M1GA0008539	6	5,084,7065	4.76E-08	2.35E-03	*VRK3*	96,686
ASGA0096650	1	137,394,981	2.50E-08	1.24E-03	*FBN1*	91,415	ALGA0036150	6	91,513,570	1.01E-08	5.00E-04	-	22,354
ALGA0005986	1	145,646,235	2.74E-08	1.35E-03	*DLL4*	97,416	M1GA0008850	6	92,900,402	2.11E-08	1.04E-03	-	47,209
ALGA0006831	1	184,400,666	1.11E-07	5.50E-03	*CORO2B*	Intragenic	M1GA0008859*	6	100,662,783	3.01E-08	1.49E-03	*CABLES1*	Intragenic
ALGA0007846*	1	238,616,307	2.30E-08	1.14E-03	-	1,025,259	M1GA0008862*	6	101,205,770	1.87E-07	9.23 E-03	*LAMA3*	Intragenic
ALGA0009614	1	287,443,993	6.74E-08	3.33E-03	*TRIM32*	1,006,260	M1GA0008864*	6	101,772,766	1.87E-07	9.23 E-03	*OSBPL1A*	Intragenic
ALGA0014534	2	98,997,793	1.61E-07	7.94E-03	*MEF2C*	112,521	M1GA0008893*	6	119,164,257	1.87E-07	9.23E-03	*PDZK1IP1*	84,766
ALGA0016010	2	137,184,334	1.59E-07	7.85E-03	*SLC27A6*	9,625	DRGA0006924	6	132,658,334	1.10E-07	5.45E-03	*LRRC7*	45,971
ALGA0018040	3	23,224,453	6.75E-08	3.33E-03	*LOC100522792*	49,643	DRGA0006951	6	134,411,824	7.30E-08	3.61E-03	*SLC35D1*	29,042
ALGA0025924	4	80,363,802	1.08E-08	5.32E-04	*RAB2A*	1,080,804	M1GA0009091	6	155,139,641	9.80E-08	4.84E-03	*LOC100737128*	61,610
M1GA0006238	4	111,511,647	1.47E-09	7.26E-05	*HMGCS2*	8,1005	M1GA0009342	7	1,611,710	1.22E-07	6.01E-03	*LOC100157526*	62,661
M1GA0006250	4	112,746,036	9.82E-09	4.85E-04	*LOC100155374*	Intragenic	M1GA0009339	7	1,780,404	9.80E-08	4.84E-03	*LOC100157526*	90,238
M1GA0006299*	4	114,151,582	1.30E-09	6.44E-05	*IGSF3*	Intragenic	M1GA0009374	7	3,571,588	9.80E-08	4.84E-03	*FARS2*	23,283
M1GA0006302*	4	114,170,369	4.84E-10	2.39E-05	*IGSF3*	4,195	M1GA0009455	7	5,113,060	9.80E-08	4.84E-03	*LOC100156744*	91,839
M1GA0006343*	4	116,750,464	3.57E-10	1.76E-05	*OLFML3*	43,229	M1GA0009500	7	7,664,403	1.27E-08	6.28E-04	*LOC100738362*	32,601
M1GA0006478	4	122,816,731	8.35E-08	4.13E-03	*LOC100737105*	Intragenic	M1GA0009555	7	1,0031,634	1.49E-07	7.34E-03	*LOC100739137*	52,139
M1GA0006616	4	134,698,568	7.21E-09	3.56E-04	*LOC100152734*	32,777	M1GA0009568	7	1,0543,393	5.71E-08	2.82E-03	*CCDC90A*	5,345
M1GA0006613	4	134,796,609	4.32E-10	2.13E-05	*LOC100155583*	26,428	M1GA0009677	7	21,031,127	1.41E-07	6.97E-03	*CMAHP*	1,162
M1GA0006828	4	141,168,185	2.40E-08	1.18E-03	*CLCA2*	1,020,511	M1GA0009735	7	25,272,090	4.20E-08	2.07E-03	*MOG*	79,658
M1GA0006854	4	141,458,072	7.05E-08	3.48E-03	*HS2ST1*	92,125	M1GA0009865	7	33,358,569	1.16E-07	5.75E-03	*BEND6*	85,410
M1GA0006869	4	141,552,372	4.36E-09	2.16E-04	*HS2ST1*	Intragenic	M1GA0010028	7	37,366,668	5.42E-08	2.68E-03	*CDKN1A*	99,505
M1GA0006938	4	142,907,745	1.43E-08	7.06E-04	*ZNHIT6*	32,699	ALGA0040777	7	41,624,144	9.12E-09	4.51E-04	*LOC100518497*	6,957
M1GA0006965	4	143,224,099	4.66E-09	2.30E-04	*DDAH1*	17,835	M1GA0011382	7	130,525,474	9.49E-08	4.69E-03	-	1,013,453
M1GA0007072	5	844,337	1.08E-08	5.32E-04	*ATXN10*	Intragenic	M1GA0011548	8	723,623	4.54E-08	2.24E-03	*WHSC2*	96,776
M1GA0007246*	5	2,608,692	6.34E-08	3.13E-03	*SULT4A1*	60,112	ALGA0047898	8	56,440,088	7.30E-08	3.61E-03	*REST*	1,000,471
M1GA0007255	5	2,683,343	9.98E-08	4.93E-03	*5S_rRNA*	10,703	ALGA0048976	8	112,525,755	1.16E-07	5.75E-03	*SYNPO2*	Intragenic
M1GA0007258*	5	2,748,616	5.73E-08	2.83E-03	*EFCAB6*	Intragenic	ASGA0042165	9	27,635,662	2.32E-08	1.146E-03	-	1,118,578
M1GA0007286	5	3,095,194	1.06E-07	5.26E-03	*SCUBE1*	Intragenic	ALGA0055314	9	139,257,822	1.58E-07	7.83E-03	*U6*	26,534
M1GA0007352	5	4,855,248	1.20E-08	5.95E-04	*XPNPEP3*	91,269	ALGA0057214	10	15,467,849	6.62E-08	3.27E-03	*CNIH3*	81,000
ALGA0030091	5	5,363,246	6.75E-08	3.33E-03	*ADSL*	85,183	H3GA0032476	11	77,940,001	2.07E-08	1.02E-03	*FGF14*	99,934
M1GA0007436	5	6,942,486	1.12E-08	5.54E-04	*CSNK1E*	39,565	H3GA0036210	13	36,964,009	1.19E-07	5.90E-03	*RBM15B*	70,931

M1GA0007494*	5	8,031,569	2.23E-08	1.10E-03	*CARD10*	65,189	DRGA0012382	13	56,900,313	1.65E-07	8.16E-03	-	57,664
M1GA0007506	5	8,303,062	2.74E-08	1.35E-03	*C1QTNF6*	99,276	DRGA0012768	13	103,171,133	6.39E-08	3.16E-03	*MME*	42,849
M1GA0007538	5	9,292,551	5.57E-08	2.75E-03	*LOC100517940*	16,264	ALGA0072425	13	159,407,802	1.05E-07	5.19E-03	*U6*	20,127
M1GA0007600	5	11,249,463	2.08E-08	1.03E-03	*SYN3*	1,011,272	ALGA0080306	14	10,7939,105	1.85E-07	9.16E-03	*A1CF*	90,844
M1GA0007630	5	14,225,240	2.78E-08	1.37E-03	*LOC100620963*	21,192	ALGA0088670	16	4,016,044	1.59E-07	7.85E-03	*TRIO*	16,110
M1GA0007662	5	16,120,363	6.18E-09	3.06E-04	*AQP6*	49,433	ALGA0088909	16	7,471,376	4.76E-08	2.35E-03	*FAM134B*	1,146,830
M1GA0007707	5	18,676,679	2.18E-08	1.08E-03	*KRT4*	84,202	ALGA0090834	16	58,445,833	5.02E-08	2.48E-03	*FAM196B*	41,336
M1GA0007772	5	30,095,581	1.66E-08	8.20E-04	*MON2*	85,519	M1GA0021097	16	59,838,300	8.72E-09	4.31E-04	*SLIT3*	47,859
M1GA0007784	5	35,802,016	1.66E-08	8.20E-04	*LOC100152555*	65,262	ALGA0091161	16	66,890,995	4.89E-09	2.42E-04	*LOC100516706*	10,282
M1GA0007824	5	65,432,242	1.09E-07	5.40E-03	*PHC1*	54,480	M1GA0021128	16	69,042,784	1.61E-10	7.94E-06	*LOC100519063*	38,294
M1GA0007840	5	66,740,501	7.96E-08	3.93E-03	*VAMP1*	92,758	M1GA0021136	16	73,257,168	8.97E-08	4.43E-03	*C5ORF4*	1,021,984
M1GA0007853	5	67,619,158	1.84E-08	9.11E-04	*KCNA5*	34,256	ALGA0091438	16	73,552,942	1.09E-07	5.38E-03	*HAND1*	1,036,732
M1GA0007928	5	69,437,477	1.97E-07	9.75E-03	*TSPAN9*	43,044	M1GA0021168	16	77,794,570	1.66E-08	8.20E-04	*FAT2*	34,727
M1GA0007944	5	69,759,629	5.35E-08	2.64E-03	*SLC6A12*	73,043	M1GA0021255	16	80,842,730	2.00E-07	9.86E-03	-	Intragenic
M1GA0008010	5	74,367,428	5.35E-08	2.64E-03	-	20,916	M1GA0021335	16	82,055,821	4.76E-08	2.35E-03	-	1,129,166
M1GA0008025	5	79,010,106	1.97E-07	9.75E-03	*SLC38A2*	1,029,163	M1GA0021462	16	84,330,710	1.87E-07	9.23E-03	*IRX4*	1,228,304
M1GA0008064	5	85,112,008	5.35E-08	2.64E-03	*LOC100738422*	51,201	M1GA0021563*	16	85,843,098	2.30E-08	1.14E-03	*LPCAT1*	5,430
M1GA0008091	5	91,724,948	1.97E-07	9.75E-03	*HAL*	769	ALGA0092903	17	5,999,136	5.10E-09	2.52E-04	*MTUS1*	53,654
M1GA0008099	5	92,805,456	1.18E-07	5.85E-03	-	1,025,329	M1GA0021675	17	20,631,316	9.80E-08	4.84E-03	*PLCB4*	Intragenic
M1GA0008133	5	100,729,561	1.97E-07	9.75E-03	*MGAT4C*	90,555	M1GA0021697	17	28,221,039	1.76E-07	8.72E-03	*BFSP1*	1,122,141
M1GA0008142	5	103,561,458	4.82E-08	2.38E-03	*U6*	92,710	ALGA0097816	18	32,645,078	7.79E-08	3.85E-03	*FOXP2*	1,020,806
M1GA0008159	5	105,349,192	2.60E-08	1.29E-03	*ACSS3*	Intragenic	ALGA0098168*	18	45,408,799	2.65E-08	1.31E-03	*GHRHR*	1,018,328
M1GA0008164	5	106,046,784	2.90E-08	1.43E-03	*OTOGL*	43,204	M1GA0008884	JH118434.1	15,0095	1.87E-07	9.23E-03	-	-
M1GA0008394	6	18,653,852	1.79E-07	8.83E-03	*CCDC135*	1,073,849	ALGA0030787	0	0	3.19E-11	1.57E-06	-	-
M1GA0008405	6	26,197,426	6.62E-08	3.27E-03	*LOC100517946*	80,522	ASGA0102104	0	0	1.94E-09	9.60E-05	-	-
M1GA0008418	6	27,367,862	6.71E-08	3.32E-03	*LOC100737013*	Intragenic	M1GA0006887	0	0	1.08E-08	5.32E-04	-	-
M1GA0008438	6	28,304,343	4.63E-08	2.29E-03	*TOX3*	1,002,255	M1GA0021138	0	0	2.90E-08	1.43E-03	-	-
M1GA0008473	6	30,900,258	4.76E-08	2.35E-03	*ZNF423*	36,367	ALGA0054046	0	0	9.71E-08	4.80E-03	-	-
ALGA0035254	6	40,760,438	1.61E-10	7.94E-06	*KIRREL2*	93,901							

[1] SNP with asterisk were associated with backfat thickness in our previous genome wide association study (Fontanesi et al., 2012d).

[2]Additional information on close genes and Ensembl ID is reported in Supplementary Table 1.

[3] SNP not assigned to any Sscrofa10.2 position are indicated with chromosome (SSC) position "0". A few SNP were assigned to unassembled scaffolds. Several of these SNP were mapped on Sscofa9.2: ALGA0030787, SSC5, position 11,032,453; M1GA0006887, SSC4, 126,275,676; M1GA0021138, SSC16, 43,907,250; M1GA0011548, SSC8, 17,610; ALGA0054046, SSC9, 46,348,451

Supplementary Table 1.

SNP with $P_{Bonferroni}$ <0.10 and annotations. If the SNP was significant in previous works related to productive traits it is reported.

SNP	SSC	Position	Praw	PBonferroni	Sign. in BFT GWAS	Closest gene (Ensembl ID)	Distance to the closest gene	Closest gene symbol	Closest gene description - name	Other genes within +/-100 kb
ALGA0030787	0	0	3.19E-11	1.57E-06						
ASGA0102104	0	0	1.94E-09	9.60E-05						
M1GA0006887	0	0	1.08E-08	0.000532279						
M1GA0021138	0	0	2.90E-08	0.001432899						
ALGA0054046	0	0	9.71E-08	0.004798405						
M1GA0011187	0	0	2.46E-07	0.012133705						
MARC0074011	0	0	4.31E-07	0.021280279						
MARC0076810	0	0	4.63E-07	0.022898186						
ALGA0000009	1	538161	1.46E-06	0.072267256	BFT	ENSSSCG00000030932	44805	CH242-271M12.1		
ALGA0000014	1	565627	2.30E-08	0.001135863		ENSSSCG00000030932	72271	CH242-271M12.1		
ALGA0003521	1	60455614	1.82E-07	0.009012631		ENSSSCG00000004287	1041253	DDX43		DEAD (Asp-Glu-Ala-Asp) box polypeptide 43
ASGA0003420	1	81652927	1.17E-06	0.057977283		ENSSSCG00000004369	51649	PRDM1	PR domain containing 1, with ZNF domain	
ALGA0004718	1	94737894	1.04E-10	5.15E-06	BFT	ENSSSCG00000004463	36318	novel pseudogene		
ALGA0004837	1	99300150	1.17E-10	5.78E-06		ENSSSCG00000004475	165442	HTR1B	5-hydroxytryptamine (serotonin) receptor 1B (HTR1B), mRNA.	
ALGA0004958	1	102500451	7.79E-08	0.003848717		ENSSSCG00000021383	1035398	MB21D1		Mab-21 domain containing 1
ASGA0096650	1	137394981	2.50E-08	0.001236816		ENSSSCG00000004658	91415	FBN1	fibrillin 1	LOC100621123,DUT

ALGA0005986	1	145646235	2.74E-08	0.001353428		ENSSSCG00000004755	97416	DLL4	delta-like 4 (Drosophila)	ZFYVE19,FAM82A2,VPS18,RHOV,C15ORF62,DNAJC17,SPINT1,LOC100519696,PPP1R14D,LOC100519405,GCHFR
ALGA0006483	1	168603720	8.73E-07	0.043125089	BFT	ENSSSCG00000028977	1306510	DOK6		
ALGA0006831	1	184400666	1.11E-07	0.005500999		ENSSSCG00000004962	INTRAgenic	CORO2B	coronin, actin binding protein, 2B	
ALGA0007846	1	238616307	2.30E-08	0.001135863	BFT	ENSSSCG00000005196	1025259			transmembrane protein C9orf123-like LOC100155846
ALGA0008236	1	254016770	1.37E-06	0.067794364		ENSSSCG00000005267	1008967	ANXA1		annexin A1
ALGA0009614	1	287443993	6.74E-08	0.003328804		ENSSSCG00000005501	1006260	TRIM32		E3 ubiquitin-protein ligase TRIM32-like
MARC0074154	1	308379812	6.85E-07	0.033849084		ENSSSCG00000029715	26519	OLFM1	noelin-like	LOC100620523
ALGA0011875	2	10706436	1.41E-06	0.069663304		ENSSSCG00000027387	30455	MS4A12	membrane-spanning 4-domains subfamily A member 12-like	LOC100627859,LOC100628049,MS4A5,MS4A1
ALGA0014534	2	98997793	1.61E-07	0.007936289		ENSSSCG00000014149	112521	MEF2C		myocyte enhancer factor 2C
ALGA0016010	2	137184334	1.59E-07	0.007846383		ENSSSCG00000014257	9625	SLC27A6		LOC100517640,ISOC1
ALGA0016219	2	140038667	6.98E-07	0.034483874		ENSSSCG00000014274	64719	PDLIM4		SLC22A4,SLC22A5
ALGA0018040	3	23224453	6.75E-08	0.003334092		ENSSSCG00000007833	49643	LOC100522792	protein kinase C beta type-like	LOC100736706
ALGA0019910	3	85838482	6.57E-07	0.032459717		ENSSSCG00000027414	1020310			
ALGA0117579	3	143791334	4.79E-07	0.023689888		ENSSSCG00000008661	1110351			
MARC0072995	4	2600559	7.49E-07	0.036996541		ENSSSCG00000005934	INTRAgenic	TRAPPC9		
ALGA0023678	4	17726338	1.47E-06	0.072664055	BFT	ENSSSCG00000030871	208865			
ALGA0023852	4	20331468	4.88E-07	0.02413651		ENSSSCG00000006000	46650	TAF2	TAF2 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 150kDa	ENPP2
ALGA0025924	4	80363802	1.08E-08	0.000532279		ENSSSCG00000006232	1080804	RAB2A		
M1GA0006144	4	103706322	4.41E-07	0.021803726		ENSSSCG00000006541	97879	DCST2	DC-STAMP domain containing 2	KCNN3,LOC100151902,PBXIP1,PMVK
DIAS0002618	4	108974555	9.21E-07	0.04552915		ENSSSCG00000025491	32495	U1		LOC100157381,RBM8A,GNRHR2,HFE2,LIX1L,POLR3GL,ITGA10,PEX11B,TXNIP,LOC100153742
M1GA0006238	4	111511647	1.47E-09	7.26E-05		ENSSSCG00000006716	81005	HMGCS2	3-hydroxy-3-methylglutaryl-CoA synthase 2 (mitochondrial)	LOC100522133,ZNF697,CH242-150C11.4,HSD3B1,PHGDH,HAO2
M1GA0006250	4	112746036	9.82E-09	0.000485001		ENSSSCG00000006726	INTRAgenic	LOC100155374	sperm-associated antigen 17-like	SPAG17,LOC100153342,WDR3,GDAP2

M1GA0006291	4	113822178	1.04E-06	0.051408361		ENSSSCG00000006735	22951	PTGFRN	prostaglandin F2 receptor negative regulator	U6
M1GA0006299	4	114151582	1.30E-09	6.44E-05	BFT	ENSSSCG00000006737	INTRAgenic	IGSF3	immunoglobulin superfamily, member 3	LOC100622379
M1GA0006302	4	114170369	4.84E-10	2.39E-05	BFT	ENSSSCG00000006737	4195	IGSF3	immunoglobulin superfamily, member 3	LOC100622379
M1GA0006343	4	116750464	3.57E-10	1.76E-05	BFT	ENSSSCG00000006759	43229	OLFML3	olfactomedin-like 3	HIPK1
ALGA0027862	4	117329189	5.79E-07	0.028596467		ENSSSCG00000006767	62738	MAGI3	membrane associated guanylate kinase, WW and PDZ domain containing 3	
M1GA0006478	4	122816731	8.35E-08	0.004125104		ENSSSCG00000006851	INTRAgenic	LOC100737105	guanine nucleotide exchange factor VAV3-like	
ALGA0028664	4	129539293	3.17E-07	0.015641385		ENSSSCG00000006864	46160	CDC14A	CDC14 cell division cycle 14 homolog A (S. cerevisiae)	LRRC39,TRMT13,RTCA,LOC100622510,LOC100623721,LOC100513729,DBT
ALGA0028834	4	131377049	6.98E-07	0.034483874	BFT	ENSSSCG00000022071	75423			
M1GA0006616	4	134698568	7.21E-09	0.000356219		ENSSSCG00000027942	32777	LOC100152734		
M1GA0006613	4	134796609	4.32E-10	2.13E-05		ENSSSCG00000027942	26428	LOC100155583 (ARHGAP29)		ABCA4,LOC100152734
M1GA0006828	4	141168185	2.40E-08	0.001183768		ENSSSCG00000006935	1020511	CLCA2		chloride channel accessory 2
M1GA0006854	4	141458072	7.05E-08	0.003481537		ENSSSCG00000022032	92125	HS2ST1	heparan sulfate 2-O-sulfotransferase 1	
M1GA0006869	4	141552372	4.36E-09	0.000215542		ENSSSCG00000022032	INTRAgenic	HS2ST1	heparan sulfate 2-O-sulfotransferase 1	
M1GA0006938	4	142907745	1.43E-08	0.000705504		ENSSSCG00000006939	32699	ZNHIT6	zinc finger, HIT-type containing 6	CYR61
M1GA0006965	4	143224099	4.66E-09	0.000230046		ENSSSCG00000006941	17835	DDAH1	dimethylarginine dimethylaminohydrolase 1	LOC100155842,SYDE2,LOC100153411,LOC100152216,BCL10
M1GA0007072	5	844337	1.08E-08	0.000532279		ENSSSCG00000024864	INTRAgenic	ATXN10	ataxin 10	
M1GA0007246	5	2608692	6.34E-08	0.003134744	BFT	ENSSSCG00000000026	60112	SULT4A1	sulfotransferase 4A1-like	LOC100523243,5S_rRNA,EFCAB6,LOC100524434
M1GA0007255	5	2683343	9.98E-08	0.004929156		ENSSSCG00000019789	10703	5S_rRNA		

M1GA0007258	5	2748616	5.73E-08	0.00283128	BFT	ENSSSCG00000000028	INTRAgenic	EFCAB6	EF-hand calcium-binding domain-containing protein 6-like	LOC100524434,5S_rRNA
M1GA0007286	5	3095194	1.06E-07	0.005256477		ENSSSCG00000000029	INTRAgenic	SCUBE1	signal peptide, CUB domain, EGF-like 1	LOC100523602,MCAT,TSPO,TTLL12
M1GA0007352	5	4855248	1.20E-08	0.00059458		ENSSSCG00000022335	91269	XPNPEP3		LOC100624537,LOC100624451,MCHR1,RPL31,ST13,LOC100517713,SLC25A17
ALGA0030091	5	5363246	6.75E-08	0.003334092		ENSSSCG00000000077	85183	ADSL	adenylosuccinate lyase	TNRC6B,LOC100516974
M1GA0007436	5	6942486	1.12E-08	0.000554255		ENSSSCG00000000107	39565	CSNK1E	casein kinase 1, epsilon	LOC100514711,LOC100736927,PLA2G6,LOC100156484,TMEM184B,MAFF
M1GA0007494	5	8031569	2.23E-08	0.001103167	BFT	ENSSSCG00000025260	65189	CARD10		LOC100511563,CYTH4,ELFN2
M1GA0007506	5	8303062	2.74E-08	0.001353428		ENSSSCG00000030325	99276	C1QTNF6	C1q and tumor necrosis factor related protein 6	LOC100512279,LOC100154490,KCTD17,TEX33,LOC100152026,TMPRSS6,MPST,TSTD1,LOC100621776,TST
M1GA0007538	5	9292551	5.57E-08	0.002750517		ENSSSCG00000000148	16264	LOC100517940	apolipoprotein L3-like	APOL3,LOC100626503
M1GA0007600	5	11249463	2.08E-08	0.00102992		ENSSSCG00000000154	1011272	SYN3		synapsin-3-like LOC100519024
M1GA0007630	5	14225240	2.78E-08	0.001373678		ENSSSCG00000000170	21192	LOC100620963	T-complex protein 11-like protein 2-like	NUAK1
M1GA0007662	5	16120363	6.18E-09	0.00030558		ENSSSCG00000000212	49433	AQP6	aquaporin 6, kidney specific	FAIM2,AQP5,AQP2,LOC100516732,LOC100515481,LOC100739053,BCDIN3D,NCKAP5L
M1GA0007687	5	17949392	4.87E-07	0.024060542		ENSSSCG00000000232	77966	ACVRL1	activin A receptor type II-like 1	GRASP,KRT84,KRT85,LOC100522015,ACVR1B
M1GA0007707	5	18676679	2.18E-08	0.001077701		ENSSSCG00000024610	84202	KRT4	keratin 4	SOAT2,KRT79,KRT18,LOC100515544,IGFBP6,KRT8,LOC100626135,EIF4B
M1GA0007772	5	30095581	1.66E-08	0.000820252		ENSSSCG00000000458	85519	MON2	protein MON2 homolog	LOC100154568,ssc-let-7i,MIRLET7I
M1GA0007784	5	35802016	1.66E-08	0.000820252		ENSSSCG00000030014	65262	LOC100152555	ras-related protein Rap-1b-like	LOC100515832,LOC100739054,RAP1A,LOC100736973,MDM2,SLC35E3
ALGA0118428	5	54100199	6.35E-07	0.031376164		ENSSSCG00000018927	96139			
M1GA0007824	5	65432242	1.09E-07	0.005401051		ENSSSCG00000000661	54480	PHC1		RIMKLB,LOC100738680,KLRG1,A2ML1,M6PR
M1GA0007840	5	66740501	7.96E-08	0.003933634		ENSSSCG00000000703	92758	VAMP1	vesicle-associated membrane protein 1-like	LOC100519841,LTBR,PLEKHG6,TNFRSF1A,CD27,LOC100520196,TAPBPL,SCNN1A,LOC100520908
M1GA0007853	5	67619158	1.84E-08	0.000911205		ENSSSCG00000021596	34256	KCNA5	potassium voltage-gated channel, shaker-related subfamily, member 5	KV1.5

M1GA0007928	5	69437477	1.97E-07	0.009749955		ENSSSCG00000000735	43044	TSPAN9	tetraspanin 9	TULP3,RHNO1,LOC100511381,FOXM1,LOC100511196,TEAD4
M1GA0007944	5	69759629	5.35E-08	0.002642588		ENSSSCG00000000747	73043	SLC6A12		IQSEC3
M1GA0008007	5	72590667	2.16E-07	0.010671976		ENSSSCG00000027190	94796	LOC100517945	peroxisome assembly protein 26-like	USP18,LOC100518253
DRGA0006426	5	72907472	1.86E-06	0.091758617		ENSSSCG00000024523	1022345	SLC2A13		
M1GA0008010	5	74367428	5.35E-08	0.002642588		ENSSSCG00000000783	20916			
ALGA0118429	5	76596356	6.35E-07	0.031376164		ENSSSCG00000027887	41933	7SK		
M1GA0008025	5	79010106	1.97E-07	0.009749955		ENSSSCG00000000808	1029163	SLC38A2		sodium-coupled neutral amino acid transporter 2-like
M1GA0008047	5	83722723	2.16E-07	0.010671976		ENSSSCG00000000843	21068	TXNRD1	thioredoxin reductase 1	LOC100514406,LOC100514233
M1GA0008064	5	85112008	5.35E-08	0.002642588		ENSSSCG00000026314	51201	LOC100738422	phenylalanine-4-hydroxylase-like	LOC100521900
M1GA0008091	5	91724948	1.97E-07	0.009749955		ENSSSCG00000024152	769	HAL	histidine ammonia-lyase-like	LTA4H,AMDHD1,CCDC38,SNRPF,SNORAD72
M1GA0008099	5	92805456	1.18E-07	0.005851355		ENSSSCG00000000911	1025329			
M1GA0008133	5	100729561	1.97E-07	0.009749955		ENSSSCG00000000931	90555	MGAT4C	mannosyl (alpha-1,3-)-glycoprotein beta-1,4-N-acetylglucosaminyltransferase, isozyme C (putative)	
M1GA0008142	5	103561458	4.82E-08	0.002382548		ENSSSCG00000021240	92710	U6		LOC100620150
M1GA0008159	5	105349192	2.60E-08	0.00128535		ENSSSCG00000000939	INTRAgenic	ACSS3	acyl-CoA synthetase short-chain family member 3	
M1GA0008164	5	106046784	2.90E-08	0.001432899		ENSSSCG00000000943	43204	OTOGL		
DRGA0006447	5	107303857	4.32E-07	0.021322437		ENSSSCG00000000953	1020019			
DRGA0006450	5	107536603	2.69E-07	0.013282353		ENSSSCG00000000943	1533023	OTOGL		
M1GA0008203	5	107862487	4.88E-07	0.02413651		ENSSSCG00000020514	1168322	5S_rRNA		
DRGA0006468	5	108606372	2.63E-07	0.012989755		ENSSSCG00000024510	1210686	CAND1		cullin-associated and neddylation-dissociated 1
M1GA0008199	5	109063391	4.18E-07	0.020648366		ENSSSCG00000023494	1024189			
DRGA0006491	6	5951885	2.69E-07	0.013282353		ENSSSCG00000002688	1042212	PLCG2		phospholipase C, gamma 2 (phosphatidylinositol-specific)
M1GA0008318	6	7416050	2.87E-07	0.014162469		ENSSSCG00000002690	51354	GAN	gigaxonin	

M1GA0008394	6	18653852	1.79E-07	0.00883389		ENSSSCG00000002817	1073849	CCDC135		coiled-coil domain-containing protein 135-like LOC100515380
DRGA0006574	6	22460362	2.69E-07	0.013282353		ENSSSCG00000002795	1247967	CDH11		cadherin 11 type 2 OB-cadherin (osteoblast)
M1GA0008405	6	26197426	6.62E-08	0.003272114		ENSSSCG00000002763	80522	LOC100517946	protein arginine N-methyltransferase 7-like	PRMT7,LOC100518128,SMPD3
M1GA0008418	6	27367862	6.71E-08	0.003317915		ENSSSCG00000002826	INTRAgenic	LOC100737013	liver carboxylesterase-like	SLC6A2,LOC100517530,CES1,LOC100517716,LPCAT2,LOC100736962
M1GA0008432	6	28215213	2.19E-07	0.010797822	BFT	ENSSSCG00000002831	44338	LOC100518611	iroquois-class homeodomain protein IRX-3-like	IRX3
M1GA0008438	6	28304343	4.63E-08	0.002288641		ENSSSCG00000002835	1002255	TOX3		TOX high mobility group box family member 3
M1GA0008445	6	29678259	3.29E-07	0.016268485		ENSSSCG00000019170	1001995			
M1GA0008473	6	30900258	4.76E-08	0.002351677		ENSSSCG00000002838	36367	ZNF423	zinc finger protein 423	
ALGA0035254	6	40760438	1.61E-10	7.94E-06		ENSSSCG00000002915	93901	KIRREL2	kin of IRRE like 2 (Drosophila)	LOC100522440,SDHAF1,U6,HCST,TYROBP,LOC100521795,PRODH2,LOC100521971,WDR62,NFKBID,LRFN3,THAP8,LOC100524441,LOC100523906,LOC100524259,LOC100524089,ALKBH6,APLP1,CLIPR-59,SYNE4
M1GA0008525	6	47265406	4.76E-08	0.002351677		ENSSSCG00000027426	97353	BCL3	B-cell CLL/lymphoma 3	CBLC,BCAM,LOC100738241,PRR2,TOMM40,APOC4,APOE,APOC2,CLPTM1
M1GA0008519	6	47321053	4.76E-08	0.002351677		ENSSSCG00000003090	INTRAgenic	PRR2	poliovirus receptor related 2	TOMM40,SYMPK,APOC4,APOE,APOC2,CLPTM1
M1GA0008536	6	50495796	1.50E-07	0.007430287		ENSSSCG00000003183	82042	RRAS	ras-related protein R-Ras-like	LOC100516558,IRF3,LOC100522621,CPT1C,LOC100522079,TSKS,BCL2L12,PTOV1,MED25,PRMT1,FUZ,AP2A1,SCAF1,LOC100519902,ADM5
M1GA0008537	6	50638891	1.45E-07	0.007151422		ENSSSCG00000003199	85602	AP2A1	adaptor-related protein complex 2, alpha 1 subunit	LOC100522621,FUZ,NUP62,LOC100523675,PNKP,PTOV1,MED25,LOC100519715,AKT1S1,LOC100519902,ATF5,IL4I1,TBC1D17
M1GA0008539	6	50847065	4.76E-08	0.002351677		ENSSSCG00000003202	96686	VRK3		LOC100524748,LOC100511457,U6,IZUMO2,LOC100524938,LOC100523004,ZNF473
M1GA0008727	6	74175566	2.10E-07	0.010383315		ENSSSCG00000003520	58742	CDC42	cell division cycle 42 (GTP binding protein, 25kDa)	WNT4
DRGA0006627	6	74312696	2.63E-07	0.012989755		ENSSSCG00000027849	1052040	CNR2		cannabinoid receptor 2-like
DRGA0006644	6	91142077	2.63E-07	0.012989755		ENSSSCG00000003683	1013692	NDUFV2		NDUFV2 NUDFV2

ALGA0036150	6	91513570	1.01E-08	0.000499685		ENSSSCG00000003676	22354			
M1GA0008850	6	92900402	2.11E-08	0.001040829		ENSSSCG00000029947	47209			
M1GA0008859	6	100662783	3.01E-08	0.00148884	BFT	ENSSSCG00000003705	INTRAgenic	CABLES1	CDK5 and ABL1 enzyme substrate 1-like	LOC100739096
M1GA0008862	6	101205770	1.87E-07	0.009233816	BFT	ENSSSCG00000025584	INTRAgenic	LAMA3		
M1GA0008864	6	101772766	1.87E-07	0.009233816	BFT	ENSSSCG00000003712	INTRAgenic	OSBPL1A	oxysterol binding protein-like 1A	U6
M1GA0008880	6	112770027	9.76E-07	0.048243917		ENSSSCG00000003742	1025364	INO80C		INO80 complex subunit C-like
M1GA0008893	6	119164257	1.87E-07	0.009233816	BFT	ENSSSCG00000003753	84766	PDZK1IP1	PDZK1 interacting protein 1	LOC100522145,MCOLN2
DRGA0006924	6	132658334	1.10E-07	0.005449763		ENSSSCG00000023754	45971	LRRC7	leucine rich repeat containing 7	
DRGA0006951	6	134411824	7.30E-08	0.003605549		ENSSSCG00000003802	29042	SLC35D1	solute carrier family 35 (UDP-glucuronic acid/UDP-N-acetylgalactosamine dual transporter), member D1	LOC100739104,C1ORF141,LOC100519721
M1GA0009091	6	155139641	9.80E-08	0.004840097		ENSSSCG00000021630	61610	LOC100737128		HYI,U6,SZT2,LOC100517478
M1GA0009118	6	155404189	3.32E-07	0.016408493		ENSSSCG00000003950	82461	LOC100627013	tyrosine-protein kinase receptor Tie-1-like	TIE-1,TMEM125,LOC100520256,LOC100517766,C1ORF210,MPL,CH242-210N13.2,LOC100519781,LOC100737741,CDC20,LOC100519374,ELOVL1,MED8,LOC100520438,EBNA1BP2,WDR65
M1GA0009342	7	1611710	1.22E-07	0.006014927		ENSSSCG00000028777	62661	LOC100157526	uncharacterized LOC100157526	MYLK4
M1GA0009339	7	1780404	9.80E-08	0.004840097		ENSSSCG00000028777	90238	LOC100157526	uncharacterized LOC100157526	MYLK4,SERPINB1,NQO2,HMSD,LOC100155921,WRNIP1
M1GA0009374	7	3571588	9.80E-08	0.004840097		ENSSSCG00000001016	23283	FARS2		
M1GA0009384	7	3699246	5.27E-07	0.02602384		ENSSSCG00000027136	3170	LOC100737220		F13A1,NRN1
M1GA0009455	7	5113060	9.80E-08	0.004840097		ENSSSCG00000001025	91839	LOC100156744	desmoplakin-like	DSP,BMP6,SNRNP48,LOC100152326
M1GA0009471	7	5307401	2.34E-07	0.011547865		ENSSSCG00000001027	INTRAgenic	BMP6	bone morphogenetic protein 6	EEF1E1

M1GA0009495	7	7536565	2.34E-07	0.011547865		ENSSSCG00000001036	INTRAgenic	TFAP2A	transcription factor AP-2 alpha (activating enhancer binding protein 2 alpha)	
M1GA0009500	7	7664403	1.27E-08	0.000627821		ENSSSCG00000001039	32601	LOC100738362	N-acetyllactosaminide beta-1,6-N-acetylglucosaminyl-transferase, isoform C-like	GCNT2,PAK1IP1,LOC100270682,LOC100156606,C6ORF52
M1GA0009515	7	8372209	2.70E-07	0.013363862		ENSSSCG00000001046	57304	LOC100738669	enhancer of filamentation 1-like	NEDD9,NEDD9
M1GA0009527	7	9302078	2.60E-07	0.012871871		ENSSSCG00000030526	1024656	SIRT5		
M1GA0009555	7	10031634	1.49E-07	0.007341401		ENSSSCG00000001054	52139	LOC100739137	glucose-fructose oxidoreductase domain-containing protein 1-like	GFOD1,PHACTR1,TBC1D7
M1GA0009568	7	10543393	5.71E-08	0.002821284		ENSSSCG00000001057	5345	CCDC90A	coiled-coil domain-containing protein 90A, mitochondrial-like	LOC100520967,CD83,LOC100153365
M1GA0009599	7	11606341	6.09E-07	0.030079184		ENSSSCG00000001057	1030372	CCDC90A		coiled-coil domain-containing protein 90A, mitochondrial-like
M1GA0009677	7	21031127	1.41E-07	0.006974312		ENSSSCG00000001099	1162	CMAHP	cytidine monophosphate-N-acetylneuraminic acid hydroxylase	CMAH
M1GA0009735	7	25272090	4.20E-08	0.00207369		ENSSSCG00000001248	79658	MOG		LOC100626789,KRAB,UBD,LOC100514715,ZFP57,OLF42-2,OLF42-3,GABBR1,LOC100738146,LOC100158165,LOC100626885,OLF42-1
M1GA0009865	7	33358569	1.16E-07	0.005752726		ENSSSCG00000001498	85410	BEND6	BEN domain-containing protein 6-like	LOC100153820
M1GA0010028	7	37366668	5.42E-08	0.002676215		ENSSSCG00000022111	99505	CDKN1A	cyclin-dependent kinase inhibitor 1-like	LOC100623143,C6ORF89,CPNE5,PPIL1,RAB44,LOC100524445,LOC100156227
M1GA0010098	7	39525131	1.54E-06	0.076298247		ENSSSCG00000001588	INTRAgenic	DNAH8	dynein, axonemal, heavy chain 8	GLP1R
ALGA0040777	7	41624144	9.12E-09	0.000450533		ENSSSCG00000027922	6957	LOC100518497		LOC100157694,LOC100518202,NFYA,LOC100152461,LOC100627430,LOC100155551,LOC100156477,LOC100517350,TSPO2
M1GA0010339	7	54693725	1.84E-06	0.090773838		ENSSSCG00000001780	97391	LOC100623036	fumarylacetoacetase-like	FAH,

M1GA0010444	7	62355933	1.42E-06	0.070216839	BFT	ENSSSCG00000001871	94216	HMG20A	high mobility group protein 20A-like	LOC100737571,LINGO1
M1GA0010490	7	87714516	1.42E-06	0.070216839	BFT	ENSSSCG00000026370	1054430	RYR3		
M1GA0010637	7	101888771	1.21E-06	0.059901053	BFT	ENSSSCG00000002333	INTRAgenic			
M1GA0010653	7	103637930	1.21E-06	0.059901053	BFT	ENSSSCG00000002366	55307	NPC2	Niemann-Pick disease, type C2	LTBP2,ISCA2,KIAA0317,LOC100514300,LOC100622552,LOC100738266
ALGA0044087	7	106203235	1.34E-06	0.066068296		ENSSSCG00000002392	49098	LOC100152768	interferon regulatory factor 2-binding protein-like	IRF2BPL,LOC100158012,AHSA1,KIAA1737,LOC100157174,ZDHHC22,LOC100626607
M1GA0010789	7	119600471	1.44E-06	0.071035842		ENSSSCG00000002440	INTRAgenic	CCDC88C	coiled-coil domain containing 88C	SMEK1
ALGA0045246	7	123997457	4.86E-07	0.024000151		ENSSSCG00000002495	53535	SYNE3	nesprin-3-like	LOC100525411,SCARNA13,GLRX5,LOC100154744
M1GA0011167	7	128674568	1.20E-06	0.059130539		ENSSSCG00000002512	80924	DEGS2	uncharacterized LOC100739329	LOC100739329,LOC100513744,SLC25A47,WARS,LOC100515556,ssc-mir-345-2,WDR25,LOC100513936,SLC25A29
M1GA0011231	7	129041116	2.95E-07	0.014563586		ENSSSCG00000002525	34346	LOC100156979	TNF receptor-associated factor 3-like	TRAF3_TV1,RCOR1
M1GA0011339	7	129926876	2.05E-07	0.010128984		ENSSSCG00000002540	INTRAgenic	PPP2R5C		LOC100737027
M1GA0011355	7	130171175	4.46E-07	0.022049006		ENSSSCG00000002542	32213			
M1GA0011382	7	130525474	9.49E-08	0.00468922		ENSSSCG00000029320	1013453			olfactory receptor 4F6-like LOC100524571
M1GA0011391	7	130844451	2.86E-07	0.014120283		ENSSSCG00000002544	19986	PPP1R13B		ZFYVE21,XRCC3,LOC100523621
M1GA0011538	7	134236731	1.83E-06	0.090567997		ENSSSCG00000002622	17730	TMEM14A	transmembrane protein 14A	LOC100526118,GSTA2,LOC100511647
M1GA0011548	8	723623	4.54E-08	0.002242144		ENSSSCG00000008681	96776	WHSC2	Wolf-Hirschhorn syndrome candidate 2	C4ORF48,LOC100515739,ZFYVE28,LOC100739155,POLN,MXD4
M1GA0011615	8	1867618	2.60E-07	0.012831711		ENSSSCG00000008701	24915	LRPAP1	low density lipoprotein receptor-related protein associated protein 1	LREAP1,DOK7,LOC100736823
ALGA0047898	8	56440088	7.30E-08	0.003605549		ENSSSCG00000008893	1000471	REST		RE1-silencing transcription factor-like
MARC0048950	8	70991761	3.49E-07	0.017251505		ENSSSCG00000024160	57187	UGT2A3		LOC100624541,LOC100624700,SULT1B1
M1GA0012006	8	102135908	5.42E-07	0.026801179		ENSSSCG00000009072	1428364	PGRMC2		progesterone receptor membrane component 2
ALGA0048976	8	112525755	1.16E-07	0.005752726		ENSSSCG00000009111	INTRAgenic	SYNPO2	synaptopodin 2	SEC24D
ASGA0042165	9	27635662	2.32E-08	0.001145629		ENSSSCG00000022342	1118578			elongation factor 1-alpha, oocyte form-like

DIAS0002588	9	50852374	1.24E-06	0.061509266		ENSSSCG00000023777	94620	ATP5L	ATP synthase, H+ transporting, mitochondrial Fo complex, subunit G	TREH,IFT46,MLL,TMEM25,LOC100622760,ARCN1,LOC100622487,TTC36
ALGA0052956	9	52177218	3.19E-07	0.015743461		ENSSSCG00000015125	87482	LOC100522562	poliovirus receptor-related protein 1-like	
MARC0049000	9	115450964	2.44E-07	0.012044145		ENSSSCG00000015426	1117782	RELN		
ALGA0055314	9	139257822	1.58E-07	0.007825715		ENSSSCG00000019156	26534	U6		
ALGA0057214	10	15467849	6.62E-08	0.003272114		ENSSSCG00000028881	81000	CNIH3		U3
ALGA0057938	10	28938242	4.87E-07	0.024060542		ENSSSCG00000010923	87988	LGR6	leucine-rich repeat containing G protein-coupled receptor 6	LOC100525658,UBE2T,PPP1R12B
MARC0074336	10	55089008	6.73E-07	0.033257404		ENSSSCG00000011058	90399	GPR158		
MARC0048895	11	1257092	3.44E-07	0.016982348		ENSSSCG00000025308	1063173	IL17D		interleukin-17D-like
ASGA0049706	11	9611445	2.16E-07	0.010671976		ENSSSCG00000023699	1014666			
ALGA0061389	11	23330891	5.09E-07	0.025143796		ENSSSCG00000022876	16501			
MARC0048926	11	54768013	3.39E-07	0.01673839		ENSSSCG00000009477	50012	EDNRB	endothelin receptor type B	
MARC0048937	11	73781506	4.32E-07	0.02134773		ENSSSCG00000009506	16152	RAP2A	RAP2A, member of RAS oncogene family	LOC100155367,RAP2A
H3GA0032476	11	77940001	2.07E-08	0.001022334		ENSSSCG00000009527	99934	FGF14		
MARC0092941	12	48747724	1.43E-06	0.070766584		ENSSSCG00000029262	64596	LOC100517676	Golgi SNAP receptor complex member 1-like	GOSR1,TUSC5,BHLHA9,U6,LOC100521580
ALGA0117762	13	10197528	7.21E-07	0.035623803		ENSSSCG00000023705	117664	U6		
ALGA0069293	13	31572374	1.62E-06	0.080193299		ENSSSCG00000011312	INTRAgenic	LARS2	leucyl-tRNA synthetase 2, mitochondrial	
H3GA0036210	13	36964009	1.19E-07	0.005901058		ENSSSCG00000011420	70931	RBM15B	putative RNA-binding protein 15B-like	LOC100519215,LOC100620322,MANF,VPRBP
H3GA0036239	13	37867796	2.22E-07	0.010950027		ENSSSCG00000011439	88406	PHF7	PHD finger protein 7	NT5DC2,NISCH,STAB1,TNNC1,SEMA3G
DRGA0012382	13	56900313	1.65E-07	0.008156839		ENSSSCG00000028669	57664			
DRGA0012768	13	103171133	6.39E-08	0.003158303		ENSSSCG00000011723	42849	MME	membrane metallo-endopeptidase	
DRGA0012775	13	105219083	1.63E-06	0.080415788		ENSSSCG00000029352	58234	CCNL1	cyclin L1	VEPH1
ALGA0072425	13	159407802	1.05E-07	0.005193013		ENSSSCG00000025709	20127	U6		MORC1

ALGA0078332	14	68085933	6.54E-07	0.03231111		ENSSSCG00000010210	41277	SLC16A9	solute carrier family 16, member 9 (monocarboxylic acid transporter 9)	CCDC6
ALGA0080306	14	107939105	1.85E-07	0.009162404		ENSSSCG00000010431	90844	A1CF	APOBEC1 complementation factor	SGMS1,ASAH2
ALGA0082530	14	142827774	6.11E-07	0.03019421		ENSSSCG00000010699	INTRAgenic	ATE1	arginyltransferase 1	NSMCE4A
ALGA0087587	15	140046718	4.88E-07	0.024103854	BFT	ENSSSCG00000016237	INTRAgenic	DOCK10	dedicator of cytokinesis 10	
ALGA0088670	16	4016044	1.59E-07	0.007846383		ENSSSCG00000029792	16110	TRIO		
ALGA0088909	16	7471376	4.76E-08	0.002351677		ENSSSCG00000016792	1146830	FAM134B		protein FAM134B-like
MARC0048886	16	11583165	2.57E-07	0.012690948		ENSSSCG00000016802	1043595	CDH10		cadherin 10 type 2 (T2-cadherin)
ALGA0090202	16	35635161	5.33E-07	0.026325697		ENSSSCG00000016912	1077577	PPAP2A		lipid phosphate phosphohydrolase 1-like
ALGA0090834	16	58445833	5.02E-08	0.002481991		ENSSSCG00000017010	41336	FAM196B	protein FAM196B-like	LOC100511664
M1GA0021097	16	59838300	8.72E-09	0.000430737		ENSSSCG00000017012	47859	SLIT3		LOC100512568,PANK3,U6,ssc-mir-103-1,MIR103-1,LOC100512746,FBLL1
ALGA0091161	16	66890995	4.89E-09	0.000241651		ENSSSCG00000017026	10282	LOC100516706	gamma-aminobutyric acid receptor subunit alpha-1-like	
M1GA0021128	16	69042784	1.61E-10	7.94E-06		ENSSSCG00000017039	38294	LOC100519063	tetratricopeptide repeat protein 1-like	TTC1,ADRA1B
M1GA0021136	16	73257168	8.97E-08	0.004432085		ENSSSCG00000017068	1021984	C5ORF4		uncharacterized protein C5orf4 homolog
ALGA0091438	16	73552942	1.09E-07	0.005381917		ENSSSCG00000017070	1036732	HAND1		heart and neural crest derivatives expressed 1
M1GA0021168	16	77794570	1.66E-08	0.000820252		ENSSSCG00000017084	34727	FAT2		SLC36A1
M1GA0021255	16	80842730	2.00E-07	0.009859432		ENSSSCG00000027191	INTRAgenic			
M1GA0021335	16	82055821	4.76E-08	0.002351677		ENSSSCG00000027191	1129166			
M1GA0021378	16	82675301	6.08E-07	0.030039176		ENSSSCG00000017109	INTRAgenic	ADAMTS16	ADAM metallopeptidase with thrombospondin type 1 motif, 16	
M1GA0021462	16	84330710	1.87E-07	0.009233816		ENSSSCG00000017112	1228304	IRX4		iroquois homeobox 4 IRX4
M1GA0021563	16	85843098	2.30E-08	0.001135863	BFT	ENSSSCG00000017116	5430	LPCAT1	lysophosphatidylcholine acyltransferase 1	SLC6A3,LOC100520648,CLPTM1L,SLC6A18,LOC100520811,TERT
ALGA0092903	17	5999136	5.10E-09	0.000251868		ENSSSCG00000006989	53654	MTUS1		U6atac,FGL1

M1GA0021675	17	20631316	9.80E-08	0.004840097		ENSSSCG00000007058	INTRAgenic	PLCB4	1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase beta-4-like	LOC100516581
M1GA0021697	17	28221039	1.76E-07	0.008718417		ENSSSCG00000007084	1122141	BFSP1		
M1GA0021987	17	41156001	6.53E-07	0.03224714		ENSSSCG00000007249	88778	C20ORF112	nucleolar protein 4-like	LOC100511968,LOC100514995,C20ORF112,COMMD7,LOC100738825
ALGA0097816	18	32645078	7.79E-08	0.003848717		ENSSSCG00000016639	1020806	FOXP2		forkhead box P2 FOXP2
ALGA0098168	18	45408799	2.65E-08	0.00130916	BFT	ENSSSCG00000016673	1018328	GHRHR		growth hormone releasing hormone receptor
MARC0073715	GL894509.1	20636	2.17E-07	0.01074429						
DRGA0006472	GL895797.1	8884	2.69E-07	0.013282353						
M1GA0008884	JH118434.1	150095	1.87E-07	0.009233816						
M1GA0009237	JH118859.1	25096	3.56E-07	0.017572743						
H3GA0051832	X	85563058	5.40E-07	0.026665719		ENSSSCG00000031003	149	CH242-442M8.1		

Supplementary table 2

GO term enrichment in the regions surrounding significant SNPs

Category	**Term**	**Count**	**%**	**P Value**	**Genes**
GOTERM_MF_FAT	GO:0005509~calcium ion binding	17	1.239067055393586	0.0025493685773515682	CLCA2, EFCAB6, SCUBE1, FBN1, ANXA1, LRPAP1, SLIT3, PLCB4, LPCAT1, SYN3, DLL4, RYR3, FAT2, PLCG2, RELN, CDH10, CDH11

GOTERM_BP_FAT	GO:0045934~negative regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	10	0.7288629737609329	0.012610400532715563	MEF2C, HTR1B, HAND1, RBM15B, BCL3, SIRT5, REST, PRDM1, ZNF423, FOXP2
GOTERM_BP_FAT	GO:0051172~negative regulation of nitrogen compound metabolic process	10	0.7288629737609329	0.013680405952819817	MEF2C, HTR1B, HAND1, RBM15B, BCL3, SIRT5, REST, PRDM1, ZNF423, FOXP2
GOTERM_BP_FAT	GO:0016481~negative regulation of transcription	9	0.6559766763848397	0.01921275356051075	MEF2C, HAND1, RBM15B, BCL3, SIRT5, REST, PRDM1, ZNF423, FOXP2
GOTERM_BP_FAT	GO:0031327~negative regulation of cellular biosynthetic process	10	0.7288629737609329	0.021573666398263757	MEF2C, HTR1B, HAND1, RBM15B, BCL3, SIRT5, REST, PRDM1, ZNF423, FOXP2
GOTERM_BP_FAT	GO:0009890~negative regulation of biosynthetic process	10	0.7288629737609329	0.02433594280141406	MEF2C, HTR1B, HAND1, RBM15B, BCL3, SIRT5, REST, PRDM1, ZNF423, FOXP2
GOTERM_BP_FAT	GO:0010629~negative regulation of gene expression	9	0.6559766763848397	0.03130660764484612	MEF2C, HAND1, RBM15B, BCL3, SIRT5, REST, PRDM1, ZNF423, FOXP2
GOTERM_BP_FAT	GO:0002312~B cell activation during immune response	2	0.1457725947521866	0.03606437988578248	PLCG2, BCL3
GOTERM_BP_FAT	GO:0002313~mature B cell differentiation during immune response	2	0.1457725947521866	0.03606437988578248	PLCG2, BCL3
GOTERM_BP_FAT	GO:0051259~protein oligomerization	5	0.36443148688046645	0.0388024606470939	CCDC88C, SCUBE1, ADSL, KCNA5, OLFM1
GOTERM_BP_FAT	GO:0007242~intracellular signaling cascade	16	1.1661807580174928	0.03926592772095401	RAB2A, RAP2A, MAGI3, GHRHR, CARD10, EDNRB, CDC42, HTR1B, PLCB4, CNIH3, PLCG2, CNR2, BCL3, RRAS, PPAP2A, DDAH1
GOTERM_BP_FAT	GO:0002335~mature B cell differentiation	2	0.1457725947521866	0.04312121833760098	PLCG2, BCL3
GOTERM_BP_FAT	GO:0051173~positive regulation of nitrogen compound metabolic process	10	0.7288629737609329	0.046017681504293674	TAF2, MEF2C, ACVRL1, HAND1, BCL3, CAND1, DDAH1, GHRHR, BMP6, ZNF423
GOTERM_BP_FAT	GO:0010558~negative regulation of macromolecule biosynthetic process	9	0.6559766763848397	0.04701143421589153	MEF2C, HAND1, RBM15B, BCL3, SIRT5, REST, PRDM1, ZNF423, FOXP2
GOTERM_BP_FAT	GO:0019216~regulation of lipid metabolic process	4	0.2915451895043732	0.04891761525034585	CDC42, NPC2, PPAP2A, BMP6
GOTERM_BP_FAT	GO:0010628~positive regulation of gene expression	9	0.6559766763848397	0.0625865981808453	TAF2, MEF2C, ACVRL1, HAND1, BCL3, CAND1, PRDM1, BMP6, ZNF423
GOTERM_BP_FAT	GO:0031328~positive regulation of cellular biosynthetic process	10	0.7288629737609329	0.06314785678025409	TAF2, MEF2C, ACVRL1, HAND1, BCL3, CAND1, DDAH1, GHRHR, BMP6, ZNF423
GOTERM_BP_FAT	GO:0006895~Golgi to endosome transport	2	0.1457725947521866	0.06398633428218667	AP2A1, MON2
GOTERM_BP_FAT	GO:0030336~negative regulation of cell migration	3	0.21865889212827988	0.06527026754545807	ACVRL1, DLL4, RRAS
GOTERM_BP_FAT	GO:0010552~positive regulation of specific transcription from RNA polymerase II promoter	3	0.21865889212827988	0.06527026754545807	TAF2, MEF2C, HAND1
GOTERM_BP_FAT	GO:0009891~positive regulation of biosynthetic process	10	0.7288629737609329	0.06787447109340661	TAF2, MEF2C, ACVRL1, HAND1, BCL3, CAND1, DDAH1, GHRHR, BMP6, ZNF423
GOTERM_MF_FAT	GO:0016840~carbon-nitrogen lyase activity	2	0.1457725947521866	0.06984773859000996	HAL, ADSL

GOTERM_BP_FAT	GO:0007507~heart development	5	0.36443148688046645	0.07293983790197968	IRX4, HAND1, SCUBE1, FBN1, NDUFV2
GOTERM_BP_FAT	GO:0040013~negative regulation of locomotion	3	0.21865889212827988	0.07346402157772032	ACVRL1, DLL4, RRAS
GOTERM_BP_FAT	GO:0051271~negative regulation of cell motion	3	0.21865889212827988	0.07767901845368015	ACVRL1, DLL4, RRAS
GOTERM_BP_FAT	GO:0030855~epithelial cell differentiation	4	0.2915451895043732	0.07915693943358637	LAMA3, SCUBE1, ANXA1, KRT4
GOTERM_BP_FAT	GO:0060429~epithelium development	5	0.36443148688046645	0.08510599651261444	LAMA3, SCUBE1, ANXA1, TFAP2A, KRT4
GOTERM_BP_FAT	GO:0045935~positive regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	9	0.6559766763848397	0.08648268629665126	TAF2, MEF2C, ACVRL1, HAND1, BCL3, CAND1, GHRHR, BMP6, ZNF423
GOTERM_BP_FAT	GO:0006869~lipid transport	4	0.2915451895043732	0.09021337262348476	NPC2, OSBPL1A, ANXA1, SLC27A6
GOTERM_MF_FAT	GO:0046872~metal ion binding	41	2.988338192419825	0.09173418358810002	SLC38A2, ACVRL1, EFCAB6, FARS2, ADAMTS16, PDLIM4, MME, REST, KCNA5, PLCB4, LPCAT1, SYN3, PGRMC2, FAT2, DDAH1, PHC1, ZNF423, TAF2, CLCA2, SCUBE1, MGAT4C, FBN1, ANXA1, SIRT5, XPNPEP3, LRPAP1, SLIT3, FOXP2, CDKN1A, C5ORF4, DLL4, RYR3, TRIM32, PLCG2, NDUFV2, RELN, PRDM1, ZNHIT6, PHF7, CDH10, CDH11
GOTERM_MF_FAT	GO:0043621~protein self-association	2	0.1457725947521866	0.09203899979319571	CCDC88C, TRIM32
GOTERM_BP_FAT	GO:0050678~regulation of epithelial cell proliferation	3	0.21865889212827988	0.09525515283220964	EDNRB, KRT4, FOXP2
GOTERM_BP_FAT	GO:0010604~positive regulation of macromolecule metabolic process	11	0.8017492711370262	0.09610025717685595	TAF2, MEF2C, EDNRB, ACVRL1, HAND1, TRIM32, BCL3, CAND1, PRDM1, BMP6, ZNF423
GOTERM_MF_FAT	GO:0015293~symporter activity	4	0.2915451895043732	0.09736253677216272	SLC2A13, SLC38A2, SLC6A12, SLC16A9
GOTERM_BP_FAT	GO:0010596~negative regulation of endothelial cell migration	2	0.1457725947521866	0.09776523367096782	ACVRL1, DLL4

Supplementary Table 3

GO terms annotation clusters related to regions surrounding significant SNPs

Annotation Cluster 1	**Enrichment Score: 1.1304470776476798**				
Category	**Term**	**Count**	**%**	**PValue**	**Genes**
GOTERM_MF_FAT	GO:0005509~calcium ion binding	17	1.239067055393586	0.0025493685773515682	CLCA2, EFCAB6, SCUBE1, FBN1, ANXA1, LRPAP1, SLIT3, PLCB4, LPCAT1, SYN3, DLL4, RYR3, FAT2, PLCG2, RELN, CDH10, CDH11

GOTERM_MF_FAT	GO:0046872~metal ion binding	41	2.988338192419825	0.09173418358810002	SLC38A2, ACVRL1, EFCAB6, FARS2, ADAMTS16, PDLIM4, MME, REST, KCNA5, PLCB4, LPCAT1, SYN3, PGRMC2, FAT2, DDAH1, PHC1, ZNF423, TAF2, CLCA2, SCUBE1, MGAT4C, FBN1, ANXA1, SIRT5, XPNPEP3, LRPAP1, SLIT3, FOXP2, CDKN1A, C5ORF4, DLL4, RYR3, TRIM32, PLCG2, NDUFV2,
GOTERM_MF_FAT	GO:0043169~cation binding	41	2.988338192419825	0.10328080253039482	SLC38A2, ACVRL1, EFCAB6, FARS2, ADAMTS16, PDLIM4, MME, REST, KCNA5, PLCB4, LPCAT1, SYN3, PGRMC2, FAT2, DDAH1, PHC1, ZNF423, TAF2, CLCA2, SCUBE1, MGAT4C, FBN1, ANXA1, SIRT5, XPNPEP3, LRPAP1, SLIT3, FOXP2, CDKN1A, C5ORF4, DLL4, RYR3, TRIM32, PLCG2, NDUFV2,
GOTERM_MF_FAT	GO:0043167~ion binding	41	2.988338192419825	0.12365304686492087	SLC38A2, ACVRL1, EFCAB6, FARS2, ADAMTS16, PDLIM4, MME, REST, KCNA5, PLCB4, LPCAT1, SYN3, PGRMC2, FAT2, DDAH1, PHC1, ZNF423, TAF2, CLCA2, SCUBE1, MGAT4C, FBN1, ANXA1, SIRT5, XPNPEP3, LRPAP1, SLIT3, FOXP2, CDKN1A, C5ORF4, DLL4, RYR3, TRIM32, PLCG2, NDUFV2,
GOTERM_MF_FAT	GO:0046914~transition metal ion binding	21	1.530612244897959	0.7457209721859258	TAF2, ACVRL1, ADAMTS16, PDLIM4, MME, SIRT5, REST, XPNPEP3, FOXP2, CDKN1A, C5ORF4, TRIM32, PGRMC2, NDUFV2, RELN, PRDM1, ZNHIT6, DDAH1, PHF7, PHC1, ZNF423
Annotation Cluster 2	**Enrichment Score: 0.9799829985388513**				
Category	**Term**	**Count**	**%**	**PValue**	**Genes**
GOTERM_BP_FAT	GO:0030336~negative regulation of cell migration	3	0.21865889212827988	0.06527026754545807	ACVRL1, DLL4, RRAS
GOTERM_BP_FAT	GO:0040013~negative regulation of locomotion	3	0.21865889212827988	0.07346402157772032	ACVRL1, DLL4, RRAS
GOTERM_BP_FAT	GO:0051271~negative regulation of cell motion	3	0.21865889212827988	0.07767901845368015	ACVRL1, DLL4, RRAS
GOTERM_BP_FAT	GO:0030334~regulation of cell migration	4	0.2915451895043732	0.12687553771908394	LAMA3, ACVRL1, DLL4, RRAS
GOTERM_BP_FAT	GO:0040012~regulation of locomotion	4	0.2915451895043732	0.1661483424365675	LAMA3, ACVRL1, DLL4, RRAS
GOTERM_BP_FAT	GO:0051270~regulation of cell motion	4	0.2915451895043732	0.16793231657235183	LAMA3, ACVRL1, DLL4, RRAS
Annotation Cluster 3	**Enrichment Score: 0.863680494039049**				
Category	**Term**	**Count**	**%**	**PValue**	**Genes**
GOTERM_BP_FAT	GO:0045934~negative regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	10	0.7288629737609329	0.012610400532715563	MEF2C, HTR1B, HAND1, RBM15B, BCL3, SIRT5, REST, PRDM1, ZNF423, FOXP2

GOTERM_BP_FAT	GO:0051172~negative regulation of nitrogen compound metabolic process	10	0.7288629737609329	0.013680405952819817	MEF2C, HTR1B, HAND1, RBM15B, BCL3, SIRT5, REST, PRDM1, ZNF423, FOXP2
GOTERM_BP_FAT	GO:0016481~negative regulation of transcription	9	0.6559766763848397	0.01921275356051075	MEF2C, HAND1, RBM15B, BCL3, SIRT5, REST, PRDM1, ZNF423, FOXP2
GOTERM_BP_FAT	GO:0031327~negative regulation of cellular biosynthetic process	10	0.7288629737609329	0.021573666398263757	MEF2C, HTR1B, HAND1, RBM15B, BCL3, SIRT5, REST, PRDM1, ZNF423, FOXP2
GOTERM_BP_FAT	GO:0009890~negative regulation of biosynthetic process	10	0.7288629737609329	0.02433594280141406	MEF2C, HTR1B, HAND1, RBM15B, BCL3, SIRT5, REST, PRDM1, ZNF423, FOXP2
GOTERM_BP_FAT	GO:0010629~negative regulation of gene expression	9	0.6559766763848397	0.03130660764484612	MEF2C, HAND1, RBM15B, BCL3, SIRT5, REST, PRDM1, ZNF423, FOXP2
GOTERM_BP_FAT	GO:0051173~positive regulation of nitrogen compound metabolic process	10	0.7288629737609329	0.046017681504293674	TAF2, MEF2C, ACVRL1, HAND1, BCL3, CAND1, DDAH1, GHRHR, BMP6, ZNF423
GOTERM_BP_FAT	GO:0010558~negative regulation of macromolecule biosynthetic process	9	0.6559766763848397	0.04701143421589153	MEF2C, HAND1, RBM15B, BCL3, SIRT5, REST, PRDM1, ZNF423, FOXP2
GOTERM_BP_FAT	GO:0010628~positive regulation of gene expression	9	0.6559766763848397	0.0625865981808453	TAF2, MEF2C, ACVRL1, HAND1, BCL3, CAND1, PRDM1, BMP6, ZNF423
GOTERM_BP_FAT	GO:0031328~positive regulation of cellular biosynthetic process	10	0.7288629737609329	0.06314785678025409	TAF2, MEF2C, ACVRL1, HAND1, BCL3, CAND1, DDAH1, GHRHR, BMP6, ZNF423
GOTERM_BP_FAT	GO:0010552~positive regulation of specific transcription from RNA polymerase II promoter	3	0.21865889212827988	0.06527026754545807	TAF2, MEF2C, HAND1
GOTERM_BP_FAT	GO:0009891~positive regulation of biosynthetic process	10	0.7288629737609329	0.06787447109340661	TAF2, MEF2C, ACVRL1, HAND1, BCL3, CAND1, DDAH1, GHRHR, BMP6, ZNF423
GOTERM_BP_FAT	GO:0045935~positive regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	9	0.6559766763848397	0.08648268629665126	TAF2, MEF2C, ACVRL1, HAND1, BCL3, CAND1, GHRHR, BMP6, ZNF423
GOTERM_BP_FAT	GO:0010604~positive regulation of macromolecule metabolic process	11	0.8017492711370262	0.09610025717685595	TAF2, MEF2C, EDNRB, ACVRL1, HAND1, TRIM32, BCL3, CAND1, PRDM1, BMP6, ZNF423
GOTERM_BP_FAT	GO:0045941~positive regulation of transcription	8	0.5830903790087464	0.11914478657423203	TAF2, MEF2C, ACVRL1, HAND1, BCL3, CAND1, BMP6, ZNF423
GOTERM_BP_FAT	GO:0045892~negative regulation of transcription, DNA-dependent	6	0.43731778425655976	0.11992543559967225	MEF2C, HAND1, SIRT5, REST, PRDM1, FOXP2
GOTERM_BP_FAT	GO:0051253~negative regulation of RNA metabolic process	6	0.43731778425655976	0.12620782934936825	MEF2C, HAND1, SIRT5, REST, PRDM1, FOXP2
GOTERM_BP_FAT	GO:0043193~positive regulation of gene-specific transcription	3	0.21865889212827988	0.13326278784728401	TAF2, MEF2C, HAND1

GOTERM_BP_FAT	GO:0045944~positive regulation of transcription from RNA polymerase II promoter	6	0.43731778425655976	0.1359124951329621	TAF2, MEF2C, HAND1, BCL3, CAND1, BMP6
GOTERM_BP_FAT	GO:0010551~regulation of specific transcription from RNA polymerase II promoter	3	0.21865889212827988	0.1508246947254292	TAF2, MEF2C, HAND1
GOTERM_BP_FAT	GO:0006357~regulation of transcription from RNA polymerase II promoter	9	0.6559766763848397	0.16262924833225423	TAF2, MEF2C, HAND1, BCL3, TFAP2A, CAND1, PRDM1, BMP6, FOXP2
GOTERM_BP_FAT	GO:0010605~negative regulation of macromolecule metabolic process	9	0.6559766763848397	0.16870983592767724	MEF2C, HAND1, RBM15B, BCL3, SIRT5, REST, PRDM1, ZNF423, FOXP2
GOTERM_BP_FAT	GO:0007517~muscle organ development	4	0.2915451895043732	0.20094247576868352	MEF2C, HAND1, NDUFV2, FOXP2
GOTERM_BP_FAT	GO:0010557~positive regulation of macromolecule biosynthetic process	8	0.5830903790087464	0.20233179670638501	TAF2, MEF2C, ACVRL1, HAND1, BCL3, CAND1, BMP6, ZNF423
GOTERM_BP_FAT	GO:0032583~regulation of gene-specific transcription	3	0.21865889212827988	0.2570366481532445	TAF2, MEF2C, HAND1
GOTERM_BP_FAT	GO:0045893~positive regulation of transcription, DNA-dependent	6	0.43731778425655976	0.2706657983663773	TAF2, MEF2C, HAND1, BCL3, CAND1, BMP6
GOTERM_BP_FAT	GO:0051254~positive regulation of RNA metabolic process	6	0.43731778425655976	0.27629431628920675	TAF2, MEF2C, HAND1, BCL3, CAND1, BMP6
GOTERM_BP_FAT	GO:0000122~negative regulation of transcription from RNA polymerase II promoter	4	0.2915451895043732	0.30865337444216523	MEF2C, HAND1, PRDM1, FOXP2
GOTERM_MF_FAT	GO:0016564~transcription repressor activity	4	0.2915451895043732	0.4663860942943846	HAND1, REST, ZNF423, FOXP2
GOTERM_MF_FAT	GO:0003700~transcription factor activity	9	0.6559766763848397	0.5252128623784147	MEF2C, IRX4, HAND1, HMG20A, BCL3, TFAP2A, PRDM1, ZNF423, FOXP2
GOTERM_BP_FAT	GO:0051252~regulation of RNA metabolic process	14	1.0204081632653061	0.5764288133554576	TAF2, MEF2C, IRX4, RBM15B, SIRT5, HMG20A, REST, FOXP2, HAND1, TFAP2A, BCL3, CAND1, PRDM1, BMP6
GOTERM_BP_FAT	GO:0045449~regulation of transcription	19	1.3848396501457727	0.6443983825224116	IRX4, TAF2, MEF2C, ACVRL1, EFCAB6, RBM15B, CCNL1, HMG20A, SIRT5, WHSC2, REST, FOXP2, HAND1, TFAP2A, BCL3, CAND1, PRDM1, BMP6, ZNF423
GOTERM_BP_FAT	GO:0006355~regulation of transcription, DNA-dependent	13	0.9475218658892128	0.6593479004864191	TAF2, MEF2C, IRX4, SIRT5, HMG20A, REST, FOXP2, HAND1, TFAP2A, BCL3, CAND1, PRDM1, BMP6
GOTERM_MF_FAT	GO:0030528~transcription regulator activity	12	0.8746355685131195	0.6788295591254021	MEF2C, IRX4, HAND1, TRIM32, HMG20A, BCL3, TFAP2A, CAND1, REST, PRDM1, ZNF423, FOXP2
GOTERM_BP_FAT	GO:0006350~transcription	15	1.0932944606413995	0.6909529055039787	TAF2, MEF2C, EFCAB6, RBM15B, CCNL1, HMG20A, WHSC2, REST, FOXP2, HAND1, TFAP2A, BCL3, CAND1, PRDM1, ZNF423

GOTERM_MF_FAT	GO:0043565~sequence-specific DNA binding	5	0.36443148688046645	0.7228176064713828	TAF2, MEF2C, IRX4, HAND1, FOXP2
GOTERM_MF_FAT	GO:0003677~DNA binding	13	0.9475218658892128	0.9726553509096934	TAF2, MEF2C, IRX4, HMG20A, REST, TOX3, FOXP2, HAND1, TFAP2A, BCL3, PRDM1, PHC1, ZNF423
Annotation Cluster 4	**Enrichment Score: 0.8451900140388257**				
Category	**Term**	**Count**	**%**	**PValue**	**Genes**
GOTERM_BP_FAT	GO:0051259~protein oligomerization	5	0.36443148688046645	0.0388024606470939	CCDC88C, SCUBE1, ADSL, KCNA5, OLFM1
GOTERM_BP_FAT	GO:0006461~protein complex assembly	7	0.5102040816326531	0.16500194727104764	TAF2, CCDC88C, SCUBE1, ADSL, KCNA5, OLFM1, CARD10
GOTERM_BP_FAT	GO:0070271~protein complex biogenesis	7	0.5102040816326531	0.16500194727104764	TAF2, CCDC88C, SCUBE1, ADSL, KCNA5, OLFM1, CARD10
GOTERM_BP_FAT	GO:0065003~macromolecular complex assembly	8	0.5830903790087464	0.21381507073526243	TAF2, CCDC88C, SCUBE1, ADSL, KCNA5, ZNHIT6, OLFM1, CARD10
GOTERM_BP_FAT	GO:0043933~macromolecular complex subunit organization	8	0.5830903790087464	0.26313232940188563	TAF2, CCDC88C, SCUBE1, ADSL, KCNA5, ZNHIT6, OLFM1, CARD10
Annotation Cluster 5	**Enrichment Score: 0.7212991520659683**				
Category	**Term**	**Count**	**%**	**PValue**	**Genes**
GOTERM_BP_FAT	GO:0006869~lipid transport	4	0.2915451895043732	0.09021337262348476	NPC2, OSBPL1A, ANXA1, SLC27A6
GOTERM_BP_FAT	GO:0010876~lipid localization	4	0.2915451895043732	0.10792633580096045	NPC2, OSBPL1A, ANXA1, SLC27A6
GOTERM_MF_FAT	GO:0008289~lipid binding	4	0.2915451895043732	0.7042132496474445	NPC2, OSBPL1A, PGRMC2, ANXA1
Annotation Cluster 6	**Enrichment Score: 0.6649398774822877**				
Category	**Term**	**Count**	**%**	**PValue**	**Genes**
GOTERM_BP_FAT	GO:0007517~muscle organ development	4	0.2915451895043732	0.20094247576868352	MEF2C, HAND1, NDUFV2, FOXP2
GOTERM_BP_FAT	GO:0014706~striated muscle tissue development	3	0.21865889212827988	0.2164755302705781	HAND1, NDUFV2, FOXP2
GOTERM_BP_FAT	GO:0060537~muscle tissue development	3	0.21865889212827988	0.23264825884879606	HAND1, NDUFV2, FOXP2
Annotation Cluster 7	**Enrichment Score: 0.6570576422251674**				
Category	**Term**	**Count**	**%**	**PValue**	**Genes**
GOTERM_BP_FAT	GO:0006928~cell motion	7	0.5102040816326531	0.1349010049901273	EDNRB, ACVRL1, ANXA1, RELN, PPAP2A, DNAH8, SLIT3
GOTERM_BP_FAT	GO:0048870~cell motility	5	0.36443148688046645	0.18757895851271814	EDNRB, ACVRL1, RELN, PPAP2A, DNAH8
GOTERM_BP_FAT	GO:0051674~localization of cell	5	0.36443148688046645	0.18757895851271814	EDNRB, ACVRL1, RELN, PPAP2A, DNAH8
GOTERM_BP_FAT	GO:0016477~cell migration	4	0.2915451895043732	0.3287298940692783	EDNRB, ACVRL1, RELN, PPAP2A
GOTERM_BP_FAT	GO:0043085~positive regulation of catalytic activity	6	0.43731778425655976	0.3322684943956695	CDC42, EDNRB, RELN, PPAP2A, GHRHR, CARD10

Annotation Cluster 8	**Enrichment Score: 0.5930283907303879**				
Category	**Term**	**Count**	**%**	**PValue**	**Genes**
GOTERM_BP_FAT	GO:0030817~regulation of cAMP biosynthetic process	3	0.21865889212827988	0.16881674343847725	EDNRB, HTR1B, GHRHR
GOTERM_BP_FAT	GO:0030814~regulation of cAMP metabolic process	3	0.21865889212827988	0.17402411010550833	EDNRB, HTR1B, GHRHR
GOTERM_BP_FAT	GO:0030808~regulation of nucleotide biosynthetic process	3	0.21865889212827988	0.1924438748808114	EDNRB, HTR1B, GHRHR
GOTERM_BP_FAT	GO:0030802~regulation of cyclic nucleotide biosynthetic process	3	0.21865889212827988	0.1924438748808114	EDNRB, HTR1B, GHRHR
GOTERM_BP_FAT	GO:0030799~regulation of cyclic nucleotide metabolic process	3	0.21865889212827988	0.2004170823296681	EDNRB, HTR1B, GHRHR
GOTERM_BP_FAT	GO:0007187~G-protein signaling, coupled to cyclic nucleotide second messenger	3	0.21865889212827988	0.20575464441448438	HTR1B, CNR2, GHRHR
GOTERM_BP_FAT	GO:0006140~regulation of nucleotide metabolic process	3	0.21865889212827988	0.20842944196417976	EDNRB, HTR1B, GHRHR
GOTERM_BP_FAT	GO:0019935~cyclic-nucleotide-mediated signaling	3	0.21865889212827988	0.24618442674430993	HTR1B, CNR2, GHRHR
GOTERM_BP_FAT	GO:0019932~second-messenger-mediated signaling	4	0.2915451895043732	0.24704906822673245	EDNRB, HTR1B, CNR2, GHRHR
KEGG_PATHWAY	hsa04080:Neuroactive ligand-receptor interaction	4	0.2915451895043732	0.3550937061475683	EDNRB, HTR1B, CNR2, GHRHR
GOTERM_BP_FAT	GO:0007610~behavior	5	0.36443148688046645	0.4506497595806731	EDNRB, HTR1B, CNR2, RELN, FOXP2
GOTERM_BP_FAT	GO:0007186~G-protein coupled receptor protein signaling pathway	7	0.5102040816326531	0.8405369735691236	EDNRB, HTR1B, GPR158, CNR2, PPAP2A, LGR6, GHRHR
Annotation Cluster 9	**Enrichment Score: 0.5914765044764794**				
Category	**Term**	**Count**	**%**	**PValue**	**Genes**
GOTERM_BP_FAT	GO:0042325~regulation of phosphorylation	7	0.5102040816326531	0.12643234844172455	CDC42, EDNRB, CDKN1A, CCDC88C, RELN, PPAP2A, CARD10
GOTERM_BP_FAT	GO:0051174~regulation of phosphorus metabolic process	7	0.5102040816326531	0.14462184840931444	CDC42, EDNRB, CDKN1A, CCDC88C, RELN, PPAP2A, CARD10
GOTERM_BP_FAT	GO:0019220~regulation of phosphate metabolic process	7	0.5102040816326531	0.14462184840931444	CDC42, EDNRB, CDKN1A, CCDC88C, RELN, PPAP2A, CARD10

GOTERM_BP_FAT	GO:0033674~positive regulation of kinase activity	4	0.2915451895043732	0.23923627254569466	CDC42, RELN, PPAP2A, CARD10
GOTERM_BP_FAT	GO:0051347~positive regulation of transferase activity	4	0.2915451895043732	0.25687179696226664	CDC42, RELN, PPAP2A, CARD10
GOTERM_BP_FAT	GO:0044093~positive regulation of molecular function	7	0.5102040816326531	0.25827587744537667	CDC42, EDNRB, PLCG2, RELN, PPAP2A, GHRHR, CARD10
GOTERM_BP_FAT	GO:0043549~regulation of kinase activity	5	0.36443148688046645	0.26525008928170735	CDC42, CDKN1A, RELN, PPAP2A, CARD10
GOTERM_BP_FAT	GO:0051338~regulation of transferase activity	5	0.36443148688046645	0.2897126033429236	CDC42, CDKN1A, RELN, PPAP2A, CARD10
GOTERM_BP_FAT	GO:0016477~cell migration	4	0.2915451895043732	0.3287298940692783	EDNRB, ACVRL1, RELN, PPAP2A
GOTERM_BP_FAT	GO:0043085~positive regulation of catalytic activity	6	0.43731778425655976	0.3322684943956695	CDC42, EDNRB, RELN, PPAP2A, GHRHR, CARD10
GOTERM_BP_FAT	GO:0045859~regulation of protein kinase activity	4	0.2915451895043732	0.46472563633671266	CDKN1A, RELN, PPAP2A, CARD10
GOTERM_BP_FAT	GO:0045860~positive regulation of protein kinase activity	3	0.21865889212827988	0.4877313898806421	RELN, PPAP2A, CARD10
Annotation Cluster 10	**Enrichment Score: 0.545492467043092**				
Category	**Term**	**Count**	**%**	**PValue**	**Genes**
GOTERM_BP_FAT	GO:0006874~cellular calcium ion homeostasis	4	0.2915451895043732	0.1503624045264695	EDNRB, RYR3, PLCG2, KCNA5
GOTERM_BP_FAT	GO:0055074~calcium ion homeostasis	4	0.2915451895043732	0.1590710719217424	EDNRB, RYR3, PLCG2, KCNA5
GOTERM_BP_FAT	GO:0006875~cellular metal ion homeostasis	4	0.2915451895043732	0.17331813773199284	EDNRB, RYR3, PLCG2, KCNA5
KEGG_PATHWAY	hsa04020:Calcium signaling pathway	4	0.2915451895043732	0.17658812734947682	EDNRB, PLCB4, RYR3, PLCG2
GOTERM_BP_FAT	GO:0055065~metal ion homeostasis	4	0.2915451895043732	0.18976215799082	EDNRB, RYR3, PLCG2, KCNA5
GOTERM_BP_FAT	GO:0051480~cytosolic calcium ion homeostasis	3	0.21865889212827988	0.2137900791406705	EDNRB, PLCG2, KCNA5
GOTERM_BP_FAT	GO:0030005~cellular di-, tri-valent inorganic cation homeostasis	4	0.2915451895043732	0.2314686657186558	EDNRB, RYR3, PLCG2, KCNA5
GOTERM_BP_FAT	GO:0055066~di-, tri-valent inorganic cation homeostasis	4	0.2915451895043732	0.2549025755511811	EDNRB, RYR3, PLCG2, KCNA5
GOTERM_BP_FAT	GO:0030003~cellular cation homeostasis	4	0.2915451895043732	0.28464187910710553	EDNRB, RYR3, PLCG2, KCNA5
GOTERM_BP_FAT	GO:0048878~chemical homeostasis	6	0.43731778425655976	0.320657967430916	EDNRB, NPC2, RYR3, PLCG2, KCNA5, GHRHR
GOTERM_BP_FAT	GO:0055080~cation homeostasis	4	0.2915451895043732	0.3488071744056349	EDNRB, RYR3, PLCG2, KCNA5
GOTERM_BP_FAT	GO:0019725~cellular homeostasis	5	0.36443148688046645	0.4457544607062853	EDNRB, RYR3, PLCG2, TXNRD1, KCNA5

GOTERM_BP_FAT	GO:0042592~homeostatic process	7	0.5102040816326531	0.4730883437752941	EDNRB, NPC2, RYR3, PLCG2, TXNRD1, KCNA5, GHRHR
GOTERM_BP_FAT	GO:0006873~cellular ion homeostasis	4	0.2915451895043732	0.5185414858461462	EDNRB, RYR3, PLCG2, KCNA5
GOTERM_BP_FAT	GO:0055082~cellular chemical homeostasis	4	0.2915451895043732	0.5293212011630171	EDNRB, RYR3, PLCG2, KCNA5
GOTERM_BP_FAT	GO:0050801~ion homeostasis	4	0.2915451895043732	0.579500702807454	EDNRB, RYR3, PLCG2, KCNA5
Annotation Cluster 11	**Enrichment Score: 0.49426242080848154**				
Category	**Term**	**Count**	**%**	**PValue**	**Genes**
GOTERM_BP_FAT	GO:0043069~negative regulation of programmed cell death	6	0.43731778425655976	0.12304759207325056	MEF2C, EDNRB, CDKN1A, PLCG2, ANXA1, BCL3
GOTERM_BP_FAT	GO:0060548~negative regulation of cell death	6	0.43731778425655976	0.1240967951305419	MEF2C, EDNRB, CDKN1A, PLCG2, ANXA1, BCL3
GOTERM_BP_FAT	GO:0043067~regulation of programmed cell death	9	0.6559766763848397	0.2429749331372102	MEF2C, EDNRB, CDKN1A, PLCG2, ANXA1, TRIO, BCL3, PPP1R13B, CARD10
GOTERM_BP_FAT	GO:0010941~regulation of cell death	9	0.6559766763848397	0.24604082159893315	MEF2C, EDNRB, CDKN1A, PLCG2, ANXA1, TRIO, BCL3, PPP1R13B, CARD10
GOTERM_BP_FAT	GO:0043066~negative regulation of apoptosis	5	0.36443148688046645	0.2604060371698833	MEF2C, EDNRB, CDKN1A, ANXA1, BCL3
GOTERM_BP_FAT	GO:0042981~regulation of apoptosis	8	0.5830903790087464	0.3743005745858848	MEF2C, EDNRB, CDKN1A, ANXA1, TRIO, BCL3, PPP1R13B, CARD10
GOTERM_BP_FAT	GO:0006917~induction of apoptosis	4	0.2915451895043732	0.41640251448465404	CDKN1A, TRIO, BCL3, PPP1R13B
GOTERM_BP_FAT	GO:0012502~induction of programmed cell death	4	0.2915451895043732	0.4183632521517583	CDKN1A, TRIO, BCL3, PPP1R13B
GOTERM_BP_FAT	GO:0043065~positive regulation of apoptosis	4	0.2915451895043732	0.6137147204271258	CDKN1A, TRIO, BCL3, PPP1R13B
GOTERM_BP_FAT	GO:0043068~positive regulation of programmed cell death	4	0.2915451895043732	0.6184503638503946	CDKN1A, TRIO, BCL3, PPP1R13B
GOTERM_BP_FAT	GO:0010942~positive regulation of cell death	4	0.2915451895043732	0.621586023054428	CDKN1A, TRIO, BCL3, PPP1R13B
Annotation Cluster 12	**Enrichment Score: 0.48515408229437657**				
Category	**Term**	**Count**	**%**	**PValue**	**Genes**
GOTERM_BP_FAT	GO:0006954~inflammatory response	5	0.36443148688046645	0.21468378388687845	IL17D, SCUBE1, CNR2, ANXA1, BMP6
GOTERM_BP_FAT	GO:0009611~response to wounding	6	0.43731778425655976	0.3468418044411133	IL17D, ACVRL1, SCUBE1, CNR2, ANXA1, BMP6
GOTERM_BP_FAT	GO:0006952~defense response	6	0.43731778425655976	0.47055140538407825	IL17D, SCUBE1, CNR2, ANXA1, BCL3, BMP6
Annotation Cluster 13	**Enrichment Score: 0.4728269607301088**				

Category	Term	Count	%	PValue	Genes
GOTERM_MF_FAT	GO:0003924~GTPase activity	4	0.2915451895043732	0.2391662614784602	RAB2A, CDC42, RAP2A, RRAS
GOTERM_MF_FAT	GO:0005525~GTP binding	5	0.36443148688046645	0.34813661092651593	RAB2A, CDC42, RAP2A, RRAS, DOCK10
GOTERM_MF_FAT	GO:0019001~guanyl nucleotide binding	5	0.36443148688046645	0.366420979393869	RAB2A, CDC42, RAP2A, RRAS, DOCK10
GOTERM_MF_FAT	GO:0032561~guanyl ribonucleotide binding	5	0.36443148688046645	0.366420979393869	RAB2A, CDC42, RAP2A, RRAS, DOCK10
GOTERM_BP_FAT	GO:0007264~small GTPase mediated signal transduction	4	0.2915451895043732	0.38677190707080134	RAB2A, CDC42, RAP2A, RRAS
Annotation Cluster 14	**Enrichment Score: 0.4680945607005565**				
Category	**Term**	**Count**	**%**	**PValue**	**Genes**
GOTERM_BP_FAT	GO:0009152~purine ribonucleotide biosynthetic process	3	0.21865889212827988	0.2111079853904024	ATP5L, ADSL, MON2
GOTERM_BP_FAT	GO:0009260~ribonucleotide biosynthetic process	3	0.21865889212827988	0.22994657268215618	ATP5L, ADSL, MON2
GOTERM_BP_FAT	GO:0009150~purine ribonucleotide metabolic process	3	0.21865889212827988	0.26789856742064233	ATP5L, ADSL, MON2
GOTERM_BP_FAT	GO:0009259~ribonucleotide metabolic process	3	0.21865889212827988	0.29232468257548144	ATP5L, ADSL, MON2
GOTERM_BP_FAT	GO:0006164~purine nucleotide biosynthetic process	3	0.21865889212827988	0.2950343578885873	ATP5L, ADSL, MON2
GOTERM_BP_FAT	GO:0009165~nucleotide biosynthetic process	3	0.21865889212827988	0.3960104955215172	ATP5L, ADSL, MON2
GOTERM_BP_FAT	GO:0006163~purine nucleotide metabolic process	3	0.21865889212827988	0.3960104955215172	ATP5L, ADSL, MON2
GOTERM_BP_FAT	GO:0034404~nucleobase, nucleoside and nucleotide biosynthetic process	3	0.21865889212827988	0.4139667788215734	ATP5L, ADSL, MON2
GOTERM_BP_FAT	GO:0034654~nucleobase, nucleoside, nucleotide and nucleic acid biosynthetic process	3	0.21865889212827988	0.4139667788215734	ATP5L, ADSL, MON2
GOTERM_BP_FAT	GO:0044271~nitrogen compound biosynthetic process	3	0.21865889212827988	0.6916238261238912	ATP5L, ADSL, MON2
Annotation Cluster 15	**Enrichment Score: 0.4139470876715953**				
Category	**Term**	**Count**	**%**	**PValue**	**Genes**
GOTERM_MF_FAT	GO:0008237~metallopeptidase activity	4	0.2915451895043732	0.1811470866754913	TAF2, ADAMTS16, MME, XPNPEP3

GOTERM_MF_FAT	GO:0008233~peptidase activity	6	0.43731778425655976	0.48935480898846284	TAF2, ADAMTS16, PPP2R5C, MME, RELN, XPNPEP3
GOTERM_MF_FAT	GO:0070011~peptidase activity, acting on L-amino acid peptides	5	0.36443148688046645	0.6464032369922802	TAF2, ADAMTS16, MME, RELN, XPNPEP3
Annotation Cluster 16	**Enrichment Score: 0.4055671936694299**				
Category	**Term**	**Count**	**%**	**PValue**	**Genes**
GOTERM_BP_FAT	GO:0031960~response to corticosteroid stimulus	3	0.21865889212827988	0.1283370434122411	CDKN1A, HTR1B, GHRHR
GOTERM_BP_FAT	GO:0048545~response to steroid hormone stimulus	3	0.21865889212827988	0.41141719891003053	CDKN1A, HTR1B, GHRHR
GOTERM_BP_FAT	GO:0009725~response to hormone stimulus	4	0.2915451895043732	0.5058040736387006	CDKN1A, HTR1B, HMGCS2, GHRHR
GOTERM_BP_FAT	GO:0009719~response to endogenous stimulus	4	0.2915451895043732	0.5727766528979135	CDKN1A, HTR1B, HMGCS2, GHRHR
GOTERM_BP_FAT	GO:0010033~response to organic substance	6	0.43731778425655976	0.6131432177529625	TAF2, CDKN1A, HTR1B, HMGCS2, PLCG2, GHRHR
Annotation Cluster 17	**Enrichment Score: 0.4011976284894864**				
Category	**Term**	**Count**	**%**	**PValue**	**Genes**
GOTERM_BP_FAT	GO:0007156~homophilic cell adhesion	3	0.21865889212827988	0.24889602928106225	FAT2, CDH10, CDH11
GOTERM_BP_FAT	GO:0007155~cell adhesion	7	0.5102040816326531	0.4061448086253882	CLCA2, LAMA3, FAT2, RELN, KIRREL2, CDH10, CDH11
GOTERM_BP_FAT	GO:0022610~biological adhesion	7	0.5102040816326531	0.40746694005296724	CLCA2, LAMA3, FAT2, RELN, KIRREL2, CDH10, CDH11
GOTERM_BP_FAT	GO:0016337~cell-cell adhesion	3	0.21865889212827988	0.6031401414347006	FAT2, CDH10, CDH11
Annotation Cluster 18	**Enrichment Score: 0.39697811947069955**				
Category	**Term**	**Count**	**%**	**PValue**	**Genes**
GOTERM_BP_FAT	GO:0048568~embryonic organ development	3	0.21865889212827988	0.3593995162717888	HAND1, TFAP2A, PRDM1
GOTERM_BP_FAT	GO:0001701~in utero embryonic development	3	0.21865889212827988	0.36994659614983694	ACVRL1, HAND1, PRDM1
GOTERM_BP_FAT	GO:0043009~chordate embryonic development	4	0.2915451895043732	0.4378519574861021	ACVRL1, HAND1, TFAP2A, PRDM1
GOTERM_BP_FAT	GO:0009792~embryonic development ending in birth or egg hatching	4	0.2915451895043732	0.4436532859292699	ACVRL1, HAND1, TFAP2A, PRDM1
Annotation Cluster 19	**Enrichment Score: 0.34906099129851875**				

Category	Term	Count	%	PValue	Genes
GOTERM_BP_FAT	GO:0001525~angiogenesis	3	0.21865889212827988	0.2950343578885873	ACVRL1, HAND1, DLL4
GOTERM_BP_FAT	GO:0048514~blood vessel morphogenesis	3	0.21865889212827988	0.4588898034365008	ACVRL1, HAND1, DLL4
GOTERM_BP_FAT	GO:0001568~blood vessel development	3	0.21865889212827988	0.5380825836278453	ACVRL1, HAND1, DLL4
GOTERM_BP_FAT	GO:0001944~vasculature development	3	0.21865889212827988	0.5512214353098084	ACVRL1, HAND1, DLL4
Annotation Cluster 20	**Enrichment Score: 0.33827861592069786**				
Category	**Term**	**Count**	**%**	**PValue**	**Genes**
GOTERM_BP_FAT	GO:0002521~leukocyte differentiation	3	0.21865889212827988	0.24889602928106225	CDC42, PLCG2, BCL3
GOTERM_BP_FAT	GO:0030097~hemopoiesis	3	0.21865889212827988	0.5178919956686858	CDC42, PLCG2, BCL3
GOTERM_BP_FAT	GO:0048534~hemopoietic or lymphoid organ development	3	0.21865889212827988	0.5704402161939721	CDC42, PLCG2, BCL3
GOTERM_BP_FAT	GO:0002520~immune system development	3	0.21865889212827988	0.6031401414347006	CDC42, PLCG2, BCL3
Annotation Cluster 21	**Enrichment Score: 0.330746139358609**				
Category	**Term**	**Count**	**%**	**PValue**	**Genes**
GOTERM_BP_FAT	GO:0006793~phosphorus metabolic process	10	0.7288629737609329	0.2817859555085786	ACVRL1, VRK3, CDC14A, CSNK1E, NDUFV2, TRIO, ATP5L, RELN, PPAP2A, MON2
GOTERM_BP_FAT	GO:0006796~phosphate metabolic process	10	0.7288629737609329	0.2817859555085786	ACVRL1, VRK3, CDC14A, CSNK1E, NDUFV2, TRIO, ATP5L, RELN, PPAP2A, MON2
GOTERM_BP_FAT	GO:0016310~phosphorylation	8	0.5830903790087464	0.3694464731727828	ACVRL1, VRK3, CSNK1E, NDUFV2, TRIO, ATP5L, RELN, MON2
GOTERM_MF_FAT	GO:0004674~protein serine/threonine kinase activity	4	0.2915451895043732	0.6744616113600137	ACVRL1, VRK3, CSNK1E, TRIO
GOTERM_MF_FAT	GO:0004672~protein kinase activity	5	0.36443148688046645	0.7216086785784992	ACVRL1, VRK3, CSNK1E, TRIO, RELN
GOTERM_BP_FAT	GO:0006468~protein amino acid phosphorylation	5	0.36443148688046645	0.7258952725789223	ACVRL1, VRK3, CSNK1E, TRIO, RELN
Annotation Cluster 22	**Enrichment Score: 0.323270384338646**				
Category	**Term**	**Count**	**%**	**PValue**	**Genes**
GOTERM_BP_FAT	GO:0055085~transmembrane transport	7	0.5102040816326531	0.23750043008798177	SLC2A13, RYR3, SLC16A9, ATP5L, AQP6, KCNA5, MON2
GOTERM_BP_FAT	GO:0015672~monovalent inorganic cation transport	4	0.2915451895043732	0.4124750430736195	SLC38A2, ATP5L, KCNA5, MON2
GOTERM_BP_FAT	GO:0006811~ion transport	7	0.5102040816326531	0.4950164709548164	CLCA2, SLC38A2, RYR3, ATP5L, AQP6, KCNA5, MON2
GOTERM_BP_FAT	GO:0006812~cation transport	5	0.36443148688046645	0.5805998668176285	SLC38A2, RYR3, ATP5L, KCNA5, MON2
GOTERM_BP_FAT	GO:0030001~metal ion transport	3	0.21865889212827988	0.8591904076733301	SLC38A2, RYR3, KCNA5

Annotation Cluster 23	**Enrichment Score: 0.24952526164097708**				
Category	**Term**	**Count**	**%**	**PValue**	**Genes**
GOTERM_BP_FAT	GO:0055085~transmembrane transport	7	0.5102040816326531	0.23750043008798177	SLC2A13, RYR3, SLC16A9, ATP5L, AQP6, KCNA5, MON2
GOTERM_BP_FAT	GO:0006811~ion transport	7	0.5102040816326531	0.4950164709548164	CLCA2, SLC38A2, RYR3, ATP5L, AQP6, KCNA5, MON2
GOTERM_MF_FAT	GO:0022838~substrate specific channel activity	4	0.2915451895043732	0.6224917150686553	CLCA2, RYR3, AQP6, KCNA5
GOTERM_MF_FAT	GO:0015267~channel activity	4	0.2915451895043732	0.6458877324709046	CLCA2, RYR3, AQP6, KCNA5
GOTERM_MF_FAT	GO:0022803~passive transmembrane transporter activity	4	0.2915451895043732	0.6475197835482753	CLCA2, RYR3, AQP6, KCNA5
GOTERM_MF_FAT	GO:0022836~gated channel activity	3	0.21865889212827988	0.7141826213444793	CLCA2, RYR3, KCNA5
GOTERM_MF_FAT	GO:0005216~ion channel activity	3	0.21865889212827988	0.8197589877184859	CLCA2, RYR3, KCNA5
Annotation Cluster 24	**Enrichment Score: 0.23573909435136317**				
Category	**Term**	**Count**	**%**	**PValue**	**Genes**
GOTERM_MF_FAT	GO:0000166~nucleotide binding	22	1.6034985422740524	0.25111527915252163	RAB2A, RAP2A, MAGI3, ACVRL1, FARS2, RBM15B, SIRT5, TRIO, LARS2, DNAH8, ACSS3, CDC42, VRK3, A1CF, CSNK1E, SYN3, NDUFV2, SLC27A6, RRAS, TXNRD1, DOCK10, DDX43
GOTERM_MF_FAT	GO:0017076~purine nucleotide binding	17	1.239067055393586	0.47212431784757036	RAB2A, RAP2A, MAGI3, ACVRL1, FARS2, TRIO, LARS2, DNAH8, ACSS3, CDC42, VRK3, CSNK1E, SYN3, RRAS, TXNRD1, DOCK10, DDX43
GOTERM_MF_FAT	GO:0032553~ribonucleotide binding	16	1.1661807580174928	0.5099785447837818	RAB2A, RAP2A, MAGI3, ACVRL1, FARS2, TRIO, LARS2, DNAH8, ACSS3, CDC42, VRK3, CSNK1E, SYN3, RRAS, DOCK10, DDX43
GOTERM_MF_FAT	GO:0032555~purine ribonucleotide binding	16	1.1661807580174928	0.5099785447837818	RAB2A, RAP2A, MAGI3, ACVRL1, FARS2, TRIO, LARS2, DNAH8, ACSS3, CDC42, VRK3, CSNK1E, SYN3, RRAS, DOCK10, DDX43
GOTERM_MF_FAT	GO:0030554~adenyl nucleotide binding	12	0.8746355685131195	0.732365322134755	MAGI3, ACVRL1, VRK3, CSNK1E, FARS2, SYN3, TRIO, TXNRD1, LARS2, DNAH8, ACSS3, DDX43
GOTERM_MF_FAT	GO:0001883~purine nucleoside binding	12	0.8746355685131195	0.7506366028477341	MAGI3, ACVRL1, VRK3, CSNK1E, FARS2, SYN3, TRIO, TXNRD1, LARS2, DNAH8, ACSS3, DDX43
GOTERM_MF_FAT	GO:0001882~nucleoside binding	12	0.8746355685131195	0.7587339389425627	MAGI3, ACVRL1, VRK3, CSNK1E, FARS2, SYN3, TRIO, TXNRD1, LARS2, DNAH8, ACSS3, DDX43
GOTERM_MF_FAT	GO:0005524~ATP binding	11	0.8017492711370262	0.759190262578398	MAGI3, ACVRL1, VRK3, CSNK1E, FARS2, SYN3, TRIO, LARS2, DNAH8, ACSS3, DDX43
GOTERM_MF_FAT	GO:0032559~adenyl ribonucleotide binding	11	0.8017492711370262	0.7739576219675556	MAGI3, ACVRL1, VRK3, CSNK1E, FARS2, SYN3, TRIO, LARS2, DNAH8, ACSS3, DDX43

Annotation Cluster 25	**Enrichment Score: 0.23000803679061255**				
Category	**Term**	**Count**	**%**	**PValue**	**Genes**
GOTERM_BP_FAT	GO:0048193~Golgi vesicle transport	3	0.21865889212827988	0.24889602928106225	RAB2A, AP2A1, MON2
GOTERM_BP_FAT	GO:0016192~vesicle-mediated transport	6	0.43731778425655976	0.4141541632194901	RAB2A, AP2A1, OSBPL1A, VAMP1, MON2, LRPAP1
GOTERM_BP_FAT	GO:0046907~intracellular transport	6	0.43731778425655976	0.5293841394465902	RAB2A, NPC2, AP2A1, RBM15B, BCL3, MON2
GOTERM_BP_FAT	GO:0008104~protein localization	5	0.36443148688046645	0.8939094796636178	RAB2A, CDC42, AP2A1, BCL3, MON2
GOTERM_BP_FAT	GO:0015031~protein transport	4	0.2915451895043732	0.9229421117296501	RAB2A, AP2A1, BCL3, MON2
GOTERM_BP_FAT	GO:0045184~establishment of protein localization	4	0.2915451895043732	0.9258334101673964	RAB2A, AP2A1, BCL3, MON2
Annotation Cluster 26	**Enrichment Score: 0.22229820217935864**				
Category	**Term**	**Count**	**%**	**PValue**	**Genes**
GOTERM_BP_FAT	GO:0008219~cell death	7	0.5102040816326531	0.4312248677697398	MEF2C, MAGI3, ATXN10, FGF14, TRIO, PPP1R13B, GAN
GOTERM_BP_FAT	GO:0016265~death	7	0.5102040816326531	0.4378043332293893	MEF2C, MAGI3, ATXN10, FGF14, TRIO, PPP1R13B, GAN
GOTERM_BP_FAT	GO:0006915~apoptosis	4	0.2915451895043732	0.8229858727541238	MEF2C, MAGI3, TRIO, PPP1R13B
GOTERM_BP_FAT	GO:0012501~programmed cell death	4	0.2915451895043732	0.8306746508804	MEF2C, MAGI3, TRIO, PPP1R13B
Annotation Cluster 27	**Enrichment Score: 0.21879304996446997**				
Category	**Term**	**Count**	**%**	**PValue**	**Genes**
GOTERM_MF_FAT	GO:0016563~transcription activator activity	5	0.36443148688046645	0.4174174401875738	HAND1, TRIM32, TFAP2A, CAND1, ZNF423
GOTERM_MF_FAT	GO:0003713~transcription coactivator activity	3	0.21865889212827988	0.5138715058631418	HAND1, TRIM32, TFAP2A
GOTERM_MF_FAT	GO:0008134~transcription factor binding	4	0.2915451895043732	0.7845210012034998	HAND1, TRIM32, BCL3, TFAP2A
GOTERM_MF_FAT	GO:0003712~transcription cofactor activity	3	0.21865889212827988	0.7921329868828253	HAND1, TRIM32, TFAP2A
Annotation Cluster 28	**Enrichment Score: 0.18787010882748315**				
Category	**Term**	**Count**	**%**	**PValue**	**Genes**
GOTERM_BP_FAT	GO:0030030~cell projection organization	4	0.2915451895043732	0.5076340780741085	CDC42, ATXN10, RELN, SLIT3
GOTERM_BP_FAT	GO:0031175~neuron projection development	3	0.21865889212827988	0.5619714383686281	ATXN10, RELN, SLIT3
GOTERM_BP_FAT	GO:0030182~neuron differentiation	4	0.2915451895043732	0.6262572403993482	CDC42, ATXN10, RELN, SLIT3
GOTERM_BP_FAT	GO:0048666~neuron development	3	0.21865889212827988	0.7137642381294199	ATXN10, RELN, SLIT3
GOTERM_BP_FAT	GO:0000902~cell morphogenesis	3	0.21865889212827988	0.7388648998300662	CDC42, RELN, SLIT3
GOTERM_BP_FAT	GO:0032989~cellular component morphogenesis	3	0.21865889212827988	0.7918478723837578	CDC42, RELN, SLIT3
Annotation Cluster 29	**Enrichment Score: 0.17107495652328142**				

Category	Term	Count	%	PValue	Genes
GOTERM_BP_FAT	GO:0006511~ubiquitin-dependent protein catabolic process	3	0.21865889212827988	0.5314162652980667	TRIM32, PPP2R5C, ATE1
GOTERM_BP_FAT	GO:0043632~modification-dependent macromolecule catabolic process	5	0.36443148688046645	0.6103427941523885	TRIM32, PPP2R5C, CAND1, GAN, ATE1
GOTERM_BP_FAT	GO:0019941~modification-dependent protein catabolic process	5	0.36443148688046645	0.6103427941523885	TRIM32, PPP2R5C, CAND1, GAN, ATE1
GOTERM_BP_FAT	GO:0051603~proteolysis involved in cellular protein catabolic process	5	0.36443148688046645	0.6453675698906989	TRIM32, PPP2R5C, CAND1, GAN, ATE1
GOTERM_BP_FAT	GO:0044257~cellular protein catabolic process	5	0.36443148688046645	0.64927618913584	TRIM32, PPP2R5C, CAND1, GAN, ATE1
GOTERM_BP_FAT	GO:0030163~protein catabolic process	5	0.36443148688046645	0.6733783530202244	TRIM32, PPP2R5C, CAND1, GAN, ATE1
GOTERM_BP_FAT	GO:0044265~cellular macromolecule catabolic process	5	0.36443148688046645	0.7841349388065852	TRIM32, PPP2R5C, CAND1, GAN, ATE1
GOTERM_BP_FAT	GO:0006508~proteolysis	7	0.5102040816326531	0.7932023602325093	ADAMTS16, TRIM32, PPP2R5C, MME, CAND1, GAN, ATE1
GOTERM_BP_FAT	GO:0009057~macromolecule catabolic process	5	0.36443148688046645	0.8307538005575561	TRIM32, PPP2R5C, CAND1, GAN, ATE1
Annotation Cluster 30	**Enrichment Score: 0.14676601041836715**				
Category	**Term**	**Count**	**%**	**PValue**	**Genes**
GOTERM_MF_FAT	GO:0046982~protein heterodimerization activity	3	0.21865889212827988	0.4987027215804495	HAND1, SCUBE1, FOXP2
GOTERM_MF_FAT	GO:0046983~protein dimerization activity	4	0.2915451895043732	0.8150413972652983	HAND1, SCUBE1, TFAP2A, FOXP2
GOTERM_MF_FAT	GO:0042802~identical protein binding	4	0.2915451895043732	0.8926486707703278	HAND1, SCUBE1, ZNHIT6, FOXP2
Annotation Cluster 31	**Enrichment Score: 0.07749853836990704**				
Category	**Term**	**Count**	**%**	**PValue**	**Genes**
GOTERM_BP_FAT	GO:0034613~cellular protein localization	3	0.21865889212827988	0.8076691389563098	CDC42, AP2A1, BCL3
GOTERM_BP_FAT	GO:0070727~cellular macromolecule localization	3	0.21865889212827988	0.8109187704540026	CDC42, AP2A1, BCL3
GOTERM_BP_FAT	GO:0008104~protein localization	5	0.36443148688046645	0.8939094796636178	RAB2A, CDC42, AP2A1, BCL3, MON2

CHAPTER 3

Copy number variants in Italian Large White pigs detected using high density single nucleotide polymorphisms and their association with back fat thickness

G. Schiavo,* M.A. Dolezal,† E. Scotti,* F. Bertolini,* D.G. Calò,‡ G. Galimberti,‡ V. Russo,* and L. Fontanesi,*§ [1]

* Department of Agricultural and Food Sciences, Division of Animal Sciences, University of Bologna, Viale Fanin 46, 40127 Bologna, Italy

† Institut für Populationsgenetik Veterinärmedizinische University Wien, Josef Baumann Gasse 1, 1210 Wien, Austria

‡ Department of Statistical Sciences "Paolo Fortunati", University of Bologna, Via delle Belle Arti 41, 40126 Bologna, Italy

§ Centre for Genome Biology, University of Bologna, Bologna, Italy

Published in Animal Genetics. 2014 Oct;45(5):745-9. doi: 10.1111/age.12180. Epub 2014 Jun 24.

Summary

The aim of this study was to identify copy number variants (CNVs) in Italian Large White pigs and test them for association with back fat thickness (BFT). Within a population of 12,000 performance tested pigs, two groups of animals with extreme and divergent BFT estimated breeding values (EBVs; 147 with negative and 150 with positive EBVs) were genotyped with the Illumina PorcineSNP60 BeadChip. CNVs were detected with PennCNV software. We identified a total of 4,146 CNV events in 170 copy number variation regions (CNVRs) located on 15 porcine

autosomes. Validation of detected CNVRs was carried out 1) by comparing CNVRs already detected by other studies and 2) by semiquantitative fluorescent multiplex (SQFM) PCR of a few CNVRs. Most of CNVRs detected in Italian Large White pigs (71.2%) were already reported in other pig breeds/populations and 82.1% of the CNV events detected by PennCNV were confirmed by SQFM-PCR. For each CNVR we compared the occurrence of CNV events between the pigs of the high and low BFT EBV tails. Sixteen regions showed $P<0.10$, 7 had $P<0.05$, but were not significant after Bonferroni correction (Fisher's exact test). These results indicated that CNVs could explain a limited fraction of the genetic variability of fat deposition in Italian Large White pigs. However, it was interesting to note that one of these CNVRs encompassed the *ZPLD1* gene. In humans a rare CNV event including this gene is associated with obesity. Studies identifying CNVs in pigs could assist in elucidating the genetic mechanisms underlying human obesity.

Text

Copy number variants (CNVs) can be defined as DNA regions that differ in number of copies compared to a reference genome. They represent the major source of genetic variation in mammalian genomes in terms of number of nucleotides involved (Redon *et al.* 2006). CNVs are known to be able to affect both gene expression and regulation; they can further alter gene structure and products with potentially large phenotypic consequences. Several examples of causative roles of CNVs on phenotypic traits have been reported in livestock. For example, copy number variation including coat colour genes affects coat colour in pigs (Johansson Moller *et al.* 1996), sheep (Norris & Whan 2008), and goats (Fontanesi *et al.* 2009). A few studies on copy number variation have been already carried out in the pig by using array comparative genome hybridization (aCGH; Fadista *et al.* 2008; Li *et al.* 2012), high density single nucleotide polymorphisms (SNPs; Ramayo-Caldas *et al.* 2010; Chen *et al.* 2012; Wang J. *et al.* 2012, 2013; Fowler *et al.* 2013; Wang L. *et al.* 2013) and next generation sequencing data (Rubin *et al.* 2012; Paudel *et al.* 2013). We recently published a genome wide association study for back fat thickness (BFT) based on a selective

genotyping approach in Italian Large White pigs with the Illumina PorcineSNP60 BeadChip (Fontanesi *et al.* 2012b). Here, we further analysed these high throughput genotyping data to identify CNVs to produce a first copy number variation map in the Italian Large White pig breed and to evaluate associations between CNVs and BFT.

Investigated animals have been previously described (Fontanesi *et al.* 2012b). Briefly, two groups of Italian Large White gilts with the most extreme and divergent BFT estimated breeding value (EBV), that were individually performance tested at the Central Test Station of the National Pig Breeder Association (ANAS) in the period 1996-2009 and were selected from a population of 12,000 pigs. Selection of the pigs was obtained ranking their BFT EBV and selecting, among the animals related at two-generation levels, only the most extreme gilt (with most positive or most negative EBV for BFT). Average BFT EBV in the negative (153 pigs) and positive (152 pigs) groups were –9.8 ± 1.6 mm and +6.6 ± 2.3 mm, respectively.

Identification of CNVs in this dataset was carried out with PennCNV software (Wang K. *et al.* 2007), after data checking and filtering as follows: (i) SNPs were mapped onto Sscrofa10.2 assembly if probes had identity >94% with only one aligned region as described (Fontanesi *et al.* 2012b); (ii) principal component analysis was employed to check for potential confounding effects (no biases detected); (iii) potential artifacts due to genomic waves were corrected using the genomic_wave.pl option of the PennCNV package; (iv) only animals with log R ratio standard deviation smaller than 0.4 and GC waviness factor smaller than 0.05 were used for the detection of CNV - this stringent quality control and filtering resulted in 147 gilts with negative BFT EBV and 150 gilts with positive BFT EBV; (v) we considered CNVs with a minimum of three and less than 25 consecutive SNPs to minimize the call of false positive events; (vi) we further filtered for CNVs in telomeric regions (spanning 500 kb from the chromosome ends); (vii) only CNVs present in more than 2% of animals were retained for further analysis and annotation, otherwise they were considered as rare mutations, not representative of the BFT EBV tail. After these steps we detected in the 297 pigs that passed the different control steps a total of 4,146 CNV events (165 gains and

3981 losses), with median length of 180.3 kb, occurring in 170 copy number variation regions (CNVRs; Table S1), determined by overlapping events as previously defined (Fontanesi *et al.* 2010, 2011), that were located on 15 out of 18 autosomes (Table S1). CNVRs included 7 'only gain', 161 'only loss' events and 2 'both gain and loss' events. A larger number of losses compared to gains seems derived by the detection tool that is biased towards the identification of loss events as already reported by others (e.g. Wang J. *et al.* 2012; Liu *et al.* 2013). The largest CNVR spanned approximately 1.7 Mb and the shortest covered 25.2 kb. Validation of detected CNVRs was carried out by comparing CNVRs already detected by other studies with those identified in Italian Large White pigs (Table S1) and second by semiquantitative fluorescent multiplex (SQFM) PCR (Fontanesi *et al.* 2009) of a few CNVRs, normalized against a gene (cathepsin H, *CTSH*) that was not included in any CNVR (Russo *et al.* 2008; Table S2). Most of CNVRs detected in this breed (71.2%) were already reported in other studies (Table S1), including the *KIT* gene region on porcine chromosome 8. SQFM PCR confirmed the presence of the called CNV events in 23 out of 28 (82.1%) analysed animals/regions (Table S2). Annotation of CNVRs was carried out by identifying genes occurring within or just 10 kb far from the CNVR borders (as determined by the first or last SNP included in the regions) using BioMart (http://www.ensembl.org/biomart/martview/). The list of official gene names identified is reported in Table S3. No Gene Ontology term was significantly enriched after False Discovery Rate correction (data not shown).

Several studies have already shown that common or low frequency CNVs may affect obesity related traits in humans (e.g. Walters *et al.* 2010; Wang K. *et al.* 2010; Wheeler *et al.* 2013). To evaluate if CNVs could affect BFT in the investigated pig population, we compared CNV calls in the two BFT EBV tails. The total number of CNV events did not differ between the two tails (2,018 in the negative tail and 2,038 in the positive tail). For each CNVR we compared the occurrence of CNV events between the pigs of the high and low BFT EBV tails considered as case and control groups and calculated Fisher's exact tests and Odd Ratios (Table 1). Sixteen regions were nominally significant $P<0.10$, 7 regions had $P<0.05$ (Fisher's exact test). Fourteen of 16 regions

included low frequency CNV events (0.02< minor allele frequency <0.05) and nine of 16 were more frequent in the fatter pigs (Table 1). Odd ratios supported Fisher's exact test results. However, after Bonferroni correction for multiple testing on Fisher's exact test results, none of these regions remained significant. These results indicated that CNVs might have a limited impact in determining fat deposition in Italian Large White pigs even if it should be considered that low frequency markers would need a larger number of subjects to be tested to reach a sufficient power of the experimental design. Nevertheless annotation information supports the need to further investigate the role of these CNVs for BFT (Table 1). For example, it is interesting to mention that a rare copy number deletion including the *ZPLD1* gene has been consistently associated with obesity in children of different ethnicity (Glessner *et al.* 2010). A similar rare CNV event including the porcine *ZPLD1* gene might be associated with increased BFT in pigs (Table 1). The role of this gene is still not clear, but it seems related to the structure of brain capillaries (Gianfrancesco *et al.* 2008). Its role might match what previous studies indicated about neuronal related genes in affecting fat deposition in humans and pigs (Willer *et al.* 2009; Fontanesi *et al.* 2012b).

This study is one of the first studies to investigate association between CNVs and an economic trait in a livestock species. Results obtained indicate that low frequency CNV events could potentially affect fat deposition in *Sus scrofa*. Particularly, the involvement of *ZPLD1* is worth of further investigation. These results might indicate that CNV studies in pigs are useful to understand the genetic factors determining BFT, completing previous studies carried out in the same pig breed (Fontanesi *et al.* 2012a, 2012b). All these results might be relevant to confirm and clarify genetic mechanisms underlying human obesity.

Competing interests

The authors declare they do not have competing interests.

Acknowledgements

We would like to thank Dr. Luca Buttazzoni (CRA) and ANAS for suggestions and for providing data and samples. This study was supported by Italian MiPAAF (INNOVAGEN project) and AGER-HEPIGET (Grant no. 2011-0279) funds.

References

Chen C., Qiao R., Wei R., *et al.* (2012) A comprehensive survey of copy number variation in 18 diverse pig populations and identification of candidate copy number variable genes associated with complex traits. BMC Genomics 13, 733.

Fadista J., Nygaard M., Holm L.E., Thomsen B. & Bendixen C. (2008) A snapshot of CNVs in the pig genome. *PLoS One* **3**, e3916.

Fontanesi L., Beretti F., Riggio V., *et al.* (2009) Copy number variation and missense mutations of the agouti signaling protein (*ASIP*) gene in goat breeds with different coat colours. *Cytogenetic and Genome Research* 126, 333-47.

Fontanesi L., Galimberti G., Calò D.G., *et al.* (2012a) Identification and association analysis of several hundred single nucleotide polymorphisms within candidate genes for back fat thickness in Italian Large White pigs using a selective genotyping approach. *Journal of Animal Science* **90**, 2450-64.

Fontanesi L., Schiavo G., Galimberti G., *et al.* (2012b) A genome wide association study for backfat thickness in Italian Large White pigs highlights new regions affecting fat deposition including neuronal genes. *BMC Genomics* **13**, 583.

Fowler KE, Pong-Wong R, Bauer J, *et al.* (2013) Genome wide analysis reveals single nucleotide polymorphisms associated with fatness and putative novel copy number variants in three pig breeds. *BMC Genomics* **14**, 784.

Gianfrancesco F., Esposito T., Penco S., *et al.* (2008) ZPLD1 gene is disrupted in a patient with balanced translocation that exhibits cerebral cavernous malformations. *Neuroscience* **155**, 345-9.

Glessner J.T., Bradfield J.P., Wang K., *et al.* (2010) A genome-wide study reveals copy number variants exclusive to childhood obesity cases. *American Journal of Human Genetics* **87**, 661-6.

Johansson Moller M., Chaudhary R., Hellmén E., Höyheim B., Chowdhary B. & Andersson L. (1996) Pigs with the dominant white coat color phenotype carry a duplication of the *KIT* gene encoding the mast/stem cell growth factor receptor. *Mammalian Genome* **7**, 822-30.

Li Y., Mei S., Zhang X., *et al.* (2012) Identification of genome-wide copy number variations among diverse pig breeds by array CGH. *BMC Genomics* **13**, 725.

Liu J., Zhang L., Xu L., *et al.* (2013) Analysis of copy number variations in the sheep genome using 50K SNP BeadChip array. *BMC Genomics* **14**, 229.

Norris B.J. & Whan V.A. (2008) A gene duplication affecting expression of the ovine *ASIP* gene is responsible for white and black sheep. *Genome Research* **18**, 1282-93.

Paudel Y., Madsen O., Megens H.J., *et al.* (2013) Evolutionary dynamics of copy number variation in pig genomes in the context of adaptation and domestication. *BMC Genomics* **14**, 449.

Ramayo-Caldas Y., Castelló A., Pena R.N., *et al.* (2010) Copy number variation in the porcine genome inferred from a 60 k SNP BeadChip. *BMC Genomics* **11**, 593.

Redon R., Ishikawa .S, Fitch K.R., *et al.* (2006) Global variation in copy number in the human genome. *Nature* **444**, 444-54.

Rubin C.J., Megens H.J., Martinez Barrio A., *et al.* (2012) Strong signatures of selection in the domestic pig genome. *Proceedings of the National Academy of Sciences of the USA* **109**, 19529-36.

Russo V., Fontanesi L., Scotti E., *et al.* (2008) Single nucleotide polymorphisms in several porcine cathepsin genes are associated with growth, carcass and production traits in Italian Large White pigs. *Journal of Animal Science* **86**, 3300-314.

Walters R.G., Jacquemont S., Valsesia A., *et al.* (2010) A new highly penetrant form of obesity due to deletions on chromosome 16p11.2. *Nature* **463**, 671-5.

Wang J., Jiang J., Fu W., *et al.* (2012) A genome-wide detection of copy number variations using SNP genotyping arrays in swine. *BMC Genomics* **13**, 273.

Wang J., Wang H., Jiang J., *et al.* (2013) Identification of genome-wide copy number variations among diverse pig breeds using SNP genotyping arrays. *PLoS One* **8**, e68683.

Wang K., Li M., Hadley D., *et al.* (2007) PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Research* **17**, 1665-74.

Wang K., Li W.D., Glessner J.T., Grant S.F., Hakonarson H. & Price R.A. (2010) Large copy-number variations are enriched in cases with moderate to extreme obesity. *Diabetes* **59**, 2690-4.

Wang L., Liu X., Zhang L., *et al.* (2013) Genome-wide copy number variations inferred from SNP genotyping arrays using a large white and minzhu intercross population. *PLoS* One **8**, e74879.

Wheeler E., Huang N., Bochukova E.G., *et al.* (2013) Genome-wide SNP and CNV analysis identifies common and low-frequency variants associated with severe early-onset obesity. *Nature Genetics* **45**, 513-7.

Willer C.J., Speliotes E.K., Loos R.J., *et al.* (2009) Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. *Nature Genetics* **41**, 25-34.

Xu Y., Peng B., Fu Y. & Amos C.I. (2011) Genome-wide algorithm for detecting CNV associations with diseases. *BMC Bioinformatics* **12**, 331

Table 1. Association between copy number variants and back fat thickness in Italian Large White pigs

CNVR[1]	Positive tail[2]	Negative tail[3]	Length of the CNVR (bp)	No. of SNP[4]	Odds Ratio[5]	Fisher's exact test (P value)	MAF (%)[6]	Distance from closest gene (bp)[7]	Closest gene (Ensembl)	Closest gene (symbol)
chr13:99619213-100035462	7	0	416249	10	Infinity	0.015	2.36	0	ENSSSCG00000025418	*MED12L*
chr13:100384911-100946188	9	1	561277	11	9.267	0.020	3.37	0	ENSSSCG00000011720	*MBNL1*
chr8:136860516-136915280	6	0	54764	4	Infinity	0.030	2.02	-	-	-
chr13:93744471-94227153	12	3	482682	8	4.157	0.031	5.05	0	ENSSSCG00000024768	*MYBPC2*
chr13:179392747-180090208	1	7	697461	13	0.135	0.035	2.69	0	ENSSSCG00000011996	*POU1F1*
chr11:11250799-12031063	8	1	780264	21	8.181	0.036	3.03	0	ENSSSCG00000021980	-
chr11:63710002-63804815	8	1	94813	4	8.181	0.036	3.03	45066	ENSSSCG00000009488	*SLITRK5*
chr11:39790178-40052324	3	10	262146	7	0.281	0.050	4.38	-	-	-
chr15:123606534-124174987	3	10	568453	17	0.281	0.050	4.38	0	ENSSSCG00000016151	*MAP2*
chr11:78986323-79304081	5	13	317758	9	0.357	0.054	6.06	-	-	-
chr11:44698910-45225077	9	2	526167	12	4.607	0.061	3.7	-	-	-
chr12:31229725-31875247	1	6	645522	13	0.159	0.065	2.36	0	ENSSSCG00000017601	*TOM1L1*
chr13:171997369-172099078	1	6	101709	4	0.159	0.065	2.36	0	ENSSSCG00000020950	*OR5H@*[8]
chr16:30264193-30841632	1	6	577439	18	0.159	0.065	2.36	0	ENSSSCG00000016878	*FGF10*
chr14:12319126-12399361	7	1	80235	4	7.110	0.067	2.69	0	ENSSSCG00000009664	*PTK2B*
chr13:165948194-166261234	12	4	313040	7	3.098	0.069	5.39	0	ENSSSCG00000011947	*ZPLD1*

[1] Copy Number Variation Region including map positions (chromosome: position of the first SNP-position of the last SNP).
[2] No. of CNV events in the positive back fat thickness estimated breeding value tail.
[3] No. of CNV events in the negative back fat thickness estimated breeding value tail.
[4] No. of SNPs included in the CNVR.
[5] Odds Ratio.
[6] Minor Allele Frequency in the whole population.
[7] Zero means that the indicated gene is included partially or completely in the CNVR. When no gene was identified close less than 100 kb from the two sides no information was reported (-).
[8] Gene of the olfactory receptor family.

Supplementary material

Table S1. Copy number variation regions identified in Italian Large White pigs and comparison with results reported by other authors (Ramayo-Caldas *et al.* 2010; Chen *et al.* 2012; Li *et al.* 2012; Rubin *et al.* 2012; Wang J. *et al.* 2012; Paudel *et al.* 2013; Wang L. *et al.* 2013). None of the CNV found by Fadista *et al.* (2008), Fowler *et al.* (2013) and Wang J. *et al.* (2013) was present in our dataset. LiftOver tool http://genome.ucsc.edu/cgi-bin/hgLiftOver) was used to transfer information from Sscrofa9.2 into Sscrofa10.2 genome versions. For studies that used previous *Sus scrofa* genome versions, SNP coordinates were transferred using scripts we developed for this purpose.

CNVR no.	CNVR (chromosome:position of the first SNP-position of the last SNP)	Length (bp)	No. of CNV events in BFT-EBV positive tail	No. of CNV events in BFT-EBV negative tail	Gain/loss in Italian Large White	Ramayo-Caldas *et al.* (2010)	Chen *et al.* (2012)	Li *et al.* (2012)	Rubin *et al.* (2012)	Wang J. *et al.* (2012)	Paudel *et al.* (2013)	Wang L. *et al.* (2013)	Match with CNVR identified in Italian Large White pigs
1	chr1:1901240-1958514	57274	2	2	loss	no	no	no	no	no	no	no	0
2	chr1:192637720-193093994	456274	28	23	loss	no	gain	no	gain	loss	gain	no	1
3	chr1:220102612-220517767	415155	5	7	loss	no	no	no	no	no	gain	no	1
4	chr1:284443528-284486575	43047	10	10	gain	no	no	no	no	loss	gain	no	1
5	chr1:70218968-70512838	293870	6	4	loss	no	gain/loss	no	loss	no	no	no	1
6	chr1:99869938-100080552	210614	6	8	loss	no	no	no	no	no	gain	no	1
7	chr2:114454568-115000618	546050	6	4	loss	no	no	no	no	no	no	no	0
8	chr2:116037401-116578393	540992	16	21	loss	no	no	no	gain	loss	gain	no	1
9	chr2:129950230-130043084	92854	15	6	loss	no	no	no	no	no	no	no	0
10	chr2:15119750-15540036	420286	11	6	gain	no	no	no	gain	no	gain	no	1
11	chr2:46542866-47166444	623578	28	28	loss	no	no	no	no	no	gain	loss	1
12	chr2:8674597-8701281	26684	2	2	loss	no	no	no	no	loss	no	no	1
13	chr5:21238084-21339891	101807	4	5	gain/loss	no	gain/loss	no	gain	no	gain	no	1
14	chr6:121307601-121517271	209670	1	3	loss	no	no	no	no	no	no	no	0
15	chr6:30428217-30511368	83151	3	2	loss	no	no	no	no	no	no	no	0
16	chr6:74605898-74682817	76919	3	3	loss	no	no	no	no	no	no	no	0
17	chr6:85994652-86135123	140471	8	7	loss	no	no	no	gain	no	no	no	1
18	chr7:133400942-133487315	86373	55	46	gain	no	gain/loss	no	gain	no	gain	gain	1
19	chr7:25773312-26018707	245395	2	2	gain/loss	no	no	no	no	no	gain	no	1
20	chr7:76806487-77057967	251480	3	1	loss	no	no	no	gain	no	no	no	1
21	chr8:100629167-100921095	291928	2	2	loss	no	loss	no	no	no	no	no	1

22	chr8:10519687-10545871	26184	8	15	loss	no	no	no	no	no	no	no	0
23	chr8:136860516-137134554	274038	10	2	loss	no	loss	no	no	no	gain	no	1
24	chr8:43651639-43878303	226664	2	6	gain	no	no	no	gain	no	gain	loss	0
25	chr8:85280494-85554837	274343	14	13	loss	no	no	no	no	no	no	no	0
26	chr9:18417346-18654588	237242	4	0	loss	no	no	no	no	no	gain	no	1
27	chr9:24548249-24662816	114567	1	3	loss	no	no	no	no	loss	no	no	1
28	chr9:93179067-93298999	119932	1	3	loss	no	no	no	no	no	no	no	0
29	chr10:19203770-19266049	62279	8	4	loss	no	no	no	no	no	no	no	1
30	chr10:6048687-6116151	67464	3	2	loss	no	loss	no	no	no	gain	no	1
31	chr10:60836899-61521348	684449	16	20	loss	no	no	no	no	no	no	no	0
32	chr10:62288178-62441849	153671	4	5	loss	no	no	no	no	no	no	no	0
33	chr11:11250799-12031063	780264	8	1	loss	no	loss	no	no	no	no	no	1
34	chr11:29592086-29994790	402704	10	11	loss	no	gain	gain	no	no	gain	no	1
35	chr11:30832976-31605299	772323	13	17	loss	no	no	no	no	loss	gain	no	1
36	chr11:32619321-32878292	258971	20	20	loss	no	gain/loss	no	no	no	no	loss	1
37	chr11:3834824-4333131	498307	6	5	loss	no	no	loss	loss	no	no	no	1
38	chr11:39790178-41062108	1271930	37	40	loss	no	loss	no	gain	no	gain	no	1
39	chr11:42642972-42745437	102465	6	6	loss	no	no	no	no	no	no	no	0
40	chr11:44698910-45910375	1211465	35	32	loss	no	loss	no	loss	no	gain	no	1
41	chr11:49681894-49765012	83118	8	7	loss	no	loss	no	no	no	no	no	1
42	chr11:51024403-51401700	377297	11	7	loss	no	loss	no	no	no	no	no	1
43	chr11:59423296-60652758	1229462	13	13	loss	no	no	no	gain	no	gain	no	1
44	chr11:62514730-62873167	358437	6	4	loss	no	no	no	no	no	no	loss	1
45	chr11:63629757-64065725	435968	31	27	loss	no	gain/loss	no	no	no	gain	no	1
46	chr11:64629699-64928776	299077	4	8	loss	no	gain	gain/loss	no	no	no	no	1
47	chr11:65366338-65806636	440298	17	27	loss	loss	no	gain/loss	no	no	gain	no	1
48	chr11:67671825-68327393	655568	2	5	loss	no	no	loss	no	no	gain	no	1
49	chr11:70656747-70967934	311187	4	4	loss	no	no	no	gain	no	gain	no	1
50	chr11:72031850-72190719	158869	1	3	loss	no	no	no	no	no	no	no	1
51	chr11:72966366-73249203	282837	9	15	loss	no	no	gain	no	no	no	no	1
52	chr11:73719476-74105845	386369	10	12	loss	no	no	no	no	no	no	no	0
53	chr11:75128937-75298434	169497	0	4	loss	no	no	no	no	no	no	no	0
54	chr11:78986323-79567420	581097	10	17	loss	no	no	no	no	loss	no	no	1
55	chr11:8638031-9216774	578743	4	4	loss	no	no	no	no	no	no	no	0
56	chr12:17534131-17667914	133783	2	3	loss	no	no	no	no	loss	no	no	1
57	chr12:18416666-18890652	473986	14	19	loss	no	gain/loss	no	no	no	no	no	1

58	chr12:22730786-23189353	458567	5	5	loss	no	no	no	no	no	no	no	0
59	chr12:23711384-23998154	286770	36	24	loss	no	no	no	no	no	gain	no	1
60	chr12:29022063-29157697	135634	3	3	loss	no	no	no	no	no	no	no	0
61	chr12:30823563-31875247	1051684	27	34	loss	no	loss	gain	gain	no	gain	no	1
62	chr12:33176030-33348283	172253	3	3	loss	no	no	no	no	no	no	no	0
63	chr12:34916437-35494915	578478	18	16	loss	no	no	no	no	no	gain	no	1
64	chr12:36015100-36364110	349010	3	1	loss	no	no	no	no	no	gain	no	1
65	chr12:37172208-37593431	421223	7	6	loss	no	no	gain	no	no	no	no	1
66	chr12:46194492-46323167	128675	9	8	loss	no	no	no	no	loss	no	no	1
67	chr12:47210754-48083984	873230	38	39	loss	no	gain	no	no	no	no	no	1
68	chr12:49872027-50620345	748318	13	11	loss	no	no	no	loss	no	gain	no	1
69	chr12:52828074-53138587	310513	8	12	loss	no	loss	no	no	no	no	no	1
70	chr13:11292387-11720192	427805	4	4	loss	no	gain/loss	no	no	no	no	no	1
71	chr13:114482305-115237462	755157	18	13	loss	no	no	no	no	gain/loss	gain	no	1
72	chr13:116290982-116727358	436376	5	3	loss	no	no	no	loss	no	no	no	1
73	chr13:121795919-122568223	772304	2	2	loss	no	no	no	no	no	gain	no	1
74	chr13:124500733-124569396	68663	2	3	loss	no	gain	no	no	loss	no	no	1
75	chr13:127478577-127657664	179087	2	2	loss	no	no	no	no	no	no	no	0
76	chr13:132017643-132800863	783220	3	6	loss	no	loss	no	no	no	no	no	1
77	chr13:134665329-134933329	268000	10	17	loss	no	no	no	no	no	no	no	0
78	chr13:139751526-139917496	165970	7	5	loss	no	no	no	no	no	no	no	0
79	chr13:141032836-141473317	440481	6	6	loss	no	no	no	gain	loss	no	no	1
80	chr13:142131038-142572873	441835	8	6	loss	no	loss	no	no	no	gain	no	1
81	chr13:143574088-144540823	966735	5	9	loss	no	gain/loss	no	no	no	gain	no	1
82	chr13:145473321-145772101	298780	23	17	loss	no	no	no	no	no	no	no	0
83	chr13:149713321-149910477	197156	3	6	loss	no	no	no	no	loss	no	loss	1
84	chr13:165948194-166379880	431686	26	18	loss	no	no	no	no	no	no	no	0
85	chr13:170944059-171115291	171232	4	1	gain	no	gain	no	gain	no	gain	no	1
86	chr13:171997369-172099078	101709	1	6	loss	no	no	no	no	no	no	gain	1
87	chr13:172999223-173838695	839472	0	4	loss	no	gain/loss	no	loss	no	gain	no	1
88	chr13:174498904-176266015	1767111	7	17	loss	no	gain/loss	no	no	no	gain	no	1
89	chr13:179392747-180090208	697461	1	7	loss	no	no	no	gain	no	no	no	1
90	chr13:183032765-183254699	221934	12	7	loss	no	no	no	no	no	no	no	0
91	chr13:188626778-188780657	153879	2	2	loss	no	no	no	no	loss	no	no	1
92	chr13:197205924-197373950	168026	1	5	loss	no	loss	no	no	no	no	no	1
93	chr13:198238451-198640665	402214	7	10	loss	no	gain	no	gain	no	gain	no	1

94	chr13:200523853-201035957	512104	17	18	loss	no	gain/loss	no	no	no	no	no	1
95	chr13:207397814-207500881	103067	2	2	loss	no	no	no	no	no	no	no	1
96	chr13:213446014-213489485	43471	9	5	gain	no	loss	no	no	loss	no	no	1
97	chr13:85882525-86411969	529444	79	83	loss	no	no	gain	no	no	no	no	1
98	chr13:8697251-8963066	265815	4	1	loss	no	gain	no	no	loss	gain	no	1
99	chr13:87583789-87817848	234059	3	1	loss	no	no	no	no	no	no	no	0
100	chr13:91170267-91528136	357869	17	8	loss	no	no	no	no	no	gain	no	1
101	chr13:91972882-92281356	308474	9	8	loss	no	gain	no	no	gain/loss	gain	no	1
102	chr13:93484102-94603103	1119001	75	59	loss	no	no	no	no	no	gain	no	1
103	chr13:99619213-100946188	1326975	67	46	loss	no	no	no	gain	no	gain	no	1
104	chr14:101124119-101995631	871512	50	44	loss	no	gain	no	gain	no	gain	no	1
105	chr14:102136878-103936325	1799447	24	14	loss	no	gain	no	gain	no	gain	no	1
106	chr14:104441890-104872741	430851	7	10	loss	no	no	no	no	no	no	no	0
107	chr14:106118402-106938671	820269	9	2	loss	no	gain/loss	no	gain	no	no	no	1
108	chr14:110367157-110541284	174127	15	13	loss	no	gain	no	no	no	no	no	1
109	chr14:111970799-112062624	91825	1	3	loss	no	no	no	no	no	no	no	0
110	chr14:113791944-113947139	155195	10	5	loss	no	no	no	no	no	no	no	0
111	chr14:115122766-115670478	547712	14	17	loss	no	gain	no	no	loss	no	no	1
112	chr14:11920612-12399361	478749	31	14	loss	no	gain/loss	no	no	no	no	no	1
113	chr14:122786004-123083682	297678	5	0	loss	no	no	no	no	no	no	no	0
114	chr14:126961187-127266590	305403	5	1	loss	no	no	no	no	no	gain	no	1
115	chr14:129315483-129682891	367408	7	8	loss	no	no	no	no	no	no	loss	1
116	chr14:16504933-16762661	257728	9	12	loss	no	no	no	no	no	no	no	0
117	chr14:19394526-19932281	537755	22	17	loss	no	loss	gain	no	no	gain	no	1
118	chr14:20209157-20912572	703415	8	6	loss	no	gain	no	gain	no	no	no	1
119	chr14:21490308-21791520	301212	4	7	loss	no	no	no	no	no	no	no	0
120	chr14:22171562-22387201	215639	8	13	loss	no	no	no	no	loss	no	no	1
121	chr14:57335272-57644094	308822	10	9	loss	no	gain	no	gain	loss	no	no	1
122	chr14:61110718-61246319	135601	3	1	loss	no	no	no	no	no	no	no	0
123	chr14:66104353-66354543	250190	62	58	loss	no	no	no	no	loss	no	no	1
124	chr14:89402153-89670530	268377	39	48	loss	no	gain	no	no	no	no	no	1
125	chr14:91290555-91513183	222628	2	5	loss	no	no	no	no	no	no	no	0
126	chr15:117636948-118086863	449915	13	12	loss	no	no	no	no	no	gain	no	1
127	chr15:119451280-120409592	958312	25	28	loss	no	loss	no	gain	no	no	no	1
128	chr15:123551835-124412433	860598	18	26	loss	no	no	no	gain	no	no	no	1
129	chr15:125590674-127082018	1491344	22	22	loss	no	gain	no	no	loss	gain	no	1

130	chr15:127761434-127952251	190817	10	7	loss	no	no	no	no	no	no	no	0
131	chr15:130374153-130765546	391393	8	3	loss	no	no	no	no	no	no	no	0
132	chr15:136139958-136195751	55793	4	9	loss	no	no	no	no	no	no	no	0
133	chr15:138046832-138799378	752546	2	3	loss	no	no	no	loss	loss	gain	no	1
134	chr15:140629329-141232228	602899	3	2	loss	no	no	no	no	no	gain	no	1
135	chr15:144911409-145467213	555804	11	9	loss	no	loss	no	no	no	gain	no	1
136	chr15:146882005-147291922	409917	8	4	loss	no	no	no	no	no	no	no	0
137	chr15:21613488-22346975	733487	10	10	loss	no	no	no	no	no	gain	gain	1
138	chr15:24182701-24649487	466786	25	22	loss	no	gain	no	gain	loss	no	no	1
139	chr15:33292968-33549017	256049	24	14	loss	no	no	no	no	gain/loss	gain	no	1
140	chr15:34635612-35158973	523361	23	24	loss	no	no	no	no	loss	gain	no	1
141	chr15:40592400-40754134	161734	5	7	loss	no	no	no	no	no	no	no	0
142	chr15:44744233-45223961	479728	3	5	loss	no	no	no	gain	no	no	no	1
143	chr15:56675162-56937444	262282	3	2	loss	no	no	no	gain	loss	gain	no	1
144	chr15:58049980-59174479	1124499	20	28	loss	no	no	no	no	gain	gain	no	1
145	chr15:63773043-63862810	89767	3	4	loss	no	no	no	no	loss	no	no	1
146	chr15:65421888-65787941	366053	46	48	loss	no	no	no	no	no	gain	no	1
147	chr15:67854572-68081323	226751	34	29	loss	no	no	no	no	gain/loss	no	no	1
148	chr15:73043848-73085844	41996	2	2	loss	no	no	no	no	no	no	no	0
149	chr15:79845945-80481877	635932	33	38	loss	no	no	no	no	no	no	no	0
150	chr15:84349124-85375509	1026385	18	18	loss	no	no	gain	no	no	no	no	1
151	chr16:15245636-15711247	465611	1	3	loss	no	loss	no	no	no	no	no	1
152	chr16:17297501-18134235	836734	15	11	loss	no	no	no	gain	no	no	no	1
153	chr16:2192302-2294773	102471	6	2	loss	no	no	no	no	no	no	no	0
154	chr16:2745009-3194852	449843	8	7	loss	no	gain	no	no	no	gain	no	1
155	chr16:30264193-30841632	577439	3	8	loss	no	no	no	gain	no	no	no	1
156	chr16:57988513-58328528	340015	2	7	loss	no	no	no	no	no	no	no	0
157	chr16:59629920-59988426	358506	5	2	loss	no	no	no	no	no	no	no	0
158	chr16:60209607-60234820	25213	2	2	loss	no	loss	no	no	no	no	no	0
159	chr16:6289550-6470509	180959	2	2	loss	no	no	no	no	no	no	no	0
160	chr16:64085517-64652254	566737	33	27	loss	no	no	no	gain	no	gain	no	1
161	chr16:68669105-69305036	635931	20	20	loss	no	no	no	gain	no	no	no	1
162	chr16:75625628-75880611	254983	19	13	loss	no	no	no	no	no	gain	no	1
163	chr16:77743350-77853044	109694	1	4	loss	no	no	no	no	no	no	no	0
164	chr16:81936588-82017160	80572	0	4	loss	no	no	no	no	no	no	no	0
165	chr16:85253075-85764477	511402	6	4	loss	no	no	no	no	no	no	no	0

166	chr17:14349492-14499321	149829	5	0	gain	no	gain	no	no	no	gain	no	1
167	chr17:3264237-4719022	1454785	37	31	loss	no	no	no	gain	no	no	no	1
168	chr17:6267307-6420022	152715	11	12	loss	no	gain	no	no	no	no	no	1
169	chr17:6930612-7817958	887346	18	25	loss	no	gain	no	gain	no	gain	no	1
170	chr17:9971346-10325411	354065	7	8	loss	no	gain	no	no	loss	no	no	1

Table S2. Validation of copy number variants by semiquantitative fluorescent multiplex (SQFM) PCR.

CNVR no.	**CNVR coordinates (chromosome: nucleotide positions)**	**Amplified region coordinates (chromosome: nucleotide positions)**	**Gene**	**PCR primers[1]**	**PCR conditions[2]**	**PennCNV results (gains/losses)**	**SQFM-PCR results in the analysed pigs – Animal ID (averaged DNA dosage)[3]**
4	chr1:284443528-284486575	chr1:284454230-284454376	salivary lipocalin (*SAL1*)	CCAGTGTGTGGATGTTTTGA CTTTTTCCCCTGTTGTTCCT	147/57/2.0/25	gains	FOG_217 (1.755), FOG_216 (1.626), FOG_353 (1.749), FOG_234 (1.094), FOG_351 (1.501), FOG_274 (1.812), FOG_177 (2.603), FOG_274 (0.461), FOG_38 (2.614), FOG_203 (2.130), FOG_248 (1.268), FOG_346 (1.336), FOG_309 (0.604), FOG_121 (3.250), FOG_120 (0.650), FOG_321 (1.268), FOG_108 (1.396), FOG_169 (1.815), FOG_105 (1.339), FOG_430 (1.429)
13	chr5:21238084-21339891	chr5:21256494-21256646	-	CGAGTAATGAAGAACCAATCG CAGCAGGGTGAGGAGAATAA	153/57/2.0/25	gains/losses	FOG_308 (1.687), FOG_118 (1.718), FOG_378 (1.056), FOG_424 (1.705), FOG_523 (1.514), FOG_164 (0.642),FOG_344 (0.220), FOG_224 (0.705)
Reference gene	-	chr7:53623176-53623352	cathepsin H (*CTSH*)	AATCTTGCCCTGGAGGAAGT GGTTAAAAATCACGCCCAAG[4]	176	-	-

[1] Primer forward was labelled at 5' with 6FAM.

[2] Annealing temperature (°C)/[$MgCl_2$] (mM)/number of PCR cycles/primer concentration (pmol/µL); for the control PCR primers, annealing temperature, [$MgCl_2$] and number of PCR cycles were the same as the CNVR tested primer pairs in the multiplex PCR analyses.

[3] Average DNA dosage relative ratio as determined by SQFM-PCR for the tested pig genomic DNA. Underlined ID of the animals was reported for the animals for which SQFM-PCR results confirmed results of the PennCNV analysis. We adopted the theoretical values of 1.5, 2.0, 2.5, and so on

for a gain of multiple of one, two, three or other copies, respectively, compared to the copy content of averaged values of pigs (23 for the CNV of chromosome 1 and 11 for CNV on chromosome 5) that were not called for CNV with PennCNV. Similarly, a loss of one set of copies (or one copy in case of a simple duplication) would theoretically result in a value of 0.5. Assignment of gain or loss events were based on the closer theoretical values to the averaged DNA dosage obtained by SQFM-PCR of the targeted samples. At least three analyses were carried out for each sample/primer pair combination, and average results were reported. A total of 43 animals for chromosome 1 (20 having CNV according to PennCNV analysis and 2 without any gain or loss in the same analysis) and 18 for chromosome 5 (8 having CNV according to PennCNV analysis and 11 without any gain or loss in the same analysis) were analysed. The normalization for comparing peaks of the amplified region and *CTSH* took into consideration animals without predicted CNV in that region and their average peak values.

[4] Primers for *CTSH* were already reported by Russo *et al.* (2008).

Table S3. Annotation of copy number variation regions (CNVR).

CNVR no.	CNVR (chromosome:position of the first SNP-position of the last SNP)	Gene names (10 kb close)
1	chr1:1901240-1958514	none
2	chr1:192637720-193093994	none
3	chr1:220102612-220517767	none
4	chr1:284443528-284486575	SAL1
5	chr1:70218968-70512838	none
6	chr1:99869938-100080552	none
7	chr2:114454568-115000618	U6
8	chr2:116037401-116578393	none
9	chr2:129950230-130043084	none
10	chr2:15119750-15540036	LOC100520069, LOC100515773, LOC100520237, LOC100520424, LOC100520951, LOC100515945, LOC100519707, LOC100515430, LOC100524068/OR4X1, LOC100515076/LOC100522784, LOC100515256, LOC100518653, LOC100518829, LOC100519014, LOC100515602, LOC100523887, LOC100522977, LOC100523167, LOC100624489
11	chr2:46542866-47166444	INSC
12	chr2:8674597-8701281	ASRGL1/LOC100517023, PHEROC
13	chr5:21238084-21339891	LOC100156296, LOC100157944, LOC100153881, LOC100152670, LOC100156762/OR6C4, LOC100155553
14	chr6:121307601-121517271	none
15	chr6:30428217-30511368	none
16	chr6:74605898-74682817	C1QA, C1QC, LOC100739136, LOC100739136
17	chr6:85994652-86135123	LOC100511641/GRIK3

18	chr7:133400942-133487315	LOC100156611, LOC100155405, OR4K5
19	chr7:25773312-26018707	LOC100152920, LOC100152520/OR5V1, LOC100155347, LOC100154120, LOC100155724, LOC100739010, LOC100739010, LOC100152090, LOC100158163, LOC100156119, LOC100512543
20	chr7:76806487-77057967	none
21	chr8:100629167-100921095	none
22	chr8:10519687-10545871	CC2D2A
23	chr8:136860516-137134554	none
24	chr8:43651639-43878303	U6
25	chr8:85280494-85554837	LOC100620970, LOC100620970
26	chr9:18417346-18654588	U2
27	chr9:24548249-24662816	none
28	chr9:93179067-93298999	AGMO
29	chr10:19203770-19266049	none
30	chr10:6048687-6116151	none
31	chr10:60836899-61521348	5S_rRNA, ITGB1
32	chr10:62288178-62441849	none
33	chr11:11250799-12031063	NBEA, MAB21L1, U4, LOC100738929
34	chr11:29592086-29994790	none
35	chr11:30832976-31605299	none
36	chr11:32619321-32878292	none
37	chr11:3834824-4333131	GPR12, USP12, RPL21, SNORD102, SNORA27, LOC100736950, LOC100524642/SSC.79080, GTF3A
38	chr11:39790178-41062108	U6
39	chr11:42642972-42745437	none
40	chr11:44698910-45910375	none
41	chr11:49681894-49765012	LOC100739028, SNORA70
42	chr11:51024403-51401700	none
43	chr11:59423296-60652758	5S_rRNA, SLITRK1
44	chr11:62514730-62873167	none
45	chr11:63629757-64065725	SLITRK5
46	chr11:64629699-64928776	none
47	chr11:65366338-65806636	none
48	chr11:67671825-68327393	U6
49	chr11:70656747-70967934	LOC100620159,
50	chr11:72031850-72190719	DZIP1, DNAJC3, LOC100155812
51	chr11:72966366-73249203	HS6ST3
52	chr11:73719476-74105845	LOC100155367, CISD1

53	chr11:75128937-75298434	DOCK9
54	chr11:78986323-79567420	none
55	chr11:8638031-9216774	LOC100621650, LOC100153603, LOC100154012, LOC100154396, U6
56	chr12:17534131-17667914	NSF, WNT3
57	chr12:18416666-18890652	MAP3K14, SPATA32, FMNL1, HEXIM2, HEXIM1, ACBD4, PLCD3, NMT1, DCAKD, C1QL1, KIF18B, GFAP, FAM187A, CCDC103, EFTUD2, HIGD1B, GJD3/GJC1, ADAM11
58	chr12:22730786-23189353	CSF3, LOC100736580, SNORD124, CSF3, LOC100512253, LOC100625993, LOC100516107, LOC100512626, LOC100513041, ZPBP2, IKZF3, GRB7, MIEN1, ERBB2, PPP1R1B, STARD3, TCAP, PNMT, PGAP3, LOC100513430
59	chr12:23711384-23998154	LOC414413, KPNB1, TBKBP1, TBX21, OSBPL7, MRPL10, LOC100519171/LRRC46, SCRN2, SP6, PNPO
60	chr12:29022063-29157697	none
61	chr12:30823563-31875247	U6, TOM1L1, LOC100739503
62	chr12:33176030-33348283	CH242-289H7.1
63	chr12:34916437-35494915	LOC100737430, CUEDC1, VEZF1, LOC100737430, LOC100737474, SRSF1, LOC100737591, LOC100512860/DYNLL2, LOC100625111
64	chr12:36015100-36364110	C17orf47, SEPT4, MTMR4, HSF5, RNF43, SUPT4H1, MIR142/ssc-mir-142, BZRAP1, MPO, LPO, MKS1, EPX, CH242-73D9.1, CU571372.3, CU571372.2
65	chr12:37172208-37593431	CH242-205G24.7/CLTC, PTRH2/LOC100519562, ssc-mir-21/LOC100524644, ssc-mir-21/MIR21, TUBD1/LOC100520097, RPS6KB1, LOC100524826/RNFT1
66	chr12:46194492-46323167	none
67	chr12:47210754-48083984	TIAF1, PIPOX, SNORA72, NUFIP2, 5S_rRNA, LOC100622172, ABHD15, TP53I13, GIT1, SSH2, ANKRD13B, CORO6
68	chr12:49872027-50620345	RPA1, LOC100511223, LOC100525608, LOC100623699, ssc-mir-132, ssc-mir-212, HIC1, SMG6, U6, PAFAH1B1, CLUH
69	chr12:52828074-53138587	PITPNM3, FAM64A/LOC100521819, AIPL1,
70	chr13:11292387-11720192	UBE2E2
71	chr13:114482305-115237462	ZBBX, SERPINI2, WDR49, PDCD10, SERPINI1
72	chr13:116290982-116727358	MECOM
73	chr13:121795919-122568223	7SK, 7SK, NAALADL2
74	chr13:124500733-124569396	none
75	chr13:127478577-127657664	none
76	chr13:132017643-132800863	LOC100621366, MAGEF1, LOC100737732, EHHADH
77	chr13:134665329-134933329	BCL6
78	chr13:139751526-139917496	LOC100522036/HRASLS, ATP13A5
79	chr13:141032836-141473317	TMEM44, LSG1, FAM43A
80	chr13:142131038-142572873	, LOC100737308, DLG1
81	chr13:143574088-144540823	U2, TNK2, MUC4, MUC20, KIAA0226, SNORD112, FYTTD1, LRCH3, SNORA31, IQCG, LOC100519675, LMLN, U6, OSBPL11, ZBP-89

82	chr13:145473321-145772101	none
83	chr13:149713321-149910477	NR1I2, LOC100154779
84	chr13:165948194-166379880	ZPLD1, SNORA18
85	chr13:170944059-171115291	LOC100627852
86	chr13:171997369-172099078	LOC100621954
87	chr13:172999223-173838695	none
88	chr13:174498904-176266015	none
89	chr13:179392747-180090208	POU1F1/PIT-I, CHMP2B, VGLL3
90	chr13:183032765-183254699	U6
91	chr13:188626778-188780657	none
92	chr13:197205924-197373950	none
93	chr13:198238451-198640665	none
94	chr13:200523853-201035957	SNORA51, U6
95	chr13:207397814-207500881	none
96	chr13:213446014-213489485	IGSF5
97	chr13:85882525-86411969	SOX14
98	chr13:8697251-8963066	none
99	chr13:87583789-87817848	LOC100512499
100	chr13:91170267-91528136	PLS1, TRPC1, PCOLCE2, PAQR9, U2SURP
101	chr13:91972882-92281356	SLC9A9,
102	chr13:93484102-94603103	U6, PLOD2, PLSCR4, PLSCR5
103	chr13:99619213-100946188	MED12L, GPR87, P2RY13, IGSF10, LOC100739184, MBNL1
104	chr14:101124119-101995631	none
105	chr14:102136878-103936325	ZWINT
106	chr14:104441890-104872741	PCDH15
107	chr14:106118402-106938671	PRKG1, CSTF2T
108	chr14:110367157-110541284	PANK1, MIR107/ssc-mir-107, KIF20B
109	chr14:111970799-112062624	PCGF5
110	chr14:113791944-113947139	HHEX
111	chr14:115122766-115670478	7SK, PLCE1, NOC3L, U6
112	chr14:11920612-12399361	LOC100739720, LOC100739720, TRIM35, PTK2B, CHRNA2, EPHX2
113	chr14:122786004-123083682	C10orf76, HPS6, LDB1, PPRC1, NOLC1, ELOVL3
114	chr14:126961187-127266590	U6
115	chr14:129315483-129682891	none
116	chr14:16504933-16762661	ADAM29, GLRA3
117	chr14:19394526-19932281	none

118	chr14:20209157-20912572	none
119	chr14:21490308-21791520	NEK1, SH3RF1
120	chr14:22171562-22387201	DDX60, 7SK
121	chr14:57335272-57644094	ZP4, RYR2
122	chr14:61110718-61246319	none
123	chr14:66104353-66354543	RET, CSGALNACT2, RASGEF1A
124	chr14:89402153-89670530	none
125	chr14:91290555-91513183	SNORA31
126	chr15:117636948-118086863	FAM117B/LOC100512564, ICA1L, WDR12/LOC100512929, CARF/LOC100627358, NBEAL1, NBEAL1
127	chr15:119451280-120409592	U6, LOC100738694, LOC100738793, U6, LOC100515588
128	chr15:123551835-124412433	LOC100522220, MAP2, UNC80
129	chr15:125590674-127082018	LOC100525789
130	chr15:127761434-127952251	none
131	chr15:130374153-130765546	FN1, 5S_rRNA
132	chr15:136139958-136195751	none
133	chr15:138046832-138799378	LOC100738149, CH242-227L8.1, MOGAT1, U6, 5S_rRNA, RAB11FIP5, LOC100625641, KCNE4, ACSL3
134	chr15:140629329-141232228	LOC100620202/NYAP2
135	chr15:144911409-145467213	SP140, LOC100516940, CAB39, ITM2C, GPR55, SPATA3, C2orf72
136	chr15:146882005-147291922	ECEL1, PRSS56, CHRND, CHRNG, EIF4E2, EFHD1, GIGYF2, KCNJ13, LOC100738429
137	chr15:21613488-22346975	,LYPD1
138	chr15:24182701-24649487	U6, LOC100738584
139	chr15:33292968-33549017	none
140	chr15:34635612-35158973	TSN, MKI67, U6, CLASP1, U4atac, 5S_rRNA, CLASP1, TFCP2L1,
141	chr15:40592400-40754134	none
142	chr15:44744233-45223961	VEGFC
143	chr15:56675162-56937444	none
144	chr15:58049980-59174479	none
145	chr15:63773043-63862810	LOC100516114
146	chr15:65421888-65787941	HS6ST1, U6,
147	chr15:67854572-68081323	LOC100738077
148	chr15:73043848-73085844	U2
149	chr15:79845945-80481877	SLC38A11, SCN3A, SCN2A
150	chr15:84349124-85375509	NOSTRIN, SPC25, LOC100518871/G6PC2, ABCB11, LOC100519057, LOC100519223, LOC100156983, FASTKD1, PPIG, CCDC173, PHOSPHO2, KLHL23, SSB, METTL5
151	chr16:15245636-15711247	none
152	chr16:17297501-18134235	none

153	chr16:2192302-2294773	none
154	chr16:2745009-3194852	none
155	chr16:30264193-30841632	FGF10
156	chr16:57988513-58328528	LCP2, C5orf58, LOC100739364, DOCK2
157	chr16:59629920-59988426	SLIT3, MIR218B/ssc-mir-218b, U6, PANK3, ssc-mir-103-1/MIR103-1, FBLL1, U6, RARS
158	chr16:60209607-60234820	TENM2/ODZ2
159	chr16:6289550-6470509	FAM134B, LOC396902/MYO10
160	chr16:64085517-64652254	none
161	chr16:68669105-69305036	CCNJL, FABP6, LOC100518878, TTC1, ADRA1B
162	chr16:75625628-75880611	SNORA18
163	chr16:77743350-77853044	FAT2, SLC36A1
164	chr16:81936588-82017160	none
165	chr16:85253075-85764477	IRX4, MRPL36, NDUFS6
166	chr17:14349492-14499321	SSC.88457,
167	chr17:3264237-4719022	TUSC3, SNORA40, SNORA40, , MSR1
168	chr17:6267307-6420022	PCM1, ASAH1, FRG1
169	chr17:6930612-7817958	none
170	chr17:9971346-10325411	CSGALNACT1, LOC100518097

CHAPTER 4

A retrospective analysis of major gene allele frequency changes during 20 years of selection in the Italian Large White pig breed

L. Fontanesi[1,2], G. Schiavo[1,*], E. Scotti[1,*], G. Galimberti[2], D.G. Calò[2], A.B. Samorè[1], M. Gallo[3], V. Russo[1] & L. Buttazzoni[4]

[1] Department of Agricultural and Food Sciences, Division of Animal Sciences, University of Bologna, Viale Fanin 46, 40127 Bologna, Italy

[2] Department of Statistical Sciences "Paolo Fortunati", Via delle Belle Arti, University of Bologna, 40126 Bologna, Italy

[3] Associazione Nazionale Allevatori Suini, Roma, Italy

[4] Centro di Ricerca per la Produzione delle Carni e il Miglioramento Genetico, Consiglio per la Ricerca e la Sperimentazione in Agricoltura, Monterotondo (Roma), Italy

Published in Journal of Animal Breeding and Genetics. 1 Mar 2015 | doi: 10.1111/jbg.12127.

Summary

In this study we investigated if a selection program based on boar genetic evaluation obtained with a classical BLUP animal model can change, in a quite short period of time, allele frequencies in a pig population. All Italian Large White boars born from 1992 to 2012 with estimated breeding value reliability >0.85 (n. = 200) were selected among all boars of this breed. Boars were genotyped with markers in major genes (*IGF2* intron3-g.3072G>A, *MC4R* p.D298N, *VRTN* PRE1 insertion, *PRKAG3* p.I199V and *FTO* g.276T>G). Genotyping data were analysed grouping boars in eight classes according to their year of birth. To evaluate the influence of time on allele frequencies of the genotyped markers, multinomial logistic regression models were computed. Four out of five

polymorphic sites (*IGF2*, *MC4R*, *VRTN* and *FTO*) showed significant ($P<0.01$) changes in allele frequencies over time due to a progressive and continuous increase of one allele (associated with higher lean meat content, higher average daily gain, and favorable feed:gain ratio) and, consequently, decrease of the other one, following the directional selection of the selection program of this pig breed. The retrospective analysis that was carried out in Italian Large White boars suggests that selection based on methodologies assuming the infinitesimal model is able to modify in a quite short period of time allele frequencies in major genes, increasing the frequency of alleles explaining a relevant (non-infinitesimal) fraction of the overall genetic variability for production traits.

Introduction

Selection in livestock populations has been mainly based on methodologies assuming the infinitesimal model that describes quantitative traits as determined by an infinite number of genes, each of them with a very small effect. By this approach, genetic improvement is not expected to significantly change allele frequencies at any particular locus in a population (Barton & Keightley 2002). However, experiments dissecting genetic variability in livestock genomes have substantially challenged the assumptions of the infinitesimal model showing that some loci explain a relevant part of the genetic variance of quantitative traits under selection. These loci have been termed quantitative trait loci (QTL), or major genes when the causative genes and mutations are known and their effects are particularly relevant. Therefore, directional selection in livestock populations could be expected to change allele frequencies of QTLs and major genes as their effects can be indirectly captured by the breeding programs designed to improve production traits. For this reason, changes in allele frequencies induced by breeding programs could remain as detectable selection signatures.

A large number of studies have already identified QTLs in the pig genome (Rothschild *et al*. 2007). Polymorphisms in several candidate genes have been shown to be the causative mutations underlying some QTL or to be associated with a significant portion of the genetic variance of

economic traits (Ernst & Steibel 2013). In particular, an imprinted QTL with large effects on muscle mass and fat deposition was identified in the telomeric end of the p arm of *Sus scrofa* chromosome (SSC) 2 (Jeon *et al.* 1999; Nezer *et al.* 1999). The causative mutation was identified in a highly conserved regulatory region of intron 3 of the *insulin-like growth factor 2* (*IGF2*) gene, in which the g.3072G>A substitution disrupts a repressor nuclear factor binding site, causing a three-fold over expression of postnatal skeletal muscle *IGF2* mRNA in pigs inheriting the mutation from their sires, leading to increased muscle mass and, in turn, reduced fat deposition (Van Laere *et al.* 2003). We showed that the *IGF2* intron3-g.3072G>A single nucleotide polymorphism (SNP) is highly associated with all traits under selection in the Italian Large White pig population. Allele A was associated with higher lean meat cut weight (LC), ham weight (HW) and average daily gain (ADG), lower back fat thickness (BFT) and favourable feed:gain ratio (FGR) in an additive way (Fontanesi *et al.* 2010b). This SNP was also the most significantly associated with BFT among about 300 candidate gene markers tested in the same pig breed using a selective genotyping approach (Fontanesi *et al.* 2012). The second most significant gene marker in a selective genotyping study carried out in the Italian Large White breed was the p.D298N (c.892G>A) missense mutation of the *melanocortin 4 receptor* (*MC4R*) gene (Fontanesi *et al.* 2012). This mutation changes a residue in a conserved position of the seventh transmembrane domain of the protein and was found to be associated with BFT at 10th rib, daily gain, and feed intake in different pig lines (Kim *et al.* 2000). The effect of this polymorphism was investigated by several subsequent studies that in most cases confirmed previous results with effects depending on the pig lines investigated. Analysis of six polymorphisms combined in four *MC4R* haplotypes in Italian Large White pigs confirmed the potential functional role of the p.D298N mutation (Fontanesi *et al.* 2013). Another putative causative mutation determined by an insertion of 291 bp of a PRE1 repeated sequence (one of the porcine short interspersed nuclear element classes; SINE) into intron 1 of the *vertnin* (*VRTN*) gene was suggested to explain the QTL identified on SSC7 affecting vertebral number (Mikawa *et al.* 2011). Allele *Q* (with the insertion) is associated with increased vertebrae number (Mikawa *et al.*

2011) that, in turn, should positively affect other traits like lean meat production and reproduction parameters (Borchers *et al.* 2004; Duijvesteijn *et al.* 2014). We showed that allele *Q* was mildly associated with lower ham weight (HW) in Italian Large White pigs (Fontanesi *et al.* 2014). Ham weight is relevant for the production of dry-cured hams and although the Italian pig breeders association calculates an estimated breeding value (EBV) for this trait, this parameter is not included in the general selection index of the boars. Meat quality is another relevant trait for the production of dry-cured ham and genes affecting meat quality parameters should be considered. The Italian Large White breed is virtually free from the *ryanodine receptor 1* (*RYR1*) c.1843C>T mutation, determining pale soft exudative meat (Fuji *et al.* 1991; Fontanesi *et al.* 2008), and from the *PRKAG3* p.R200Q mutation, determining the acid meat defect (Milan *et al.* 2000; Fontanesi *et al.* 2003). Other mutations in the *PRKAG3* gene have been shown to affect meat quality parameters (Ciobanu *et al.* 2001). In particular, the p.I199V substitution was associated with meat pH_1 in Italian Large White pigs (Fontanesi *et al.* 2008). A marker in another candidate gene, *fat and mass obesity associated* or *FTO* (g.276T>G, AM931150), was shown to affect intermuscular fat deposition in Italian Duroc and FGR in Italian Large White pigs (Fontanesi *et al.* 2009). Results on fat deposition in heavy pigs have been subsequently confirmed (Fontanesi *et al.* 2010a; Fontanesi & Russo 2013). Other polymorphisms in this gene further supported the association between *FTO* and several production traits in different pig populations (Fan *et al.* 2009). *FTO* is the most important gene associated with common obesity in humans (Frayling *et al.* 2007).

According to the results of the mentioned studies, it might be expected that a selection system based on boar genetic evaluation carried out by classical BLUP animal models was able to change allele frequencies in the pig population over a quite short period of time. To verify this hypothesis we genotyped the five above mentioned gene markers (*IGF2* intron3-g.3072G>A, *MC4R* p.D298N, *VTNR* PRE1 insertion, *PRKAG3* p.I199V and *FTO* g.276T>G) in 200 Italian Large White boars. The boars were born over a 20 year period, they were all proven and approved by the National Pig Breeders Association (ANAS) and their EBV had the highest reliability among all boars in the same

period. Results indicated that allele frequencies at several marker loci changed over the considered period according to the implemented directional artificial selection.

Materials and Methods

Animals and traits

All Italian Large White Boars born from 1992 (two years after the onset of the current selection scheme for this breed) to 2012 (n. = 5983), approved for reproduction based on their genetic merit after evaluation by ANAS, were ranked according to the reliability of their EBVs (calculated in 2012). All boars with reliability >0.85 (n. = 200) were selected for this study. The selected boars were used to constitute eight groups including boars that were born in a 2-4 year period. Averaged data (EBVs and reliability) for each boar group are reported in Table 1. No boars with EBV reliability >0.85 were available in the years 2011 and 2012.

Evaluation of the candidate boars is based on sib-testing as previously described (Fontanesi *et al.* 2008, 2010b, 2013). Briefly, triplets of siblings from the same full-sib litter (two females and one castrated male) of the candidate boar are individually performance tested at the Central Test Station of ANAS. Performance test evaluation period starts at 30 to 45 days of age and it finishes when pigs reach 155 ± 5 kg live weight. Feed intake is recorded daily and body weight is measured bimonthly, then ADG and FGR is calculated. At the end of test, performance tested animals are moved to a commercial slaughterhouse where, after slaughtering, BFT at the level of *Musculus gluteus medius*, weight of LC (necks and loins), and HW are measured. EBVs are calculated for ADG (expressed in g), LC and HW (expressed in kg), BFT (expressed in mm) and FGR using a BLUP-Multiple Trait-Animal Model with different models per each trait. Depending on the trait, models include the fixed effects of sex, batch on trial, inbreeding coefficient of the animal, interaction of sex by age at slaughtering, date of slaughtering and the random effects of litter and animal. Trends of sire EBVs reported in Fig. 1 show the results of the selection program in the Italian Large White population over the 1992-2012 period for the traits of interest in this paper.

During the considered period, selection objectives for these traits were: ADG and LC, increase; FGR, decrease; decrease and then maintenance of BFT; increase and then maintenance of HW.

Genotyping

Blood samples collected from the candidate boars were lyophilized and stored at room temperature till DNA extraction. Extraction was carried out using the NucleoSpin® Blood Kit (Macherey-Nagel GmbH & Co. KG, Düren, Germany). Five DNA markers (*IGF2* intron3-g.3072G>A, *MC4R* p.Asp298Asn, *VTNR* PRE1 insertion, *PRKAG3* p.I199V and *FTO* g.276T>G) were genotyped by PCR-RFLP or fragment length analysis. Genotyping conditions are reported in Table 2. Genotyping products (DNA after restriction enzyme digestion or after amplification) were electrophoresed on TBE1X 1.5%-2.0% agarose gels or 15% polyacrylamide:bisacrylamide 24:1 gels and visualized with 1X GelRed Nucleic Acid Gel Stain (Biotium Inc., Hayward, CA, USA).

Statistical analyses

Statistical analyses were carried out by fitting multinomial logistic regression models (Tutz 2012). In particular, for each DNA marker, the genotype was treated as a categorical dependent variable with three categories depending on allele occurrences; therefore, each multinomial logistic regression model consisted in two separate equations. Moreover, two alternative specifications were considered: one specification assuming time invariance of allele frequencies (i.e. both equations contained only a constant temr), the other one including time as a covariate (i.e. both equations contained a constat term and a regressin coefficient accounting for the effect of time). In the latter specification, time was coded as an integer ranging from 1 to 8, according to group membership described in Table 1. In order to assess the effect of time on allele frequencies, the two specifications were compared by the likelihood ratio test. According to likelihood theory, the null distribution of the test-statistic was approximated by a chi-square distribution with 2 degrees of freedom, since the latter specification contained two additional parameters compared to the first

one. These analyses were performed in R (R Core Team 2013), using the nnet package (Venables & Ripley 2002).

Results and discussion

This study was conducted on boars with the highest EBV reliabilities out of the Italian Large White boars evaluated over 20 years (1992-2012). As EBV reliability in a BLUP animal model is influenced by the number of information coming from relatives, including descendants, the selection criterion used (reliability >0.85) identified the most influencing boars in the Italian Large White population over a period approximately corresponding to 12-14 sire generations. Therefore allele frequencies in these boars are expected to offer a good approximation of the allele frequencies of the whole boar population, and indirectly, of the whole population.

Changes in allele frequencies of the five investigated gene markers in the genotyped Italian Large White boars, divided in eight groups according to their year of birth, are shown in Fig. 2. Table 3 reports statistics of the multinomial logistic regression models for the same markers. Four out of five polymorphic sites showed significant (after Bonferroni correction) changes in allele frequencies over time due to a progressive and continuous increase of one allele and, consequently, decrease of the other one. The only marker that did not change allele frequencies over the monitored period was the *PRKAG3* p.I199V mutation. Allele I, considered to be the positive one in terms of meat quality, was the less frequent in all groups (Fig. 2), confirming our previous survey on this gene polymorphism (Fontanesi *et al.* 2008). Allele frequencies at this mutation did not change in the considered period preobably because it does not significantly directly affect any traits included in the Italian Large White breeding program.

Among the five genes, allele A at the *IGF2* intron3-g.3072G>A SNP increased from 0.650 (1992-1995) to 0.938 (2008-2010). Its constant increase over the eight temporal windows was highly significant (P<0.001). This allele causes higher LM deposition than the alternative allele. Its effect on LC positively correlates with ADG, favourably with feed conversion rate and negatively

with fat deposition. The strong effects of this allele on the mentioned traits were already shown in Italian Large White pigs (Fontanesi *et al.* 2010b) and they could justify its rapid increase in about 20 years, providing a proof of concept that markers with high effects can change their allele frequencies in a relatively short time following the direction of the selection pressure. Given the increase of about 30% of allele A over the considered period in the most influencing boars, it could be expected that this allele would reach fixation in the Italian Large White boar population within a few years, should the selection continue its pressure towards higher LC content. The trend toward the elimination of allele G in the population should be carefully monitored to avoid a reduction in reproduction performances of the sows as this allele is positively associated with prolificacy-related traits (Stinckens *et al.* 2010).

The frequency of the *MC4R* p.298N allele showed a similar increase (Fig. 2) over the same time period ($P<0.001$). This allele is associated with higher ADG, higher LC, higher HW, lower BFT and favourable FGR in Italian Large White pigs (Fontanesi *et al.* 2013). Its frequency was 0.575 at the beginning of the current selection scheme (years: 1992-1995) and reached 0.826 in the last group of genotyped boars (years: 2008-2010). In this period, its frequency showed a constant increase that reached about 25 percent points. The trend of this allele is specular of the trend on the EBVs in the boar population (Fig. 1).

While both the *IGF2* intron3-g.3072A allele and the *MC4R* p.298N allele have indirectly been selected starting from quite high similar frequencies in the first time window (boars born in 1992-1995), *FTO* and *VRTN* alleles have increased their frequencies in the same period even if they started from low frequencies (~20% for both alleles in boars born in 1992-95). This means that the effectiveness of the current selection programs in changing relevant allele frequencies appears to hold also for low/medium frequency alleles. In particular, *FTO* g.276G allele frequency moved from 0.180 to 0.479 ($P<0.01$), with an increase of about 30 points, even higher than the observed change in the *IGF2* and *MC4R* alleles. Allele G at the *FTO* gene is favorably associated with several traits which have been selected for the last 20 years: ADG, LC, and FGR (Fontanesi *et al.*

2009, 2010). This allele is also negatively associated with BFT (Fontanesi *et al.* 2010, 2013), a trait that decreased during the first years of the selection program and later on it has been substantially maintained constant in order to select pigs with the appropriate BFT needed for dry-cured ham production (Fig. 1). Looking at Fig. 2, one can see that over the considered two decades of selection the frequency of the G allele of the *FTO* gene increased in the first 15 years while it stabilized in the last period. It is tempting to note that that this tendency seems consistent with the general trend of the BFT EBV in the boar population.

During the investigated period, the *VRTN Q* allele increased from a frequency of about 0.23 to 0.42 ($P<0.001$). Allele *Q* is associated with higher vertebrae number as compared to the wild type allele (Mikawa *et al.* 2011). The latter still is the most frequent allele in Italian Large White breed. Vertebrae number is not recorded in the Italian Large White performance tested slaughtered animals and no direct information on this trait is used in the Italian Large White breeding scheme. As vertebral number is correlated with body length that, in turn, is correlated with other production and reproductive traits (*e.g.*, Borchers *et al.* 2004), it appears that the observed increase of about 20% of allele *Q* in the investigated boars might be derived by indirect selection due to the existing correlations between vertebrae number and other traits under selection in this population. In a previous study on Italian Large White performance tested pigs, we identified a mild association between allele *Q* and lower HW (Fontanesi *et al.* 2104). Since HW increased during the last 20 years of selection in the Italian Large White breed (Fig. 1), probably because of its correlation with lean meat content that is part of the general selection index of the breed, one should expect a decrease, not an increase in the frequency of *VRTN Q* allele (i.e. *IGF2* intron3-g.3072A and *MC4r* p.298N alleles), that might play a more important role on this trait. Therefore, in addition to the potential pleiotropic effects on other traits under selection, it could be also considered that allele *Q* might be increased due to hitchhiking effects determined by its closeness to other important QTL for production traits under selection. The *VRTN* gene is located on SSC7, in the middle of a QTL region for fat deposition traits (Mikawa *et al.* 2011).

The effects of directional selection on allele frequency changes have also been described in a few other livestock species. An interesting example was given by the Norwegian White Sheep breed for the myostatin 3'-UTR mutation causing muscle hypertrophia (Boman *et al.* 2011). The introduction of BLUP based breeding values and a classification system for carcass quality determined a rapid increase of the mutated allele in a short period of time without using any molecular genotyping (Boman *et al.* 2011). Similar changes were observed in the Italian Large White boars for a few gene markers that affect production traits, starting from the beginning of the selection scheme in this breed. Before the starting of the current selection strategy, the Italian pig population was quite heterogeneous and mass selection was the only approach used. That means that the introduction of BLUP based EBV started to shape the genetic pool of the Large White population moving allele frequencies for markers affecting or associated with production traits under directional artificial selection.

The applied statistical approach was able to capture the increase over the investigated period by using frequency calculated in different windows (eight groups of 2-4 years) that were defined to include a similar number of boars in each group. As the Italian Large White selection nucleus was quite limited in terms of number of animals during the time, and pedigree data indicate that all animals are, to some extent, related, the only criterion that was used to chose the boars to be genotyped was the reliability of the EBVs that was considered as a proxy that identified the most influencing boars of the population. The multinomial logistic regression models that were used to test allele frequency changes reduced the possibility that genetic drift would be the main cause of observed significant differences in allele frequencies between the beginning and the end of the considered period of time, as intermediate time points were used to fit the models. The results obtained in particular for the *IGF2* intron3-g.3072G>A and *MC4R* p.D298N polymorphisms confirm indirectly that the design of this study might obtain a good approximation of allele frequency changes over time in a pig population under directional artificial selection based on BLUP breeding values.

Conclusions

This study indicates that it is possible to design a simple experiment to evaluate allele frequency changes over time providing that DNA from animals borne years ago is available. The retrospective analysis that was carried out in Italian Large White boars suggested that selection based on methodologies assuming the infinitesimal model is able to modify in a quite short period of time allele frequencies in major genes, increasing the frequency of alleles explaining a relevant (non-infinitesimal) fraction of the overall genetic variability for production traits.

Acknowledgements

This work was funded by MiPAAF Innovagen project.

References

Barton N.H., Keightley P.D. (2002) Understanding quantitative genetic variation. *Nat. Rev. Genet*., **3**, 11-21.

Boman I.A., Klemetsdal G., Nafstad O., Blichfeldt T., Våge D.I. (2011) Selection based on progeny testing induces rapid changes in myostatin allele frequencies - a case study in sheep. *J. Anim. Breed. Genet.*, **128**, 52-55.

Borchers N., Reinsch N., Kalm E. (2004) The number of ribs and vertebrae in a Piétrain cross: variation, heritability and effects on performance traits. *J. Anim. Breed. Genet.*, **121**, 392-403.

Ciobanu D., Bastiaansen J., Malek M., Helm J., Woollard J., Plastow G., Rothschild M. (2001) Evidence for new alleles in the protein kinase adenosine monophosphate-activated gamma(3)-subunit gene associated with low glycogen content in pig skeletal muscle and improved meat quality. *Genetics*, **159**, 1151-1162.

Duijvesteijn N., Veltmaat J.M., Knol E.F., Harlizius B. (2014) High-resolution association mapping of number of teats in pigs reveals regions controlling vertebral development. *BMC Genomics*, **15**, 542.

Ernst C.W., Steibel J.P. (2013) Molecular advances in QTL discovery and application in pig breeding. *Trends Genet.*, **29**, 215-224.

Fan B., Du Z.Q., Rothschild MF. (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.

Fontanesi L., Davoli R., Nanni Costa L., Scotti E., Russo V. (2003) Study of candidate genes for glycolytic potential of porcine skeletal muscle: identification and analysis of mutations, linkage and physical mapping and association with meat quality traits in pigs. *Cytogenet. Genome Res.*, **102**, 145-151.

Fontanesi L., Davoli R., Nanni Costa L., Beretti F., Scotti E., Tazzoli, M., Tassone F., Colombo M., Buttazzoni L., Russo V. (2008) Investigation of candidate genes for glycolytic potential of porcine skeletal muscle: association with meat quality and production traits in Italian Large White pigs. *Meat Sci.*, **80**, 780-787.

Fontanesi L., Scotti E., Buttazzoni L., Davoli R., Russo V. (2009). The porcine fat mass and obesity associated (*FTO*) gene is associated with fat deposition in Italian Duroc pigs. *Anim. Genet.*, **40**, 90-93.

Fontanesi L., Scotti E., Buttazzoni L., Dall'Olio S., Bagnato A., Lo Fiego D.P., Davoli R., Russo V. (2010a) Confirmed association between a single nucleotide polymorphism in the *FTO* gene and obesity-related traits in heavy pigs. *Mol. Biol. Rep.*, **37**, 461-466.

Fontanesi L., Speroni C., Buttazzoni L., Scotti E., Dall'Olio S., Nanni Costa L., Davoli R., Russo V. (2010b) The *IGF2 intron3-g.3072G>A* polymorphism is not the only *Sus scrofa* chromosome 2p mutation affecting meat production and carcass traits in pigs: evidences from the effects of a cathepsin D (*CTSD*) gene polymorphism. *J. Anim. Sci.*, **88**, 2235-2245.

Fontanesi L., Galimberti G., Calò D.G., Fronza R., Martelli P.L., Scotti E., Colombo M., Schiavo G., Casadio R., Buttazzoni L., Russo V. (2012) Identification and association analysis of several hundred single nucleotide polymorphisms within candidate genes for back fat thickness in Italian Large White pigs using a selective genotyping approach. *J. Anim. Sci.*, **90**, 2450-2464.

Fontanesi L., Buttazzoni L., Galimberti G., Calò D.G., Scotti E., Russo V. (2013) Association between melanocortin 4 receptor (*MC4R*) gene haplotypes and carcass and production traits in Italian Large White pigs evaluated with a selective genotyping approach. *Liv. Sci.*, **157**, 48-56.

Fontanesi L., Russo V. (2013) Nucleotide variability and haplotype heterogeneity at the porcine *fat mass and obesity associated* (*FTO*) gene. *Anim. Genet.*, **44**, 96-100.

Fontanesi L., Scotti E., Buttazzoni L., Dall'Olio S., Russo V. (2014) Investigation of a short interspersed nuclear element polymorphic site in the porcine vertnin gene: Allele frequencies and association study with meat quality, carcass and production traits in Italian Large White pigs. *Ital. J. Anim. Sci.*, **13**, 61-65.

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., Smith G.D., 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*, **316**, 889-894.

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**, 448-451.

Jeon J.T., Carlborg O., Törnsten A., Giuffra E., Amarger V., Chardon P., Andersson-Eklund L., Andersson K., Hansson I., Lundström K., Andersson L. (1999) A paternally expressed QTL affecting skeletal and cardiac muscle mass in pigs maps to the *IGF2* locus. *Nat. Genet.*, **21**, 157-158.

Kim K.S., Larsen N., Short T., Plastow G., Rothschild M.F. (2000) A missense variant of the porcine melanocortin-4 receptor (*MC4R*) gene is associated with fatness, growth, and feed intake traits. *Mamm. Genome*, **11**, 131-135.

Mikawa S., Sato S., Nii M., Morozumi T., Yoshioka G., Imaeda N., Yamaguchi T., Hayashi T., Awata T. (2011) Identification of a second gene associated with variation in vertebral number in domestic pigs. *BMC Genet.*, **12**, 5.

Milan D., Jeon J.T., Looft C., Amarger V., Robic A., Thelander M., Rogel-Gaillard C., Paul S., Iannuccelli N., Rask L., Ronne H., Lundström 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**, 1248-1251.

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**, 155-156.

R Core Team (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Rothschild M.F., Hu Z.L., Jiang Z. (2007) Advances in QTL mapping in pigs. *Int. J. Biol. Sci.*, **3**, 192-197.

Stinckens A., Mathur P., Janssens S., Bruggeman V., Onagbesan O.M., Schroyen M., Spincemaille G., Decuypere E., Georges M., Buys N. (2010) Indirect effect of IGF2 intron3 g.3072G>A mutation on prolificacy in sows. *Anim. Genet.*, **41**, 493-498.

Tutz G. (2012) *Regression for Categorical Data*, Cambridge University Press, Cambridge.

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**, 832-836.

Venables W.N., Ripley B.D. (2002) *Modern Applied Statistics with S. Fourth Edition*. Springer, New York.

Table 1. Averaged reliability of estimated breeding values (EBVs) ± SD and averaged EBVs for several production traits (ADG = average daily gain; BFT = back fat thickness; FGR = feed gain ratio; LC = lean meat cuts; HW = ham weight) of the investigated Italian Large White boars divided in 8 different groups according to the year of birth.

Years	N. of boars	EBV reliability	ADG EBV (g)	BFT EBV (mm)	FGR EBV	HW EBV (kg)	LC EBV (kg)
1992-1995	22	0.904 ± 0.012	9.500 ± 30.226	-2.448 ± 2.888	-0.069 ± 0.177	0.438 ± 0.599	0.817 ± 1.399
1996-1997	24	0.914 ± 0.017	36.375 ± 33.618	-3.726 ±3.577	-0.210 ± 0.179	0.567 ± 0.639	2.099 ±1.736
1998-1999	25	0.928 ± 0.028	40.720 ± 30.867	-3.748 ± 3.400	-0.168 ± 0.164	0.679 ± 0.581	2.858 ± 1.684
2000-2001	26	0.941 ± 0.025	35.615 ± 28.625	-2.364 ± 3.457	-0.127 ± 0.150	0.558 ± 0.690	2.370 ± 1.821
2002-2003	27	0.937 ± 0.028	47.333 ± 26.482	-2.381 ± 3.482	-0.188 ± 0.126	0.648 ± 0.602	3.507 ± 1.492
2004-2005	30	0.932 ± 0.024	43.100 ± 18.415	-1.095 ± 3.162	-0.152 ± 0.088	0.539 ± 0.455	3.242 ± 1.276
2006-2007	21	0.939 ± 0.029	49.000 ± 15.473	-0.362 ± 2.157	-0.177 ± 0.092	0.732 ± 0.443	3.732 ± 1.318
2008-2010 [1]	25	0.940 ± 0.026	52.680 ± 27.536	-1.721 ± 2.262	-0.188 ± 0.133	0.668 ± 0.614	4.324 ± 1.430

[1] No boars with EBV reliability >0.85 were available in the years 2011 and 2012.

Table 2. PCR and genotyping protocols of the investigated gene markers.

Gene markers	**Primers 5'-3' (forward and *reverse*)**	**PCR conditions[1]**	**Genotyping protocols**	**References[2]**
IGF2 intron3-g.3072G>A	GACCGAGCCAGGGACGAG *CGCGCCCCACGCGCTCCCACGCTG*	62/2.5	PCR-RFLP (*Ade*I): 85 bp = allele G; 65 + 20 bp = allele A	Van Laere *et al.* (2003), Fontanesi *et al.* (2010b)
MC4R p.D298N (c.892G>A)	TACCCTGACCATCTTGATTG *ATAGCAACAGATGATCTCTTTG*	54/2.5	PCR-RFLP (*Taq*I): 226 bp = allele N (A); 156 + 70 bp allele D (G)	Kim *et al.* (2000)
VRTN PRE1 insertion	GGCAGGGAAGGTGTTTGTTA *GACTGGCCTCTGTCCCTTG*	56/1.5	Fragment analysis: 411 bp = allele *Q*; 120 bp = allele wild type (WT)	Mikawa *et al.* (2011)
PRKAG3 p.I199V	GGAGCAAATGTGCAGACAAG *CCCACGAAGCTCTGCTTCTT*	55/3.0	PCR-RFLP (*Bsa*HI): 257 bp = allele I; 211 + 46 bp = allele V	Ciobanu *et al.* (2002)
FTO g.276T>G	ACAGGCCCTGAAGAGGAAAG *AGTAACCTGGAGTTCCTGTGG*	60/2.0	PRC-RFLP (*Tai*I): 397 bp = allele T; 275 + 122 bp = allele G	Fontanesi *et al.* (2009)

[1]Annealing temperature (°C)/[$MgCl_2$] or [$MgSO_4$] (mM) for the *IGF2* intron3-g.3072G>A. For this marker PCR included also 0.3X of PCRx Enhancer Solution (PCRx Enhancer System, Invitrogen, Carlsbad, CA, USA).

[2] References that described the polymorphisms and the genotyping protocols.

Table 3. Effect of time on allele frequency changes over the considered period.

Gene markers	**Logit deviance differences**[1]	***P***[2]
IGF2 intron3-g.3072G>A	17.584	0.000152
MC4R p.D298N (c.892G>A)	17.847	0.000133
FTO g.276T>G	12.770	0.0017
VRTN PRE1 insertion	23.318	0.000009
PRKAG3 p.I199V	3.146	0.207

[1] Critical chi-square = 5.991 (2 df, $P = 0.05$)

[2] Chi-square *P* value.

Figure 1. Trends of estimated breeding values (EBVs) in the Italian Large White boar population over the 1992-2012 period for several production traits (ADG = average daily gain, g; BFT = back fat thickness, mm; FGR = feed gain ratio; HW = ham weight, kg; LC = lean meat cuts, kg).

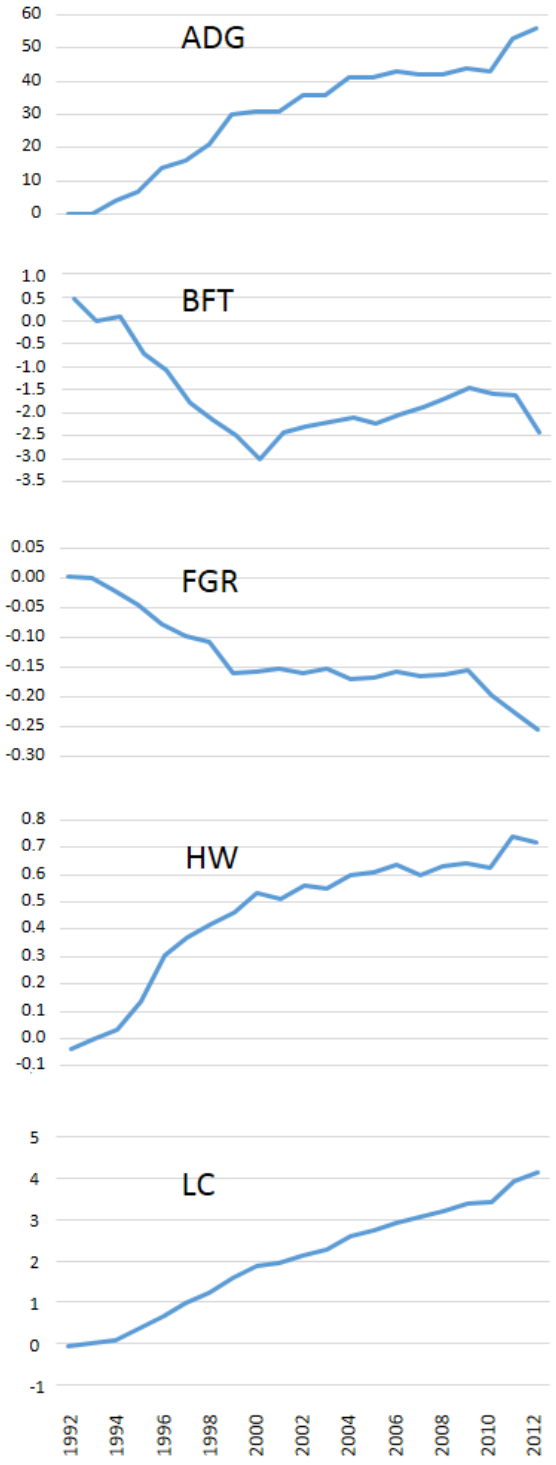
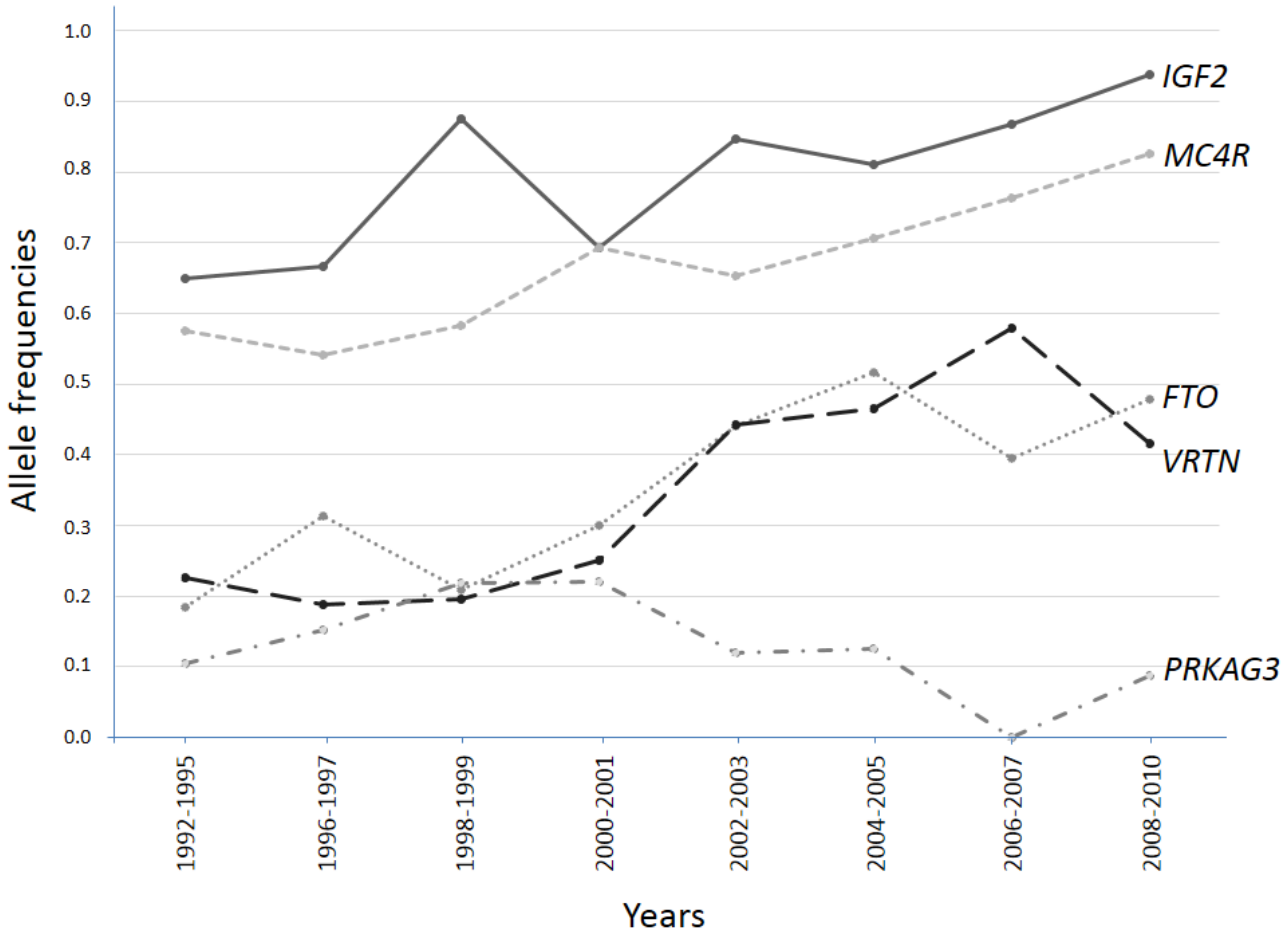


Figure 2. Frequency changes of the *IGF2* intron3-g.3072A, *MC4R* p.298N, *VRTN Q*, *FTO* g.276G and *PRKAG3* p.199I alleles in the Italian Large White boars with estimated breeding value reliability >0.85.



CHAPTER 5

Twenty years of artificial directional selection have shaped the genome of the Italian Large White pig breed

G. Schiavo*, G. Galimberti†, D.G. Calò†, A.B. Samorè*, F. Bertolini*, V. Russo*, M. Gallo‡, L. Buttazzoni§ and L. Fontanesi*

* Department of Agricultural and Food Sciences, Division of Animal Sciences, University of Bologna, Viale Fanin 46, 40127 Bologna, Italy

† Department of Statistical Sciences "Paolo Fortunati", Via delle Belle Arti, University of Bologna, 40126 Bologna, Italy

‡ Associazione Nazionale Allevatori Suini, Via L. Spallanzani 4, 00161 Roma, Italy

§ Centro di Ricerca per la Produzione delle Carni e il Miglioramento Genetico, Consiglio per la Ricerca e la Sperimentazione in Agricoltura, Monterotondo (Roma), Italy

Correspondence:
L. Fontanesi, Department of Agricultural and Food Sciences, Division of Animal Sciences, University of Bologna, Viale Fanin 46, 40127 Bologna, Italy
Tel: +39 051 2096535
Fax: +39 051 2096355
E-mail: luca.fontanesi@unibo.it

Running title: Genome wide effects of selection in pigs

Submitted to Animal Genetics, March 2015

Summary

In this study we investigated at the genome wide level if 20 years of artificial directional selection based on boar genetic evaluation obtained with a classical BLUP animal model shaped the genome of the Italian Large White pig breed. The most influencing boars of this breed (n. 192), born from 1992 (the beginning of the selection program of this breed) to 2012, with estimated breeding value reliability >0.85, were genotyped with the Illumina Porcine SNP60 BeadChip. After grouping the boars in eight classes according to their year of birth, filtered single nucleotide polymorphisms (SNPs) were used to evaluate the effects of time on genotype frequency changes using multinomial logistic regression models. Of these markers, 493 had a $P_{Bonferroni} < 0.10$. However, there was an increasing number of SNPs with a decreasing level of allele frequency changes over time, representing a continuos profile across the genome. The largest proportion of the 493 SNPs was on porcine chromosome (SSC) 7, SSC2, SSC8 and SSC18 for a total of 204 haploblocks. Functional annotations of genomic regions including the 493 shifted SNPs reported a few Gene Ontology terms that might underly the biological processes that contributed to increase performances of the pigs over the 20 years of the selection program. The obtained results indicated that the genome of the Italian Large White pigs was shaped by a directional selection program derived by the application of methodologies assuming the infinitesimal model that captured a continuos trend of allele frequency changes in the boar population.

Keywords: selection signature; genome wide analysis; directional selection; SNP; pig breed.

Introduction

The domestication process, determined by complex spatial and temporal long-term and continuos genetic changes derived by population admixture and isolation events, started to shape the genome of farm animals (Larson & Burger 2013). This process continued in a more organized way with the subsequent constitution of the livestock breeds when animals of similar morphological traits were preferentially mated in close populations. These human-driven breeding events led to the fixation of few phenotypes (for example coat colour, stature, horns, etc.) that differentiated domesticated animals from wild ancestors and, within domesticated species, produced different breeds. These processes left selection signatures in the animal genomes at loci affecting these traits, usually determined by one or few major genes (Rubin *et al.* 2012; Kemper *et al.* 2014).

However, only the introduction of modern quantitative genetics methodologies in the 20th century allowed for a substantial improvement of economically relevant traits, considered as quantitative traits. According to the infinitesimal model underlying quantitative genetics, the genetic structure of quantitative traits is determined by an infinite number of genes, each of them with a very small effect. Therefore, directional selection derived by modern breeding goals is not expected to significantly change allele frequencies at any loci in the population (Barton & Keightley 2002; Walsh 2007). However, it is now well known that a detectable part of the genetic variance of complex traits subject to genetic improvement plans is explained by specific loci, as demonstrated by many QTL studies thus far produced in pigs and in several other livestock species (see Hu *et al.* 2013). Therefore, changes in allele frequencies at some of these loci may occur. The direction of these changes should be consistent with the directional selection expected by the implemented breeding programs, producing also for quantitative traits some detectable signatures of selection in a quite short time, depending on the magnitude of their effects.

Boman *et al.* (2011) showed that selection based on progeny testing in the Norwegian White Sheep population favouring muscled carcasses indirectly induced a rapid frequency change of a causative mutation in the 3'-untranslated region of the myostatin gene. In pigs, we recently showed

that directional selection, operated by a BLUP Animal Model under a sib-testing scheme, significantly changed, in quite a short period of time, allele frequencies of a few major genes in the Italian Large White breed (Fontanesi *et al.* 2015). These genes and their mutations (*IGF2* intron3-g.3072G>A, *MC4R* p.D298N, *VTNR* PRE1 insertion and *FTO* g.276T>G) are already well known to affect production traits in different pig popuations (Van Laere *et al.* 2003; Kim *et al.* 2000; Mikawa *et al.* 2011; Fontanesi *et al.* 2009), including the Italian Large White breed (Fontanesi *et al.* 2009; 2010; 2012a; 2013; 2014a). In our previous study that detected a significant allele shift in the Italian Large White breed (Fontanesi *et al.* 2015), these polymorphisms were genotyped in the most important boars of this breed, born over a 20 year time span (from the beginning of the selection program in 1992 to 2012). The significant changes in allele frequencies detected over time were due to a progressive and continuous increase of one allele (associated with higher lean meat content, higher average daily gain, and favorable feed:gain ratio), following the direction of the selection program of this pig breed. These results suggest that selection based on methodologies assuming the infinitesimal model is also able to increase the frequency of alleles explaining a relevant (non-infinitesimal) fraction of the overall genetic variability for production traits and they provided a proof of concept that prompted the genome wide investigation we carried out in this study. Taking advantage of a commercial platform for high throughput genotyping of single nucleotide polymorphisms (SNPs) in pig (Illumina *Porcine SNP60 BeadChip*; Ramos *et al.* 2009), in this work we genotyped the same Italian Large White boars investigated in a previous study (Fontanesi *et al.* 2015). Results indicated that allele frequencies at many SNP positions, covering several chromosome regions containing QTLs for production traits (as defined in several other studies), significantly changed over the considered period, similarly to what was obtained for a few investigated candidate genes (Fontanesi *et al.* 2015). In this way, we obtained a genome wide picture of allele frequency shifts generated by the artificial selection operated in the Italian Large White breed over the last two decades.

Materials and Methods

Animals and traits

A total of 5,983 Italian Large White boars, born from 1992 (two years after the onset of the current selection scheme for this breed) to 2012 and approved for reproduction based on their genetic merit after evaluation by the National Pig Breeder Association (ANAS), were ranked according to the reliability of their EBVs calculated in 2012. All boars with reliability >0.85 (n. = 200) were selected for this study (Fontanesi *et al.* 2015). Of the 200 selected animals, 192 were genotyped as genotyping plates were multiple of 96. Selected boars were divided according to their year of birth in eight groups, each spanning 2-4 years (Fontanesi *et al.* 2015). Average EBVs and reliability for each boar group are reported in Table 1. No boars with EBV reliability >0.85 were available from those that were born in the years 2011 and 2012. Candidate boars are evaluated by ANAS using a sib-testing scheme already described in previous works (Fontanesi *et al.* 2010, 2012a). Briefly, two females and one castrated male littermate of a candidate boar (triplets of siblings) are individually performance tested at the Central Test Station of ANAS.

Performance test of the animals starts at 30 to 45 days of age and ends when pigs reach 155 ± 5 kg live weight. On each animal, body weight is measured every two weeks and feed intake is recorded daily, then average daily gain (ADG) and feed:gain ratio (FGR) are calculated. At the end of test, animals (of the performance tested triplets) are slaughtered in a commercial abattoir. Then back fat thickness (BFT) is measured on each carcass at the level of *Musculus gluteus medius* and weight of lean cuts (LC; necks and loins) and of hams (HW) are measured after carcass dissection.

Estimated breeding values are calculated for the indicated traits (ADG, expressed in g; LC and HW, expressed in kg; BFT, expressed in mm; and FGR) by a BLUP-Multiple Trait-Animal Model with different models per each trait including (depending on the traits): the fixed effects of sex, batch on trial, inbreeding coefficient of the animal, interaction of sex by age at slaughtering, date of slaughtering and the random effects of litter and animal. The genetic progress for ADG, BFT, FGR, HW and LC in the Italian Large White population over the 1992-2012 period is

summarized in the trends reported in Figure S1. Selection objectives during the considered period were: ADG and LC, increase; FGR, decrease; BFT, decrease and then maintenance; HW, increase and then maintenance (Fontanesi *et al.* 2015).

Genotyping

Blood samples were collected from candidate boars and then lyophilized and stored at room temperature. DNA extraction from lyophilized blood was carried out with the NucleoSpin® Blood Kit (Macherey-Nagel GmbH & Co. KG, Düren, Germany) according to the manufacturer's instructions. Genomic DNA was genotyped using the Illumina PorcineSNP60 v2 BeadChip array (Illumina, San Diego, CA, USA) according to the manufacturer's protocol.

Plink software (Purcell *et al.* 2007) was used to filter data: SNPs with total minor allele frequency (MAF) < 0.05, or call rate < 0.9, or Hardy-Weinberg equilibrium test < 0.001 were removed. SNPs were mapped on the Sscrofa 10.2 genome assembly, setting a similarity threshold at 94% and discarding non uniquely mapped or unmapped SNPs as described in previous works (Fontanesi *et al.* 2012b, 2014). A total of 38,662 autosomal and 4 pseudoautosomal SNPs passed the filtering steps and were subsequently used for statistical analyses.

Statistical and bioinformatic analyses

In order to assess the effect of time on allele frequencies, the multinomial logistic regression model was used. This model allows to study how a nominal outcome variable Y with more than two categories depends on $P \geq 1$ covariates (Hosmer *et al.* 2013). For each SNP marker, genotype represents a nominal variable with 3 categories, coded as 0, 1 and 2 according to allele occurrence. Thus, a model consisting of the following two equations has been fitted by maximum likelihood:

$$\begin{cases} \log\left[\dfrac{P(Y=1\,|\,t)}{P(Y=0\,|\,t)}\right] = \beta_{10} + \beta_{11}t \\ \log\left[\dfrac{P(Y=2\,|\,t)}{P(Y=0\,|\,t)}\right] = \beta_{20} + \beta_{21}t \end{cases}$$

where Y=0 is taken as the reference baseline genotype and covariate t denotes time (coded as integers ranging from 1 to 8 according to group membership described in Table 1). In order to assess the overall effect of time on all the genotype frequencies, the following system of hypotheses has been considered:

$$H_0 : \begin{cases} \beta_{11} = 0 \\ \beta_{21} = 0 \end{cases}$$

It leads to a nested model which assumes time invariance for all genotype frequencies. A significance test has been performed according to likelihood ratio theory. In particular, the Wilks test-statistics were computed by taking minus two times the difference in the log-likelihood of two models (one assuming time-invariance, according to H_0, and the other one without any restriction on the model parameters). The corresponding *P*-value was computed using a chi-square distribution with 2 degrees of freedom, which represents an asymptotic approximation of the distribution of the test-statistics under H_0. The multinomial logistic regression model allowed also to perform specific analyses in order to evaluate changes in time in the heterozygous genotype frequency and in the minor homozygous genotype frequency separately. In this way, it was possible to attribute the effect on allele frequency changes. This was carried out by testing the following hypotheses:

$$H_0 : \beta_{11} = 0$$

for checking time invariance in the heterozygous genotype frequency and

$$H_0 : \beta_{21} = 0$$

for checking time invariance in the minor homozygous genotype frequency.

Following likelihood ratio theory, each of these two hypotheses can be tested using the Wald test-statistics, obtained by dividing the maximum likelihood estimate of the corresponding coefficient by its estimated standard error. Under H_0, the asymptotic distribution of this test-statistics can be approximated by a standardized Gaussian distribution that was used to compute (two-tail) *P*-values. Model fitting and *P*-value computation were performed using the *nnet* package (Venables and Ripley, 2002) in R (R Core Team, 2013). Bonferroni correction was used to identify the significant thresholds from the models that evaluated genotype frequency changes ($P_{Bonferroni}$ = 0.01 and $P_{Bonferroni}$ = 0.05) and the suggestive threshold ($P_{Bonferroni}$ = 0.10) that corresponded to $P_{nominal\ values}$ of 2.59E-07, 1.29E-06, and 2.59E-06, respectively. For SNPs with at least $P_{Bonferroni} < 0.10$, allele frequency changes was evaluated considering $P_{nominal\ values} <0.05$.

A sliding window approach was used to evaluate the enrichment of SNPs with $P_{Bonferroni} < 0.10$ in the pig autosomal genome. Window size was set up in order to apply Fisher's exact test with Yates correction according to the expected number of at least 1 SNP with $P_{Bonferroni} < 0.10$ under the null hypothesis. Dividing the number of autosomal filtered SNPs (38,662) by the number of SNPs with $P_{Bonferroni} < 0.10$ (493) we identified the number of contiguos SNPs (78) in the windows that overlapped by 39 SNPs when slided. A total of 477 non overlapping windows and 954 overlapped windows were obtained (these numbers were derived considering that some windows at the ends of the chromosomes could not be defined appropriately and were discarded). Then, Fisher's exact tests with Yates correction were computed in the 954 slided windows based on the observed SNPs with $P_{Bonferroni} < 0.10$ among the 78 SNPs of the considered window.

Haploview software (Barret *et al.* 2005) was used to compute linkage disequilibrium values in a pairwise comparison between SNPs. Haploblocks were obtained in the following way: for all SNPs with $P_{Bonferroni} < 0.10$ (tag SNPs) all other SNPs with a value of $r^2 \geq 0.4$ and distance < 500 kb were selected.

Annotation of genomic regions close to the SNPs with $P_{Bonferroni} < 0.10$ was based following different criteria: i) considering all genes included in a window of ± 500 kb around these SNPs, ii)

considering the genes in the enriched genomic regions determined with the sliding window approach (see above) and iii) considering genes in the haploblock regions determined as described above. Genes were retrieved from the Sscrofa10.2 genome version using the Ensembl Biomart tool v. 7.0 (http://www.ensembl.org/biomart/martview/). Analysis of Gene Ontology (GO) term enrichment was performed using the official gene names with the DAVID Bioinformatics Resources v. 6.7 (http://david.abcc.ncifcrf.gov/; Huang *et al.* 2009).

QTL information was retrieved from the PigQTLdb (http://www.animalgenome.org/cgi-bin/QTLdb/SS/index; Hu *et al.* 2013), release 25 (Dec 29, 2014). QTL positions on Sscrofa10.2 reported in this database were overlapped i) with regions ± 500 kb around SNPs with $P_{Bonferroni} < 0.10$, ii) with enriched genome windows for SNPs having $P_{Bonferroni} < 0.10$, as determined above and iii) with haploblocks identified as decribed above. Moreover, SNPs with P <0.001 in two previous genome wide association (GWA) studies for BFT and ADG carried out in the Italian Large White breed (Fontanesi *et al.* 2012b; 2014b) were used to identify overlapping regions (using median haploblock size, defined above) containing SNPs with the same significant values as in this study.

Results

The Italian Large White boars with most reliable EBVs born over a period of 20 years were genotyped with the Illumina PorcineSNP60 v2 BeadChip array (Illumina). Results of the logistic regression showed that continuous genotype frequency changes during this time produced a $P_{Bonferroni} < 0.10$ for 493 SNPs. Of these SNPs, 399 had a $P_{Bonferroni} < 0.05$, and for 222 of them $P_{Bonferroni}$ was < 0.01. Distribution of these SNPs on the different chromosomes is reported in Table S1. The largest number of SNPs with $P_{Bonferroni} < 0.10$ were on porcine chromosome (SSC) 7 (expected proportion = 0.065; observed proportion = 0.156), SSC2 (expected proportion = 0.068; observed proportion = 0.128), SSC8 (expected proportion = 0.054; observed proportion = 0.128), and SSC18 (expected proportion = 0.025; observed proportion = 0.073). These results are shown on Figure 1 in which genotype frequency changes of SNPs from the three threshold intervals are

positioned in a heat map representation of the pig autosomes. A larger density of SNPs with $P_{Bonferroni} < 0.10$ is evident in the first part of SSC7, in the middle of SSC8 and in a few regions of SSC2, SSC5, SSC12, SSC13, SSC14 and SSC18, as also more precisely evaluated with the enrichment analysis obtained with sliding windows (Table S2). A total of 33 windows ($P_{nominal\ value} <0.05$; for 6 of them $P_{Bonferroni}<0.10$) were enriched of suggestively significant SNPs. In particular, from this analysis, about 25 Mbp of SSC8 (from positions 51.1 Mbp to 76.5 Mbp, merging overlapped windows) included the largest number of these SNPs.

Figure 2 reports the Manhattan plot designed with the significant values of the logit model for all autosomal SNPs. A few strait peaks are evident below the most significant SNPs that were located on the chromosomes already mentioned (e.g. SSC8, SSC18, SSC5, and SSC7). More detailed statistics and information for SNPs with $P_{nominal\ values}$ <1.00E-10 are reported in Table 2 (data on the remaining SNPs with $P_{Bonferroni} < 0.10$ are included in Table S3). The plotted allele frequencies of the most significant SNPs in the eight time points defined in the 20 years window are reported in Figure S2. Increased levels of one allele (or decrease of the other allele) for the 493 SNPs ($P_{Bonferroni} < 0.10$) were on average of 34.9 ± 7.8 frequency points (and ranged from 15.8 to 55.8 frequency points; see Table 2 and Table S3). For example, the first two most significant SNPs (ASGA0085562 and ASGA0088966, $P_{nominal\ values}$ = 5.00E-15), mapping very close on SSC8 and in complete linkage disequilibrium, reduced progressively and constantly the frequency of the minor allele (0.425 at the starting time) that almost disappeared in the last time point (0.021). That means that the alternative allele reached almost fixation in these boars. The starting frequency (0.225) of the minor allele of the third top significant marker (ASGA0085170 on SSC8, not in linkage disequilibrium with the previous ones) increased of more than 50 points resulting the major allele (0.750) in the last time period.

Linkage disequilibrium analysis around the 493 SNPs with $P_{Bonferroni}$ <0.10 produced 204 haploblocks including a total of 391 of these SNPs and encompassing from a minimum of 2 (haploblocks on SSC2, SSC7, SSC8, SSC13, SSC16, SSC17 AND SSC18) to a maximum of 14

SNPs (a hapoblock on SSC14), considering also SNPs around these 391 markers that were used as seeds, and spanning from about 1 kbp (on SSC16) to about 1 Mbp (on SSC2 and SSC18) with a median of 166.58 kbp included (Table S4).

Functional annotation of genomic regions including SNPs with $P_{Bonferroni} < 0.10$ was obtained using different approaches. First, genes were retrieved within regions ± 500 kbp from these SNPs. Using this approach, at least one gene was identified in this range for 462 out of 493 SNPs and summing up a total of 2198 annotated genes. A total of 988 annotated genes were retrieved in the genome windows enriched of these SNPs (considering a $P_{nominal\ value} <0.05$ to detect windows significantly enriched of these SNPs). Haploblocks described above contained a total of 391 annotated genes. Results of the GO enrichment analysis are reported in Table S5. The most significant GO terms using the first approach (± 500 kbp) were chromatin assembly or disassembly (GO:0006333), protein-DNA complex assembly (GO:0065004), nucleosome organization (GO:0034728), nucleosome assembly (GO:0006334), DNA packaging (GO:0006323) and chromatin assembly (GO:0031497) (not significant after Bonferroni correction) that were also identified using the second approach (windows) in addition to several other with related or different functions (advanced glycation end-product receptor activity, GO:0050785; positive regulation of smooth muscle cell proliferation GO:0048661; MHC class I protein binding, GO:0042288), most of which significant after Bonferroni correction. The lower number of genes retrieved with the haploblock method did not obtain any highly enriched GO term. In addition to the general overview that could be captured by the GO analyses, functional information could be retrieved looking at the closest genes to the most significant SNPs. For example, the closest gene to the two most significant SNPs (ASGA0085562 and ASGA0088966 at positions 3625594 and 3630272 of SSC8, respectively) was the prosaposin-like 1 gene (*PSAPL1*). Prosaponin is a protein that regulates lysosomal enzyme function within the cell whereas it plays a neuroprotective and glioprotective role outsise the cell (Meyer *et al.* 2014). However the function of the *PSAPL1* gene is not known in detail yet. The third most significant SNP (ASGA0085170 at position 3378292 on SSC8) was close

to the transcriptional adaptor 2B (*TADA2B*) gene. TADA2B, also known as ADA2B, is a transcriptional adaptor protein that potentiates transcription by coordinating histone acetyltransferase activity and by linking activation factors to basal transcriptional machinery (Barley *et al.* 2003).

The genome regions around or encompassing SNPs with $P_{Bonferroni} < 0.10$ overlapped a large number of QTL reported in the PigQTLdb (Hu *et al.* 2013). A total of 3800 QTL for many economically relevant traits were already reported in genome regions ± 500 kb from these SNPs. Information for 2966 and for 965 QTL entries were projected in genome regions overlapped with enriched windows ($P<0.05$) of significant or suggestively significant SNPs and with the haploblocks described above, respectively (data not shown). Even if this comparison should be considered with caution, it seems that many chromosome regions in which allele frequencies of the boar population shifted over the last 20 years might contain QTLs whose effect could be captured by the directional selection currently ongoing in the Italian Large White population. Similar evidences could be obtained by overlapping information coming from our previous GWA studies on BFT and ADG in the Italian Large White breed (Fontanesi *et al.* 2012b; 2014b). Single nucleotide polymorphisms with $P_{\text{nominal value}} <0.001$ in these two GWA studies were included in genomic regions close to SNPs with the same $P_{\text{nominal value}}$ level in the current investigation: 98% and 92% of th SNPs from previous BFT and ADG GWA studies (Fontanesi *et al.* 2012b; 2014b) were captured by 100% and 98% of the SNPs with the same P value threshold of the current study (overlapping determined using median haplotype block size of ±170 kb, estimated from the haplotype blocks of Table S4). The use of a relaxed threshold for this rough evaluation between different studies (even if they calculated the P value using different assumptions) could be taken as a tentative way to explain what is reported in Figure 3. There is a continuos and increasing number of SNPs which changed allele frequency with decreasing differences during time, approximating what would be expected based on the results on the GWA studies: a few markers with effects that

trespassed the stringent multiple testing defined thresholds and many others with lower and decreasing effects according to what woud be expected for complex quantitative traits.

Discussion

The aim of our study was to give a first picture of the changes that the genome of the Italian Large White pig population has experienced since early 90's, when the genetic improvement program for this breed began.

The selection nucleus of the Italian Large White breed is a relatively small population in which all animals are, to some extent, related. Therefore, the only criteria that was used to select boars for genotyping among those evaluated in the period 1992-2012 was the reliability of their EBVs. This parameter largely depends on the progeny number, and therefore it identifies the most influencing Italian Large White boars (Fontanesi *et al.* 2015). Using this approach, allele frequencies in the selected boars may represent a good approximation of the frequencies in all boars of the breed, and, indirectly in the whole population. If some loci actually have a detectable effect on genetic variation of the considered production traits in Italian Large White pigs, their frequency changes over time should reflect, at least in part, the trends of boar EBVs during the same time period of 20 years (Figure S1), as we have already shown in a previous study on allele frequency changes of major genes affecting production traits in pigs (Fontanesi *et al.* 2015). To statistically test the continuous increase of the frequency of one SNP allele (and *vice versa* the decrease of the frequency of other allele) we used a logistic regression model that, fitting several time points, can minimize (but not completely exclude) the potential effect of genetic drift, that, by definition, can have a random and alternate effect on allele frequencies (Falconer & Mackay 1996).

The presented results in boars indicate that for quite a large number of marker alleles, frequencies changed during the past two decades and left peculiar selection signatures of the implemented genetic program in the Italian Large White breed since its beginning, including the effects of directional selection in a semi-close nucleous with a low effective population size. As the

selection program was for several traits through the use of a total genetic merit index, it would be impossible to attribute genome changes to a specific trait. Our results contradict the general principles of the infinitesimal model for which allele frequencies are not expected to substantially change during time, as allele effects on complex traits would be infinitesimally small (Walsh 2007), but they agree with the evolution of this theory that was proposed after the introduction of genomic information in animal breeding (Hill, 2014). In our study, the genome wide evaluation of allele frequency shifts was probably limited by the problem of multiple testing that needed stringent thresholds to declare significant or suggestively significant changes (222, 399 and 493 SNPs were below the $P_{Bonferroni}$ = 0.01, 0.05 and 0.10 thresholds, respectively). In this way we excluded several other relevant genome regions. For example, changes for the candidate gene markers (*IGF2* intron3-g.3072G>A, *MC4R* p.D298N, *FTO* g.276T>G and *VRTN* PRE1 insertion) that we detected in a previous work (Fontanesi *et al.* 2015) would not be considered significant in the context of a genome wide study as their P values obtained with the same statistics did not reach the $P_{Bonferroni}$ = 0.10 threshold (their $P_{nominal\ values}$ were 0.000152, 0.000133, 0.0017 and 0.000009, respectively). Anyway, the effects of these candidate genes on production traits have been already demonstrated by several other studies and their allele frequencies actually changed according to directional selection aiming at improving growth and carcass traits (Fontanesi *et al.* 2015). However, allowing for more relaxed thresholds, a larger number of SNPs changed allele frequencies over time. For example, 1665 and 3613 SNPs had a $P_{nominal\ values} < 0.0001$ or <0.001, respectively (see Figure 3 for the distribution of the analysed SNPs). That means that for many other SNPs, the shift over time was smoother and more difficult to capture as it could be counfounded with genetic drift or with the error of the models. It is tempting to see in these results a similar picture that, from a different perspective, it has been depicted on the distribution of gene effects on quantitative traits in livestock: many genes (markers) with small effects (close to the infinitesimal model) and few genes (markers) with large effects (Hayes & Goddard 2001). In our study we observed for several markers large allele frequency changes and for many others decreasing and smaller allele frequency

changes. This parallel could potentially be modelled by the magnitude and distribution of their decreasing effects.

Kemper *et al.* (2014) reported little discernible selection sweeps in dairy cattle genome despite the very strong and recent artificial selection for milk production traits. They supported their conclusions by within breed haplotype homozygosity measures, assuming homozygosity levels as the final result of selection leading to haplotype fixation in those regions. However, as artificial selection can be conceived as a continuos process, its effects should be probably better evaluated by studying the addititive substitution effects deriving from allele accumulation in time detectable from past time windows as we proposed in our study and before by Glick *et al.* (2012). These Authors reported a high correlation between selection intensity for milk production and reproduction traits, by comparing changes in haplotype frequencies in a population of about 1000 Israeli Holstein bulls (genotyped with the Illumina BovineSNP50 BeadChip array) over a period spanning about 5 bull generations, and those derived by trait-based analysis of the cow population. This study on cattle confirmed that it is possible to detect selection sweeps derived by recent selection programs on complex traits by measuring haplotype (or allele) frequency changes. Other methodological advances and integrations to the logistic regression model applied in our study are needed to correlate selection intensities for the traits of the Italian Large White genetic program and changes in allele or haplotype frequencies. This could be done by linking information on the regression slopes with the percentage of variance of EBV explained by SNPs (or haplotypes) that changed significantly over this period, but more genotyping data and integration of the applied methodology with other information on the pig population investigated would be needed.

In our study, we did not consider changes in haplotype frequencies for a few reasons: the number of genotyped animals for each time point could not be large enough to obtain reliable estimates of haplotype frequencies in case of more than two haplotypes and a more complex model would be needed to fit more than three categories (the three genotypes at each locus in case of biallelic markers). Haplotype information was obtained around significant and suggestively

significant SNPs as a subsequent step to define population structures of chromosome regions that have been shifted over a defined period and presumably by directional artificial selection (Table S4).

Some chromosome regions reported a large number of significant and suggestively significant SNPs (Figure 1). This might be due to hitch-hiking effects associated with the indirect selection of favourable QTL alleles of by multiple linked QTLs (Smith & Haigh 1974; Liu *et al.* 2014). These regions potentially include QTLs for production traits that are under selection in the Italian Large White breed and the information obtained in this study could be useful to identify those that might segregate in this breed. For some of the positive QTL alleles we can expect that the directional selection operated over the last 20 years could lead to fixation or already reached fixation, according to the observed allele frequency shifts of putative linked SNPs over this period (Table 2, Table S3 and Figure S2). This is what we already observed for the *IGF2* intron3-g.3072G>A polymorphism for which allele A, positively affecting growth and lean meat deposition, is expected to reach fixation in the breed if selection pressure on these traits will be maintained costant for the next years (Fontanesi *et al.* 2015). The functional annotation of genes localized in chromosome regions including SNPs that changed allele frequency over time reported GO categories that are related to basic biology processes involved in cell replication, proliferation and antigen production. This annotation could suggest that the directional selection acted on specific biological mechanisms that might indirectly, and in combination with specific function of each gene in these regions, explain the complex economic traits that were improved by the implemented selection program.

In conclusion, this study, based on genotyping data of the most influencing boars, reported for the first time a genome wide retrospective analysis of allele frequency shift over time in a pig breed. The obtained results indicated that the genome of the Italian Large White breed was shaped by the artificial directional selection program derived by the application of methodologies assuming the infinitesimal model that captured a continuos trend of allele frequency changes across the genome.

This study was indirectly able to represent the dynamic effects of artificial forces that could be regarded as the latest "evolutionary" events that acted in the investigated pig population.

Conflict of Interests

The authors declare that they do not have any conflict of interests.

Acknowledgements

We thank Emilio Scotti (University of Bologna) for technical assistance. This study was supported by Italian MiPAAF (INNOVAGEN project) funds.

References

Barlev N.A., Emelyanov A.V., Castagnino P., Zegerman P., Bannister A.J., Sepulveda M.A., Robert F., Tora L., Kouzarides T., Birshtein B.K. & Berger S.L. (2003) A novel human Ada2 homologue functions with Gcn5 or Brg1 to coactivate transcription. *Molecular and Cell Biology* **23**, 6944-57.

Barrett J.C., Fry B., Maller J. & Daly M.J. (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263-5.

Barton N.H. & Keightley P.D. (2002) Understanding quantitative genetic variation. *Nature Review Genet*ics **3**, 11-21.

Boman I.A., Klemetsdal G., Nafstad O., Blichfeldt T. & Våge D.I. (2011) Selection based on progeny testing induces rapid changes in myostatin allele frequencies - a case study in sheep. *Journal of Animal Breeding and Genetics* **128**, 52-5.

Falconer D.S. & Mackay T.F.C. (1996) *Introduction to Quantitative Genetics*, Fourth Edition. Longmans Green, Harlow, Essex, UK.

Fontanesi L., Scotti E., Buttazzoni L., Davoli R. & Russo V. (2009). The porcine fat mass and obesity associated (*FTO*) gene is associated with fat deposition in Italian Duroc pigs. *Animal Genetics* **40**, 90-3.

Fontanesi L., Speroni C., Buttazzoni L., Scotti E., Dall'Olio S., Nanni Costa L., Davoli R., Russo V. (2010) The *IGF2 intron3-g.3072G>A* polymorphism is not the only *Sus scrofa* chromosome 2p mutation affecting meat production and carcass traits in pigs: evidences from the effects of a cathepsin D (*CTSD*) gene polymorphism. *Journal of Animal Science* **88**, 2235-45.**88**, 2235-2245. **88**, 2235-2245.**888**

88, 2235-2245.

Fontanesi L., Galimberti G., Calò D.G., Fronza R., Martelli P.L., Scotti E., Colombo M., Schiavo G., Casadio R., Buttazzoni L. & Russo V. (2012a) Identification and association analysis of several hundred single nucleotide polymorphisms within candidate genes for back fat thickness in Italian Large White pigs using a selective genotyping approach. *Journal of Animal Science* **90**, 2450-64.

Fontanesi L., Schiavo G., Galimberti G., Calò D.G., Scotti E., Martelli P.L., Buttazzoni L., Casadio R. & Russo V. (2012b) A genome wide association study for backfat thickness in Italian Large White pigs highlights new regions affecting fat deposition including neuronal genes. *BMC Genomics* **13**, 583.

Fontanesi L., Buttazzoni L., Galimberti G., Calò D.G., Scotti E. & Russo V. (2013) Association between melanocortin 4 receptor (*MC4R*) gene haplotypes and carcass and production traits in Italian Large White pigs evaluated with a selective genotyping approach. *Livestock Science*, **157**, 48-56.

Fontanesi L., Schiavo G., Galimberti G., Calò D.G. & Russo V. (2014a) A genomewide association study for average daily gain in Italian Large White pigs. *Journal of Animal Science* **92**, 1385-94.

Fontanesi L., Scotti E., Buttazzoni L., Dall'Olio S. & Russo V. (2014b) Investigation of a short interspersed nuclear element polymorphic site in the porcine vertnin gene: Allele frequencies and association study with meat quality, carcass and production traits in Italian Large White pigs. *Italian Journal of Animal Science* **13**, 61-5.

Fontanesi L., Schiavo G., Scotti E., Galimberti G., Calò D.G., Samorè A.B., Gallo M., Russo V. & Buttazzoni L. (2015) A retrospective analysis of allele frequency changes of major genes during 20 years of selection in the Italian Large White pig breed. *Journal of Animal Breeding and Genetics*, doi:10.1111/jbg.12127.

Glick G., Shirak A., Uliel S., Zeron Y., Ezra E., Seroussi E., Ron M. & Weller J.I. (2012) Signatures of contemporary selection in the Israeli Holstein dairy cattle. *Animal Genetics* **43** (Suppl. 1), 45-55.

Hayes B. & Goddard M.E. (2001) The distribution of the effects of genes affecting quantitative traits in livestock. *Genetics Selection Evolution* **33**, 209-29.

Hill W.G. (2014) Applications of population genetics to animal breeding, from Wright, Fisher and Lush to genomic prediction. *Genetics* **196**, 1-16.

Hosmer D.W., Lemeshow S. & Sturdivant R.X. (2013) *Applied logistic regression*. Third edition, Wiley, New York.

Hu Z.L., Park C.A., Wu X.L. & Reecy J.M. (2013) Animal QTLdb: an improved database tool for livestock animal QTL/association data dissemination in the post-genome era. *Nucleic Acids Research* **41** (Database issue), D871-9.

Huang da W., Sherman B.T. & Lempicki R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols* **4**, 44-57.

Kemper K.E., Saxton S.J., Bolormaa S., Hayes B.J. & Goddard M.E. (2014) Selection for complex traits leaves little or no classic signatures of selection. *BMC Genomics* **15**, 246.

Kim K.S., Larsen N., Short T., Plastow G. & Rothschild M.F. (2000) A missense variant of the porcine melanocortin-4 receptor (*MC4R*) gene is associated with fatness, growth, and feed intake traits. *Mammalian Genome* **11**, 131-5.

Larson G. & Burger J. (2013) A population genetics view of animal domestication. *Trends in Genetics* **29**, 197-205.

Liu H., Sørensen A.C., Meuwissen T.H. & Berg P (2014) Allele frequency changes due to hitch-hiking in genomic selection programs. *Genetics Selection Evolution* **46**, 8.

Meyer R.C., Giddens M.M., Coleman B.M. & Hall R.A. (2014) The protective role of prosaposin and its receptors in the nervous system. *Brain Research* **1585**, 1-12.

Mikawa S., Sato S., Nii M., Morozumi T., Yoshioka G., Imaeda N., Yamaguchi T., Hayashi T. & Awata T. (2011) Identification of a second gene associated with variation in vertebral number in domestic pigs. *BMC Genetics* **12**, 5.

Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M.A., Bender D., Maller J., Sklar P., de Bakker P.I., Daly M.J. & Sham P.C. (2007) PLINK: a toolset for whole-genome association and population-based linkage analysis. *American Journal of Human Genetics* **81**, 559-75

Ramos A.M., Crooijmans R.P., Affara N.A., Amaral A.J., Archibald A.L., Beever J.E., Bendixen C., Churcher C., Clark R., Dehais P., Hansen M.S., Hedegaard J., Hu Z.L., Kerstens H.H., Law A.S., Megens H.J., Milan D., Nonneman D.J., Rohrer G.A., Rothschild M.F., Smith T.P., Schnabel R.D., Van Tassell C.P., Taylor J.F., Wiedmann R.T., Schook L.B. & Groenen M.A. (2009) Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology. *PLoS One* **4**, e6524.

R Core Team (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Rubin C.J., Megens H.J., Martinez Barrio A., Maqbool K., Sayyab S., Schwochow D., Wang C., Carlborg Ö., Jern P., Jørgensen C.B., Archibald A.L., Fredholm M., Groenen M.A. &

Andersson L. (2012) Strong signatures of selection in the domestic pig genome. *Proceedings of the National Academy of Sciences of the USA* **109**, 19529-36.

Smith J.M. & Haigh J. (1974) The hitch-hiking effect of a favourable gene. *Genetical Research* **23**, 23-35.

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.

Venables W.N. & Ripley B.D. (2002) *Modern Applied Statistics with S.* Fourth Edition. Springer, New York.

Walsh B. (2007) Evolutionary Quantitative Genetics, pp. 533-586, in Balding D.J., Bishop M. & Cannings C. (Ed.) *Handbook of Statistical Genetics*, Third Edition, Vol. 1, John Wiley & Sons, Ltd., Chichester, England.

Figures

Figure 1. Distribution of single nucleotide polymorphisms with $P_{Bonferroni}$ <0.10 (green), <0.05 (blue) and <0.01 (red) along all pig autosomes.

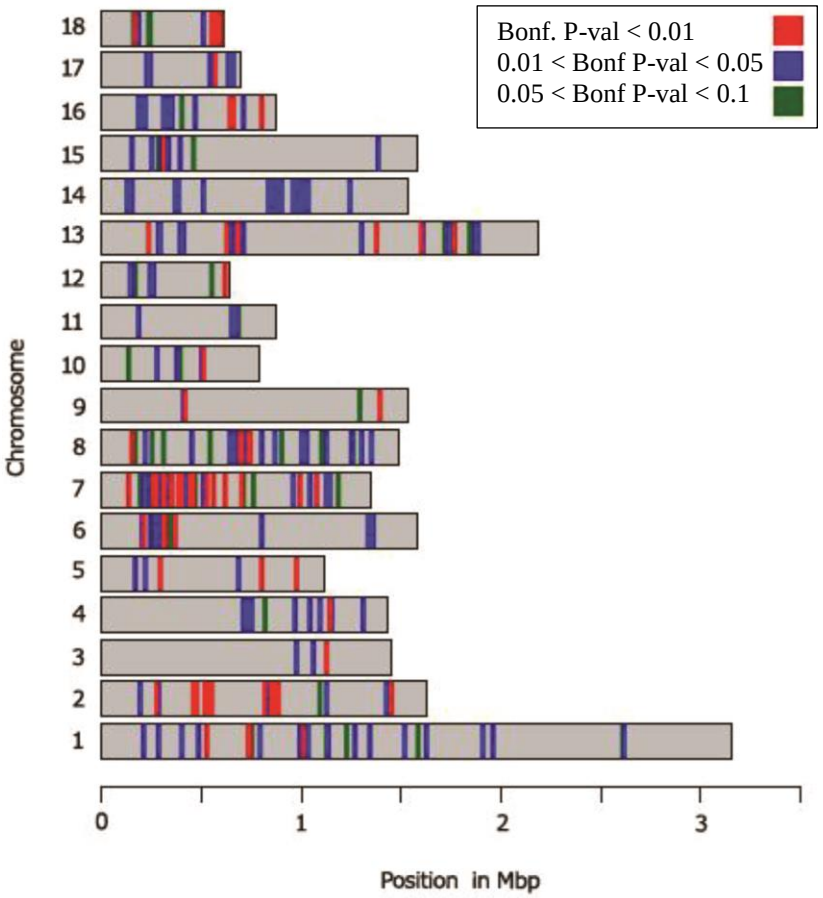


Figure 2. Manhattan plot for the genotype frequency change tests over the considered period, including the three thresholds ($P_{Bonferroni}$<0.10, green line; $P_{Bonferroni}$<0.05, blue line; $P_{Bonferroni}$<0.01, red line).

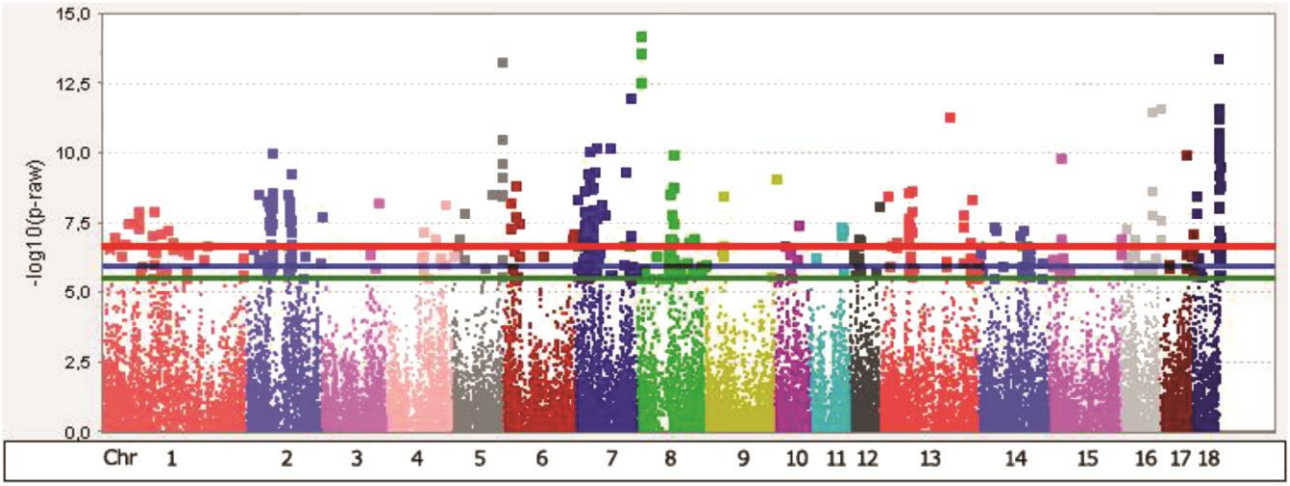
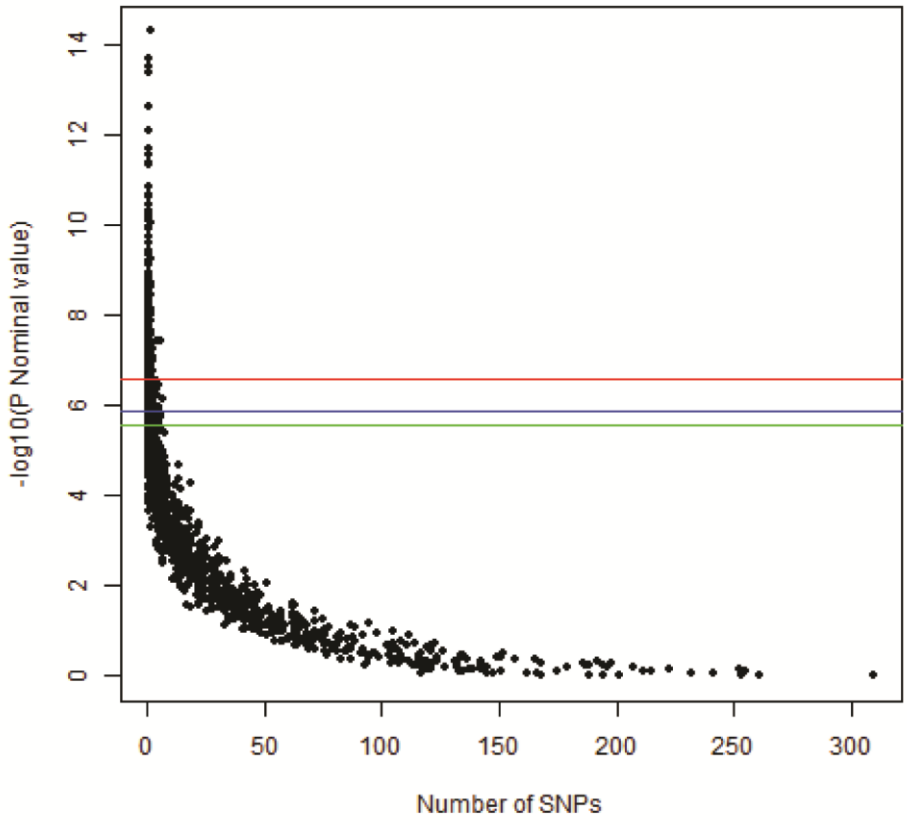


Figure 3. Distribution of the number of single nucleotide polymorphisms (SNPs) with different P values among the considered 38666 SNPs. The –log10 scaled P values (Y axis) were divided in 1333 ranges with a step of 0.005 for each bin. The number of SNPs withing these bins is reported on the X axis.



Tables

Table 1. Details about the boars genotyped and divided according to their year of birth in eight different groups. Averaged reliability of estimated breeding values (EBVs) ± SD and averaged EBVs for several production traits are reported (ADG = average daily gain; BFT = back fat thickness; FGR = feed gain ratio; LC = lean meat cuts; HW = ham weight).

Years	N. of boars	EBV reliability	ADG EBV (g)	BFT EBV (mm)	FGR EBV	HW EBV (kg)	LC EBV (kg)
1992-1995	20	0.904 ± 0.013	11.100 ± 31.236	-2.450 ± 2.803	-0.069 ± 0.177	0.455 ± 0.626	0.968 ± 1.379
1996-1997	24	0.914 ± 0.017	36.375 ± 33.618	-3.726 ±3.577	-0.210 ± 0.179	0.568 ± 0.639	2.100 ±1.736
1998-1999	24	0.930 ± 0.027	40.542 ± 31.518	-3.787 ± 3.463	-0.168 ± 0.164	0.691 ± 0.590	2.894 ± 1.711
2000-2001	26	0.941 ± 0.025	35.615 ± 29.471	-2.364 ± 3.403	-0.127 ± 0.150	0.558 ± 0.690	2.370 ± 1.870
2002-2003	26	0.937 ± 0.028	47.423 ± 26.117	-2.227 ± 3.436	-0.188 ± 0.126	0.629 ± 0.621	3.458 ± 1.478
2004-2005	29	0.932 ± 0.025	41.759 ± 17.061	-1.139 ± 3.208	-0.152 ± 0.088	0.506 ± 0.415	3.241 ± 1.359
2006-2007	19	0.940 ± 0.030	50.000 ± 15.320	-0.497 ± 2.164	-0.177 ± 0.092	0.785 ± 0.403	3.806 ± 1.311
2008-2010 [1]	24	0.941 ± 0.026	51.875 ± 27.934	-1.750 ± 2.329	-0.188 ± 0.133	0.659 ± 0.633	4.321 ± 1.494

[1] No boars with EBV reliability >0.85 were available in the years 2011 and 2012.

Table 2. List of the most significant ($P_{nominal\ value}$ <1.0E-10) single nucleotide polymorphisms (SNPs; in increasing order of P value) with the closest annotated gene in Sscrofa10.2. All other SNPs with $P_{Bonferroni}$<0.10 are reported in Table S3.

SNP	SSC	Position	A1/A2	F1	F8	$P_{Bonferroni}$	$P_{Nominal\ value}$	P(AB)	P(BB)	Ensembl ID	Distance	Gene symbol	Gene description
ASGA0085562	8	3625594	A/G	0.425	0.021	1.93E-10	5.00E-15	4.87E-02	4.41E-03	ENSSSCG00000008714	46334	PSAPL1	proactivator polypeptide-like 1-like
ASGA0088966	8	3630272	G/A	0.425	0.021	1.93E-10	5.00E-15	4.87E-02	4.41E-03	ENSSSCG00000008714	41656	PSAPL1	proactivator polypeptide-like 1-like
ASGA0085170	8	3378292	A/G	0.225	0.750	8.12E-10	2.10E-14	1.12E-03	5.00E-10	ENSSSCG00000028085	492	CCDC96	coiled-coil domain containing 96
MARC0051792	18	53199966	G/A	0.125	0.562	1.12E-09	2.90E-14	1.26E-02	6.91E-07	ENSSSCG00000016719	78035	STK31	serine/threonine kinase 31
ASGA0026958	5	103291948	G/A	1.000	0.562	1.55E-09	4.00E-14	2.28E-01	1.19E-02	ENSSSCG00000021240	175646	-	-
MARC0003410	8	3607452	G/A	0.450	0.062	8.74E-09	2.26E-13	1.29E-02	6.59E-06	ENSSSCG00000008714	64476	PSAPL1	proactivator polypeptide-like 1-like
ALGA0044524	7	115091161	A/G	0.150	0.625	2.99E-08	7.72E-13	8.46E-02	9.43E-09	-	-	-	-
MARC0078879	16	81266398	C/A	0.150	0.708	7.60E-08	1.97E-12	4.76E-04	1.54E-09	ENSSSCG00000017102	316138	PAPD7	PAP associated domain containing 7
ALGA0098588	18	52227085	A/G	0.075	0.479	7.76E-08	2.01E-12	1.65E-01	3.77E-05	ENSSSCG00000016715	35068	OSBPL3	oxysterol binding protein-like 3
MARC0039691	16	62193743	A/G	1.000	0.625	1.02E-07	2.64E-12	6.51E-01	8.96E-03	ENSSSCG00000020610	355081	U6	-
MARC0036292	13	150958740	A/G	0.500	0.021	1.59E-07	4.12E-12	4.41E-02	8.25E-04	ENSSSCG00000011908	319416	-	-
ALGA0098742	18	55295182	G/A	0.175	0.625	1.68E-07	4.35E-12	3.90E-02	5.09E-07	ENSSSCG00000016737	4719	PPIA	peptidyl-prolyl cis-trans isomerase A-like
MARC0055353	18	53659823	A/G	0.050	0.396	5.24E-07	1.35E-11	2.01E-01	6.85E-03	ENSSSCG00000016725	34466	TNS3	tensin 3
ALGA0098607	18	52307963	A/G	0.075	0.479	7.97E-07	2.06E-11	7.96E-02	2.27E-05	ENSSSCG00000016715	45810	OSBPL3	oxysterol binding protein-like 3
DRGA0006307	5	102834057	A/C	0.050	0.437	8.81E-07	2.28E-11	1.85E-01	1.53E-02	-	-	-	-
ASGA0080235	18	53188217	A/G	0.175	0.562	1.40E-06	3.61E-11	2.09E-02	2.41E-06	ENSSSCG00000016719	89784	STK31	serine/threonine kinase 31
ALGA0098601	18	52282193	G/A	0.100	0.479	1.77E-06	4.59E-11	1.82E-01	3.37E-05	ENSSSCG00000016715	20040	OSBPL3	oxysterol binding protein-like 3
ALGA0040669	7	40659636	G/A	0.125	0.500	1.98E-06	5.12E-11	6.37E-02	3.40E-06	ENSSSCG00000019053	180510	-	-
MARC0026254	7	70013552	A/C	0.026	0.326	2.05E-06	5.30E-11	2.29E-01	2.54E-01	ENSSSCG00000024478	4019	U1	-
DIAS0000349	7	26927141	A/G	0.075	0.375	2.56E-06	6.61E-11	8.02E-02	3.89E-03	ENSSSCG00000001376	2861	DHX16	DEAH (Asp-Glu-Ala-His) box polypeptide 16
MARC0071707	18	51896701	G/A	0.025	0.396	2.69E-06	6.95E-11	1.86E-01	5.71E-05	ENSSSCG00000016714	139367	CYCS	cytochrome c. somatic
H3GA0006772	2	50664844	A/G	0.028	0.417	2.91E-06	7.53E-11	1.94E-01	3.12E-05	ENSSSCG00000013401	35141	DKK3	dickkopf 3 homolog (Xenopus laevis)
ALGA0095466	17	53392101	G/A	0.100	0.437	3.17E-06	8.19E-11	6.84E-01	1.03E-04	ENSSSCG00000020659	8620		plasmin trypsin inhibitor
ALGA0117771	8	74529183	G/A	0.150	0.625	3.47E-06	8.97E-11	1.84E-01	6.25E-08	ENSSSCG00000008962	9566	EREG	epiregulin
MARC0006179	8	74321384	A/G	0.150	0.625	3.47E-06	8.97E-11	1.84E-01	6.25E-08	ENSSSCG00000008959	9396	CXCL2	chemokine (C-X-C motif) ligand 2
ALGA0098596	18	52256414	A/G	0.125	0.542	4.32E-06	1.12E-10	5.88E-02	1.78E-05	ENSSSCG00000016715	5739	OSBPL3	oxysterol binding protein-like 3
ASGA0068926	15	22635811	C/A	0.100	0.521	4.66E-06	1.20E-10	1.04E-02	5.28E-06	ENSSSCG00000015709	41058	SLC35F5	solute carrier family 35. member F5
MARC0065282	5	103318589	A/C	1.000	0.687	7.09E-06	1.83E-10	3.63E-01	1.02E-01	ENSSSCG00000021240	149005	-	-
ASGA0090395	18	56764782	G/A	1.000	0.685	9.15E-06	2.37E-10	7.01E-01	1.24E-02	ENSSSCG00000025060	83673	MRPL32	mitochondrial ribosomal protein L32
M1GA0010050	7	38163487	A/G	0.150	0.521	1.37E-05	3.55E-10	8.53E-03	3.58E-06	ENSSSCG00000001583	9839	-	-
MARC0027367	7	105314744	G/A	0.450	0.854	1.39E-05	3.60E-10	1.30E-02	2.67E-07	ENSSSCG00000002386	15440	IFT43	intraflagellar transport 43 homolog
DIAS0004400	2	90946055	C/A	0.075	0.500	1.52E-05	3.92E-10	6.08E-03	9.17E-07	ENSSSCG00000014121	7194	-	-
H3GA0020313	7	22947295	A/G	0.150	0.625	1.67E-05	4.33E-10	1.94E-03	8.01E-08	ENSSSCG00000024278	8861	-	zinc finger protein 184-like
ASGA0080153	18	52001599	A/C	0.025	0.396	1.81E-05	4.69E-10	3.30E-01	1.45E-04	ENSSSCG00000016714	34469	CYCS	cytochrome c. somatic
MARC0004117	5	103193204	A/G	0.050	0.437	2.14E-05	5.53E-10	1.57E-01	1.69E-02	ENSSSCG00000021240	274390	-	-
MARC0035348	5	103272927	G/A	0.05	0.437	2.14E-05	5.53E-10	1.57E-01	1.69E-02	ENSSSCG00000021240	194667	-	-

MARC0024811	9	152921459	A/C	0.275	0.750	2.40E-05	6.21E-10	2.67E-04	1.53E-08	ENSSSCG00000029423	112676	-	coiled-coil alpha-helical rod protein 1
INRA0024421	7	27280485	G/A	0.05	0.479	2.83E-05	7.32E-10	1.02E-01	4.50E-04	ENSSSCG00000001391	4874	CCHCR1	-

SNP: SNP name; SSC: *Sus scrofa* chromosome; Position: nucleotide position on the chromosome (Sscrofa10.2 genome version); A1/A2: allele 1 and allele 2; F1: frequency of allele 1 in the first time period (1992-1995); F8: frequency of allele 1 in the last time period (2008-2010); $P_{Bonferroni}$ value: $P_{nominal\ value}$ of genotype frequency change during time; P(AB): P nominal value for the genotype frequency change toward AB genotype; P(BB): $P_{nominal\ value}$ for the genotype frequency change toward BB genotype; Ensembl ID of the closest gene to the SNP (Ssccrofa10.2); Distance: distance of the SNP from the closest gene (in bp); Gene symbol: Symbol of the closest gene; Gene name: Name of the closest gene.

Supplementary Material
Supplementary Tables

Table S1. Distribution of single nucleotide polymorphisms (SNPs) mapped on all autosomes and chromosome X (pseudoautosomal region) and the number of SNPs with $P_{Bonferroni} < 0.10$, < 0.05 and 0.01. The expected and observed number of SNPs with $P_{Bonferroni} < 0.10$ for the different chromosomes is reported.

Chromosome	**No. of SNPs**				**Proportions (P < 0.10)**	
	All	**P < 0.10**	**P < 0.05**	**P < 0.01**	**Expected**	**Observed**
SSC1	4548	41	30	22	0.118	0.083
SSC2	2635	63	57	39	0.068	0.128
SSC3	2096	3	3	1	0.054	0.006
SSC4	2515	15	11	4	0.065	0.030
SSC5	1753	15	14	11	0.045	0.030
SSC6	2390	16	13	10	0.062	0.032
SSC7	2498	77	62	40	0.065	0.156
SSC8	2103	63	45	16	0.054	0.128
SSC9	2500	6	5	4	0.065	0.012
SSC10	1382	7	4	2	0.036	0.014
SSC11	1401	8	6	5	0.036	0.016
SSC12	1230	18	11	6	0.032	0.037
SSC13	3059	39	32	19	0.079	0.079
SSC14	2884	35	22	7	0.075	0.071
SSC15	2071	22	12	4	0.054	0.045
SSC16	1369	19	16	8	0.035	0.039
SSC17	1267	10	9	2	0.033	0.020
SSC18	961	36	28	22	0.025	0.073
SSCX (pseudoaut.)	4	0	0	0	0.000	0.000
Total	38666	493	399	222	1.000	1.000

Table S2. Genomic windows (with the chromosome number and chromosome coordinates) enriched of single nucleotide polymorphisms (SNPs) with PBonferroni <0.10. The P nominal value and the P Bonferroni corrected for the number of analysed windows of the Fisher exact test with Yates correction is reported.

Window	No. SNPs $P_{Bonferroni}<0.10$	No. SNPs $P_{Bonferroni}>0.10$	$P_{Nominal\ value}$	$P_{Bonf\ corrected}$
8:58493243-70033400	31	47	4.01E-09	3.83E-06
2:46185210-50664844	22	56	2.99E-06	0.00285
8:66787880-74321384	20	58	1.16E-05	0.01111
14:99922357-104459994	17	61	8.41E-05	0.08026
7:14569027-17891007	17	61	8.41E-05	0.08026
7:16370134-19232432	17	61	8.41E-05	0.08026
18:50782981-54194183	16	62	0.00016	0.15273
18:52282193-56764782	16	62	0.00016	0.15273
2:48805571-54417713	15	63	0.00030	0.28853
8:51141055-66562625	13	65	0.00106	1
2:89162092-92950820	12	66	0.00196	1
7:31045816-33926273	12	66	0.00196	1
2:44299936-48747164	11	67	0.00361	1
7:21072637-27352011	11	67	0.00361	1
13:60268089-65209288	10	68	0.00660	1
14:97494784-102209608	10	68	0.00660	1
2:87534079-90892618	10	68	0.00660	1
7:32456532-35460037	10	68	0.00660	1
1:72448760-76956941	9	69	0.01204	1
14:102310696-106744337	9	69	0.01204	1
5:100547769-105397108	9	69	0.01204	1
5:98732126-103517562	9	69	0.01204	1
1:69968051-74916127	8	70	0.02186	1
12:11918085-15000845	8	70	0.02186	1
15:19187327-22667341	8	70	0.02186	1
11:63710002-67401150	7	71	0.03955	1
13:58558615-61913936	7	71	0.03955	1
2:38566779-42708585	7	71	0.03955	1
2:85759506-89137126	7	71	0.03955	1
2:90946055-96053472	7	71	0.039551	1
4:71081959-76189394	7	71	0.039551	1
7:23229208-29913003	7	71	0.039551	1
8:71042714-76480189	7	71	0.039551	1

Table S3. List of the remaing most significant ($P_{Bonferroni}$ <0.10) single nucleotide polymorphisms (SNPs; in increasing order of P value) with the closest annotated gene in Sscrofa10.2. The 38 most significant SNPs are reported in Table 2. Nomenclature for Table S3 is the same of Table 2

SNP	SSC	POS	A1/A2	F1	F8	$P_{BonfErroni}$	$P_{nominal\ valuE}$	P(AB)	P(BB)	Ensembl ID	Distance	Gene symbol	Gene description
ASGA0027882	6	23697334	A/G	0.25	0.625	4.24456015E-05	1.09775E-09	0.014667539	3E-08	ENSSSCG00000002795	10995	CDH11	cadherin 11, type 2, OB-cadherin (osteoblast)
ASGA0080358	18	56551972	G/A	1	0.6875	4.308707044E-05	1.11434E-09	0.637386475	0.013188044	ENSSSCG00000027468	78865	-	-
ALGA0040052	7	33032424	G/A	0.45	0.8864	4.714506714E-05	1.21929E-09	0.00022292	2.53E-08	ENSSSCG00000001493	9456	-	primase, DNA, polypeptide 2 (58kDa)
ALGA0048092	8	73810886	G/A	0.175	0.6458	5.236420382E-05	1.35427E-09	0.105452869	8.41E-08	ENSSSCG00000008950	3930	AFM	afamin
ALGA0098680	18	53865344	G/A	0.125	0.4167	5.422945166E-05	1.40251E-09	0.122926878	0.00237759	ENSSSCG00000016725	22201	TNS3	tensin 3
ALGA0039917	7	31179740	A/C	0.325	0.8542	5.616120502E-05	1.45247E-09	0.004016786	1E-07	ENSSSCG00000001485	36817	-	-
ALGA0091060	16	64539317	T/A	0.025	0.375	6.460083284E-05	1.67074E-09	0.510099545	0.006663235	ENSSSCG00000022115	103903	-	-
ALGA0070531	13	66227544	A/C	0.125	0.5208	6.492446726E-05	1.67911E-09	0.226548472	1.08E-05	ENSSSCG00000011530	39753	CRBN	protein cereblon-like
ALGA0038950	7	16780766	A/C	0.075	0.4583	6.757424824E-05	1.74764E-09	0.138498661	6.47E-05	ENSSSCG00000001079	1687	E2F3	E2F transcription factor 3
DRGA0012439	13	60647697	G/A	0.1	0.5625	7.580971958E-05	1.96063E-09	0.161189014	6.87E-06	ENSSSCG00000011521	205807	PDZRN3	PDZ domain containing ring finger 3
MARC0030786	13	60725783	G/A	0.1	0.5625	7.580971958E-05	1.96063E-09	0.161189014	6.87E-06	ENSSSCG00000011521	283893	PDZRN3	PDZ domain containing ring finger 3
ALGA0013695	2	51273448	G/A	0.025	0.3958	7.811846644E-05	2.02034E-09	0.30210255	0.000369677	ENSSSCG00000013403	38379	GALNT18	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 18
M1GA0008041	5	81575651	A/G	0.275	0.02083	8.378264878E-05	2.16683E-09	0.28677048	0.004094302	ENSSSCG00000000816	22539	SENP1	SUMO1/sentrin specific peptidase 1
INRA0008879	2	84097613	A/G	1	0.6667	8.54750596E-05	2.2106E-09	0.135654124	0.001037061	ENSSSCG00000014067	51486	TMEM174	transmembrane protein 174
ALGA0108373	2	18722241	A/G	0.1	0.3958	8.807844138E-05	2.27793E-09	0.458649244	0.000248219	ENSSSCG00000013276	9699	PRDM11	-
ASGA0096260	2	18727675	G/A	0.1	0.3958	8.807844138E-05	2.27793E-09	0.458649244	0.000248219	ENSSSCG00000013276	15133	PRDM11	-
ALGA0033847	5	102816448	G/A	1	0.6875	8.985746404E-05	2.32394E-09	0.334191668	0.10640783	-	-	-	-
ALGA0048053	8	68896436	A/G	0.175	0.6458	8.985746404E-05	2.32394E-09	0.207775499	1.74E-07	ENSSSCG00000021957	452984	5S_rRNA	-
ALGA0068042	13	13619831	A/C	0.45	0.1875	9.83199048E-05	2.5428E-09	0.005438922	2.31E-07	ENSSSCG00000011212	401887	RARB	retinoic acid receptor beta
ALGA0052304	9	35333897	G/A	0.225	0.6042	0.0001028454	2.65984E-09	0.075368591	9.58E-06	ENSSSCG00000014975	193058	CNTN5	contactin 5
ALGA0112743	18	4825126	A/G	0.2	0.4792	0.0001049229	2.71357E-09	0.916164796	7.71E-06	ENSSSCG00000016427	112192	XRCC2	X-ray repair complementing defective repair in Chinese hamster cells 2
ALGA0033870	5	103046152	A/G	0.05	0.4375	0.0001075793	2.78227E-09	0.145635795	0.017802076	ENSSSCG00000021240	421442	-	-
ALGA0013646	2	48747164	G/A	0.15	0.5833	0.0001200881	3.10578E-09	0.004521519	9.95E-08	ENSSSCG00000021170	84678	FAR1	fatty acyl CoA reductase 1
M1GA0009342	7	1611710	A/G	0.125	0.5208	0.000132056	3.4153E-09	0.427773951	4.71E-05	ENSSSCG00000028777	62661	MYLK4	myosin light chain kinase family, member 4
ASGA0094162	2	50426070	G/A	0.05	0.4583	0.0001333834	3.44963E-09	0.105395157	6.33E-06	ENSSSCG00000013400	59322	-	-
ASGA0095968	2	50415007	A/G	0.05	0.4583	0.0001333834	3.44963E-09	0.105395157	6.33E-06	ENSSSCG00000013400	70385	-	-
ASGA0059621	13	198288173	A/C	0.4	0.75	0.0001388268	3.59041E-09	9.64E-05	1.57E-07	-	-	-	-
MARC0064216	2	40774796	G/A	0.175	0.6667	0.0001526619	3.94822E-09	0.001652768	3.5E-08	ENSSSCG00000019003	245408	U6	-
MARC0017959	2	88913126	A/G	0.075	0.4792	0.0001670391	4.32005E-09	0.011546876	2.63E-06	ENSSSCG00000028536	33893	LHFPL2	lipoma HMGIC fusion partner-like 2
MARC0083378	6	10527042	G/A	0.425	0.8958	0.0001756032	4.54154E-09	0.000158425	1.77E-05	ENSSSCG00000019476	151760	5S_rRNA	
MARC0043998	3	121035255	G/A	0.475	0.7917	0.0001836859	4.75058E-09	4.56E-05	6.96E-07	ENSSSCG00000008577	6902	DNAJC27	DnaJ (Hsp40) homolog, subfamily C, member 27
H3GA0021445	7	52184508	A/G	0.475	0.7917	0.0001999817	5.17203E-09	0.000990134	4.96E-07	ENSSSCG00000001746	154479	-	-
MARC0077792	2	47352392	A/G	0.175	0.6042	0.0002040219	5.27652E-09	0.001547092	7.93E-08	ENSSSCG00000013386	1804	CRSP3	ENSSSCG00000013386
DRGA0005119	4	123202552	G/A	0.125	0.3542	0.0002210141	5.71598E-09	0.133485079	0.000114251	ENSSSCG00000006852	7695	NTNG1	netrin G1
MARC0095075	2	50226694	G/A	0.1	0.4792	0.0002277451	5.89006E-09	0.059385678	4.61E-06	ENSSSCG00000028827	66538	PARVA	parvin, alpha
H3GA0034945	12	59563353	G/A	0.3	0.02083	0.0002370396	6.13044E-09	0.179027536	0.000577631	ENSSSCG00000018016	153756	MAP2K4	mitogen-activated protein kinase kinase 4

ASGA0097161	18	52755721	G/A	0.225	0.6667	0.0002394218	6.19205E-09	0.000909372	4.32E-07	ENSSSCG00000016717	94267	MPP6	membrane protein, palmitoylated 6 (MAGUK p55 subfamily member 6)
ALGA0013197	2	41998239	A/C	0.175	0.5833	0.0002542266	6.57494E-09	0.009366147	3.49E-07	ENSSSCG00000013349	17821	HTATIP2	oxidoreductase HTATIP2-like
ALGA0098616	18	52425175	G/A	0.375	0.8125	0.0002726603	7.05168E-09	0.001440626	7.95E-08	ENSSSCG00000016716	9636	DFNA5	deafness, autosomal dominant 5
ASGA0097814	2	50036255	G/A	0.175	0.5833	0.0002795625	7.23019E-09	0.008200383	1.3E-07	ENSSSCG00000013399	28904	TEF1	transcriptional enhancer factor 1
H3GA0006751	2	48767966	C/A	0.15	0.5833	0.0002809642	7.26644E-09	0.00238895	1.65E-07	ENSSSCG00000021170	63876	FAR1	fatty acyl CoA reductase 1
ALGA0108178	2	89176304	G/A	0.075	0.4792	0.00028664	7.41323E-09	0.008088239	2.62E-06	ENSSSCG00000029280	93921	-	arylsulfatase B-like
ASGA0010757	2	89137126	A/G	0.075	0.4792	0.00028664	7.41323E-09	0.008088239	2.62E-06	ENSSSCG00000014103	101993	SCAMP1	
ASGA0032302	7	32957768	A/G	0.4167	0.91304	0.0003120698	8.07091E-09	0.000795095	2.06E-07	ENSSSCG00000001493	65200	-	primase, DNA, polypeptide 2 (58kDa)
ASGA0032266	7	32543114	A/C	0.475	0.95833	0.0003363749	8.6995E-09	0.000708439	2E-07	ENSSSCG00000001492	100319	-	-
ALGA0005150	1	106839739	C/A	0.475	0.08333	0.0003536214	9.14554E-09	0.056539458	3.38E-05	ENSSSCG00000004505	61726	SMAD2	mothers against decapentaplegic homolog 2-like
ASGA0032316	7	33115876	G/A	0.375	0.75	0.0003553939	9.19138E-09	0.000528711	1.83E-06	ENSSSCG00000029539	3916	RAB23	RAB23, member RAS oncogene family
ALGA0013592	2	47598609	A/G	0.175	0.6042	0.0003680528	9.51877E-09	0.001219122	1.05E-07	ENSSSCG00000013388	20398	-	-
ASGA0031485	7	16738844	A/C	0.15	0.6667	0.0003698979	9.56649E-09	0.001060337	2.03E-07	ENSSSCG00000001079	27084	E2F3	E2F transcription factor 3
ASGA0057875	13	67836973	A/G	0.425	0.04167	0.0003830733	9.90724E-09	0.049388418	2.49E-05	ENSSSCG00000011534	11	BHLHE40	basic helix-loop-helix family, member e40
CASI0009829	7	16686700	A/C	0.15	0.6667	0.0003849935	9.9569E-09	0.000989037	2.01E-07	ENSSSCG00000001079	79228	E2F3	E2F transcription factor 3
INRA0002726	1	74096746	G/A	0.25	0.04167	0.0003869229	1.00068E-08	0.10790448	0.989253142	ENSSSCG00000018969	77110	U6	
INRA0018824	5	20429749	A/G	0.075	0.3333	0.000396717	1.02601E-08	0.011283524	0.001258188	ENSSSCG00000000292	4719	ZNF385A	zinc finger protein 385A
MARC0029189	18	4829959	G/A	0.2	0.4792	0.0004254845	1.10041E-08	0.749811266	8.93E-06	ENSSSCG00000016427	107359	XRCC2	X-ray repair complementing defective repair in Chinese hamster cells 2
ALGA0042201	7	59986169	A/G	0.2	0.5	0.0004725875	1.22223E-08	0.174028796	0.000144665	ENSSSCG00000001841	35830	RHCG	-
ASGA0073698	16	64957827	C/A	0.1	0.5833	0.0004773395	1.23452E-08	0.023022998	1.31E-05	ENSSSCG00000019195	138804	U6	-
ASGA0033314	7	48120121	G/A	0.425	0.08333	0.0005145169	1.33067E-08	0.03610382	0.001496536	ENSSSCG00000001720	768	-	mitochondrial uncoupling protein 4-like
DRGA0008622	8	69973631	A/G	0.475	0.8125	0.0005145169	1.33067E-08	0.005189134	3.83E-06	ENSSSCG00000029658	11152	-	transmembrane protease, serine 11D
MARC0065723	13	177964129	A/G	0.475	0.9375	0.0005196865	1.34404E-08	0.087110837	9.91E-06	ENSSSCG00000026130	154014	EPHA3	EPH receptor A3
MARC0006685	6	22591223	A/G	0.225	0.5625	0.0005222926	1.35078E-08	0.073147606	3.11E-07	-	-	-	-
ASGA0101764	2	160400099	A/G	0.2895	0.7727	0.0005772215	1.49284E-08	0.004764861	5.04E-07	ENSSSCG00000024147	41152	-	olfactory receptor 4C46-like
ALGA0034926	6	23679290	A/G	0.25	0.6042	0.0005918334	1.53063E-08	0.012919698	1.3E-07	ENSSSCG00000002795	29039	CDH11	cadherin 11, type 2, OB-cadherin (osteoblast)
ASGA0003244	1	74203064	A/G	0.275	0.06522	0.0006098595	1.57725E-08	0.142647317	0.744894404	ENSSSCG00000004346	183159	POU3F2	POU class 3 homeobox 2
M1GA0010098	7	40341063	G/A	1	0.7708	0.0006508145	1.68317E-08	0.741710382	0.716124105	ENSSSCG00000001595	24063	DAAM2	dishevelled associated activator of morphogenesis 2
MARC0035674	2	89151064	A/C	0.1	0.4792	0.0006606511	1.70861E-08	0.007967189	2.74E-06	ENSSSCG00000014103	115931	SCAMP1	
ALGA0070336	13	60317837	A/G	0.15	0.5625	0.0006706347	1.73443E-08	0.403927929	3.08E-05	ENSSSCG00000011521	104865	PDZRN3	PDZ domain containing ring finger 3
ALGA0121935	16	81322083	A/G	0.325	0.8125	0.0007085545	1.8325E-08	0.002053273	1.19E-07	ENSSSCG00000017102	260453	PAPD7	PAP associated domain containing 7
ASGA0010866	2	91637623	G/A	0.3	0.75	0.0007714138	1.99507E-08	0.001293128	1.28E-07	ENSSSCG00000020126	24823	-	-
INRA0024695	7	33388999	A/G	0.15	0.5417	0.0007988937	2.06614E-08	0.346552843	0.000289776	ENSSSCG00000025172	14287	-	-
ALGA0013598	2	47862965	A/G	0.175	0.6042	0.0008232223	2.12906E-08	0.002181464	1.65E-07	ENSSSCG00000013391	103053	-	proteasome (prosome, macropain) subunit, alpha type, 1
ASGA0010692	2	87792348	A/G	0.125	0.4792	0.0008232223	2.12906E-08	0.007458739	2.22E-06	ENSSSCG00000014097	27020	PDE8B	phosphodiesterase 8B
DBWU0000851	2	41979860	A/G	0.35	0.7708	0.0008525466	2.2049E-08	0.000360275	2.66E-07	ENSSSCG00000013349	558	HTATIP2	oxidoreductase HTATIP2-like
ALGA0039018	7	18009952	G/A	0.225	0.6875	0.0008741261	2.26071E-08	0.002876444	1.42E-07	ENSSSCG00000001081	155529	SOX4	SRY (sex determining region Y)-box 4

MARC0079844	2	51317481	A/G	0.05	0.3958	0.0009281812	2.40051E-08	0.623536669	0.000906241	ENSSSCG00000013403	5654	GALNT18	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 18
ASGA0031783	7	22240376	G/A	0.125	0.5	0.0009516785	2.46128E-08	0.16011927	3.78E-05	ENSSSCG00000026926	2349	-	olfactory receptor 2M2-like
ALGA0109366	2	49486882	A/G	0.1	0.4792	0.000985577	2.54895E-08	0.062196615	1.05E-05	ENSSSCG00000028546	23972	-	-
H3GA0001561	1	48586569	G/A	0.5	0.2083	0.0009905147	2.56172E-08	8.35E-05	4.69E-07	ENSSSCG00000004249	35338	CEP85L	centrosomal protein 85kDa-like
H3GA0024948	8	75092587	A/G	0.225	0.625	0.0010054861	2.60044E-08	0.127741682	4.83E-07	ENSSSCG00000008966	52364	PARM1	prostate androgen-regulated mucin-like protein 1
MARC0049625	6	29304575	A/C	0.15	0.5833	0.0010054861	2.60044E-08	0.000238949	2.86E-07	ENSSSCG00000002835	2023	TOX3	TOX high mobility group box family member 3
ASGA0010360	2	50721181	A/G	0.025	0.3333	0.0010465224	2.70657E-08	0.171603772	0.000540945	ENSSSCG00000013401	21196	DKK3	dickkopf 3 homolog (Xenopus laevis)
ASGA0032304	7	33017627	G/A	0.45	0.8958	0.0010517655	2.72013E-08	0.001753881	4.29E-07	ENSSSCG00000001493	5341	-	primase, DNA, polypeptide 2 (58kDa)
MARC0064720	2	41937766	A/G	0.2	0.5833	0.0010946886	2.83114E-08	0.010656086	6.39E-07	ENSSSCG00000023239	7685	-	
ALGA0038890	7	16218600	G/A	0.125	0.625	0.0011224005	2.90281E-08	0.016607009	5.99E-07	ENSSSCG00000024312	35609	ID4	inhibitor of DNA binding 4, dominant negative helix-loop-helix protein
ALGA0058647	10	46295711	G/A	0.425	0.0625	0.0012098205	3.1289E-08	0.009598029	0.000169574	ENSSSCG00000025628	83340	-	-
SIRI0000656	11	64368096	A/G	0.175	0.6042	0.0012404478	3.20811E-08	6.28E-05	2.04E-07	-	-	-	-
M1GA0018460	14	29858736	G/A	0.175	0.5208	0.0013040456	3.37259E-08	0.552860317	3.53E-05	ENSSSCG00000009755	13033	AACS	-
ALGA0062575	11	64041222	G/A	0.225	0.6458	0.0013777778	3.56328E-08	0.001482274	1.82E-07	ENSSSCG00000009488	376286	SLITRK5	SLIT and NTRK-like family, member 5
ALGA0039910	7	31157864	G/A	0.225	0.7292	0.0013916241	3.59909E-08	0.000105638	4.1E-07	ENSSSCG00000001485	14941	-	-
ALGA0072779	13	177981732	A/G	0.5	0.0625	0.0013916241	3.59909E-08	0.156478918	3E-05	ENSSSCG00000018474	148435	-	-
ASGA0098788	11	64362716	C/A	0.175	0.6042	0.0013985995	3.61713E-08	6.28E-05	2.04E-07	-	-	-	-
MARC0111972	13	62625298	G/A	0.375	0.7708	0.0013985995	3.61713E-08	0.005914762	1.95E-07	ENSSSCG00000011524	175057	CHL1	cell adhesion molecule L1-like
ALGA0003981	1	74287712	A/G	0.275	0.0625	0.0014056096	3.63526E-08	0.081067606	0.748667824	ENSSSCG00000004346	98511	POU3F2	POU class 3 homeobox 2
ALGA0003995	1	74410201	A/G	0.275	0.0625	0.0014056096	3.63526E-08	0.081067606	0.748667824	ENSSSCG00000004347	12158	FBXL4	F-box and leucine-rich repeat protein 4
ALGA0004000	1	74467285	G/A	0.275	0.0625	0.0014056096	3.63526E-08	0.081067606	0.748667824	ENSSSCG00000004347	17583	FBXL4	F-box and leucine-rich repeat protein 4
ALGA0004002	1	74502751	A/G	0.275	0.0625	0.0014056096	3.63526E-08	0.081067606	0.748667824	ENSSSCG00000004347	17883	FBXL4	F-box and leucine-rich repeat protein 4
ALGA0004005	1	74583833	A/G	0.275	0.0625	0.0014056096	3.63526E-08	0.081067606	0.748667824	ENSSSCG00000004347	98965	FBXL4	F-box and leucine-rich repeat protein 4
MARC0075306	1	74264584	A/G	0.275	0.0625	0.0014056096	3.63526E-08	0.081067606	0.748667824	ENSSSCG00000004346	121639	POU3F2	POU class 3 homeobox 2
ALGA0070329	13	60367610	G/A	0.125	0.5625	0.0015612248	4.03772E-08	0.146081197	1.52E-05	ENSSSCG00000011521	55092	PDZRN3	PDZ domain containing ring finger 3
ALGA0109170	6	9491043	A/G	0.325	0.8542	0.0015690508	4.05796E-08	0.000613711	1.71E-07	ENSSSCG00000002694	149096	CLEC3A	C-type lectin domain family 3 member A-like
DRGA0015754	16	7300411	A/G	0.05263	0.3542	0.0015848188	4.09874E-08	0.087693012	0.000911343	ENSSSCG00000020244	43633	7SK	-
INRA0036882	11	66176962	A/G	0.025	0.3125	0.0016249387	4.2025E-08	0.145848873	0.000458028	ENSSSCG00000018617	433098	ssc-mir-17	microRNA mir-17
ALGA0076653	14	30970606	C/A	0.15	0.4167	0.0016577661	4.2874E-08	0.193988267	0.00134076	ENSSSCG00000009767	9743	-	-
ASGA0004417	1	138070556	C/A	0.3	0.7292	0.0016577661	4.2874E-08	7.9E-05	2.62E-07	ENSSSCG00000004663	25723	SEMA6D	sema domain, transmembrane domain (TM), and cytoplasmic domain, (semaphorin) 6D
MARC0109399	18	51864618	A/C	0.15	0.5208	0.0016828217	4.3522E-08	0.010160927	1.73E-06	ENSSSCG00000016714	171450	CYCS	ENSSSCG00000016714
ALGA0013597	2	47821412	A/G	0.1	0.4792	0.0017082523	4.41797E-08	0.060060259	1.26E-05	ENSSSCG00000013390	118189	CALCB	calcitonin-related polypeptide beta
H3GA0041266	14	93589870	G/A	0.375	0.875	0.0018505277	4.78593E-08	0.004422334	2.82E-06	ENSSSCG00000021303	12030	-	glutamate receptor ionotropic, delta-1-like
ALGA0014325	2	90444137	G/A	0.5	0.125	0.00188791	4.88261E-08	0.00337865	1.45E-05	ENSSSCG00000014116	47435	-	metaxin-3-like
ALGA0104357	2	90455142	A/G	0.5	0.125	0.00188791	4.88261E-08	0.00337865	1.45E-05	ENSSSCG00000014117	47133	-	-
ALGA0025668	4	75590861	C/A	0.4	0.7708	0.0020046542	5.18454E-08	0.008248644	6.03E-07	ENSSSCG00000022092	181617	-	25-hydroxycholesterol 7-alpha-hydroxylase-like
ALGA0038883	7	16239090	A/C	0.15	0.5417	0.0020147035	5.21053E-08	0.059290509	8.32E-06	ENSSSCG00000024312	56099	ID4	inhibitor of DNA binding 4, dominant negative helix-loop-helix protein

ALGA0040291	7	35564543	A/G	0.2	0.7083	0.0020969229	5.42317E-08	0.012973892	4.72E-07	ENSSSCG00000001534	7912	ANKS1A	ankyrin repeat and sterile alpha motif domain containing 1A
DRGA0012492	13	65588898	G/A	0.225	0.7083	0.0021179997	5.47768E-08	0.003006576	2.93E-07	ENSSSCG00000027386	18733	-	-
ALGA0039015	7	17970658	G/A	0.225	0.6875	0.0022377522	5.78739E-08	0.002374004	3.43E-07	ENSSSCG00000001081	116235	SOX4	SRY (sex determining region Y)-box 4
DRGA0007230	7	17994605	C/A	0.225	0.6875	0.0022377522	5.78739E-08	0.002374004	3.43E-07	ENSSSCG00000001081	140182	SOX4	SRY (sex determining region Y)-box 4
MARC0087843	6	149273185	A/G	0.05263	0.4583	0.0022489692	5.8164E-08	0.00894861	1E-05	ENSSSCG00000003877	33721	FAF1	Fas (TNFRSF6) associated factor 1
ASGA0004239	1	121978687	A/C	0.2	0.5625	0.0022602404	5.84555E-08	0.015291462	3.97E-07	ENSSSCG00000004575	12543	-	60S ribosomal protein L23a-like
MARC0035967	1	121999146	A/G	0.2	0.5625	0.0022602404	5.84555E-08	0.015291462	3.97E-07	ENSSSCG00000004575	33002	-	60S ribosomal protein L23a-like
H3GA0049787	17	64982346	A/G	0.1	0.4375	0.0024484897	6.33241E-08	0.47153132	0.00023572	ENSSSCG00000007504	20134	RNPC1	RNA binding motif protein 38
MARC0022069	7	115286590	G/A	0.2	0.5833	0.002485493	6.42811E-08	0.04721138	2.86E-06	-	-	-	-
ALGA0079160	14	85857890	G/A	0.4	0.08333	0.0025611894	6.62388E-08	0.361640357	0.002565139	ENSSSCG00000010325	61308	KCNMA1	potassium large conductance calcium-activated channel, subfamily M, alpha member 1
ALGA0005010	1	104500632	A/G	0.4	0.08333	0.0025998941	6.72398E-08	0.012638141	0.00048948	ENSSSCG00000004490	6369	SETBP1	SET binding protein 1
ALGA0005012	1	104531602	A/G	0.4	0.08333	0.0025998941	6.72398E-08	0.012638141	0.00048948	ENSSSCG00000004490	24601	SETBP1	SET binding protein 1
MARC0009324	11	64065725	A/G	0.275	0.6667	0.0026260246	6.79156E-08	0.001619862	3.59E-07	ENSSSCG00000009488	400789	SLITRK5	SLIT and NTRK-like family, member 5
ALGA0001463	1	19798005	A/G	0.3	0.5625	0.0030357875	7.85131E-08	0.074426129	2.04E-06	ENSSSCG00000021079	64211	U6	-
DIAS0000880	18	56431168	A/C	0.175	0.5625	0.0032396695	8.3786E-08	0.420996009	1.54E-05	ENSSSCG00000027468	2	-	-
MARC0041144	16	81251409	A/G	0.125	0.5417	0.0032886322	8.50523E-08	0.129226094	1.57E-05	ENSSSCG00000027191	324754	-	-
ASGA0031545	7	18170018	G/A	0.225	0.6667	0.0033550643	8.67704E-08	0.019324088	7.62E-07	ENSSSCG00000001082	242475	PRL	prolactin
ASGA0031546	7	18197612	A/G	0.225	0.6667	0.0033550643	8.67704E-08	0.019324088	7.62E-07	ENSSSCG00000001082	214881	PRL	prolactin
M1GA0006114	4	101720321	A/G	0.125	0.5	0.0033887849	8.76425E-08	0.017376429	1.37E-05	ENSSSCG00000006461	11303	ARHGEF11	Rho guanine nucleotide exchange factor (GEF) 11
M1GA0020805	15	152354557	G/A	0.2	0.6458	0.0033887849	8.76425E-08	0.025021225	8.12E-06	ENSSSCG00000025228	9039	TRAF3IP1	TNF receptor-associated factor 3 interacting protein 1
CASI0008857	6	146248545	G/A	0.35	0.75	0.0034057709	8.80818E-08	0.096171773	1.26E-06	ENSSSCG00000028748	33523	NDC1	NDC1 transmembrane nucleoporin
ALGA0048019	8	67115392	A/G	0.5	0.1875	0.0034228419	8.85233E-08	0.009956708	1.17E-05	ENSSSCG00000019754	21160	U6	-
ALGA0030510	5	10866994	C/A	0.35	0.7083	0.003439998	8.8967E-08	0.002475394	6.91E-07	-	-	-	-
MARC0046672	12	15740923	A/C	1	0.8125	0.0035270816	9.12192E-08	0.72292923	0.689278957	ENSSSCG00000017298	122927	TANC2	tetratricopeptide repeat, ankyrin repeat and coiled-coil containing 2
MARC0039041	16	24398034	A/C	0.325	0.6667	0.0035625306	9.2136E-08	0.000320192	3.3E-07	ENSSSCG00000016846	258064	WDR70	WD repeat domain 70
INRA0048907	15	22899746	G/A	0.025	0.2917	0.0035803865	9.25978E-08	0.870169137	0.078428754	ENSSSCG00000030127	34804	U6	-
ASGA0057740	13	60379153	G/A	0.1	0.5	0.0036527152	9.44684E-08	0.172693869	3.96E-05	ENSSSCG00000011521	43549	PDZRN3	PDZ domain containing ring finger 3
ALGA0065273	12	14961107	A/G	0.125	0.4583	0.0036710274	9.4942E-08	0.003861011	6.39E-05	ENSSSCG00000017280	2807	ICAM2	intercellular adhesion molecule-2
MARC0112630	12	14932788	G/A	0.125	0.4583	0.0036710274	9.4942E-08	0.003861011	6.39E-05	ENSSSCG00000017279	13854	ERN1	endoplasmic reticulum to nucleus signaling 1
ALGA0049233	8	120608511	G/A	0.35	0.7292	0.0037265053	9.63768E-08	2.8E-05	1.27E-06	ENSSSCG00000009139	12096	PLA2G12A	phospholipase A2, group XIIA
ASGA0038898	8	70033400	G/A	0.4	0.7917	0.0037828262	9.78334E-08	0.000905837	1.08E-06	ENSSSCG00000008926	11998	TMPRSS11D	transmembrane protease serine 11D-like
MARC0020652	12	16142622	A/G	0.2	0.6875	0.0038399939	9.93119E-08	0.034598876	5.66E-07	ENSSSCG00000017304	8705	EFCAB3	EF-hand calcium-binding domain-containing protein 3-like
H3GA0020623	7	32858692	G/A	0.5	0.1042	0.0039175618	1.01318E-07	0.003436396	1.79E-06	ENSSSCG00000019940	74205	U6	-
MARC0080097	7	32851504	G/A	0.5	0.1042	0.0039175618	1.01318E-07	0.003436396	1.79E-06	ENSSSCG00000019940	67017	U6	-
ASGA0085258	4	102103941	A/G	0.025	0.2708	0.0039372041	1.01826E-07	0.673277131	0.642821047	ENSSSCG00000006477	8132	BCAN	brevican
ASGA0038884	8	69393260	A/G	0.5	0.1875	0.0045288712	1.17128E-07	0.00941523	1.26E-05	ENSSSCG00000020337	501	SNORA31	-
ALGA0039405	7	22888300	A/C	0.275	0.6458	0.004620355	1.19494E-07	0.000340854	5.14E-07	ENSSSCG00000029919	9462	-	zinc finger protein 391

MARC0048245	13	32393169	A/G	0.225	0.6875	0.0046901858	1.213E-07	0.000312149	4.39E-07	ENSSSCG00000024270	22390	CCR2	c-C chemokine receptor type 2-like
DRGA0012449	13	60980258	A/C	0.1	0.4792	0.0047373197	1.22519E-07	0.188537129	9E-05	-	-	-	-
H3GA0019865	7	9853300	C/A	0.45	0.8125	0.0047373197	1.22519E-07	0.139582633	2.22E-05	ENSSSCG00000025580	37066	-	-
MARC0056240	1	147576466	A/G	0.475	0.8333	0.0048815825	1.2625E-07	0.005896471	1.33E-05	ENSSSCG00000030198	281611	ENV	-
ASGA0038938	8	75190777	G/A	0.225	0.625	0.0049306497	1.27519E-07	0.074591388	1.7E-06	ENSSSCG00000029348	115019	THAP6	THAP domain containing 6
H3GA0025242	8	109273139	G/A	0.075	0.4583	0.0049306497	1.27519E-07	0.052914487	3.41E-06	ENSSSCG00000009092	51546	TRPC3	transient receptor potential cation channel, subfamily C, member 3
ALGA0121208	13	196120269	G/A	0.325	0.08333	0.005005159	1.29446E-07	0.108003979	0.027083706	ENSSSCG00000018744	61037	U6	-
ALGA0013605	2	47984857	A/G	0.275	0.7083	0.0050808284	1.31403E-07	0.005322716	4.45E-07	ENSSSCG00000013392	2016	COPB1	coatomer protein complex, subunit beta 1
DRGA0017057	18	52952847	A/G	0.275	0.5833	0.0051062706	1.32061E-07	0.009934508	5.57E-06	ENSSSCG00000016718	15280	NPY	neuropeptide Y
H3GA0001513	1	43607958	A/G	0.075	0.25	0.0053413212	1.3814E-07	0.736437922	0.713550964	ENSSSCG00000027613	76502	TRDN	triadin-like
ASGA0010881	2	91598043	A/G	0.05	0.3542	0.0055592815	1.43777E-07	0.061180927	0.00037756	ENSSSCG00000020126	14673	-	-
ALGA0007548	1	221482056	G/A	0.375	0.7708	0.0058151731	1.50395E-07	5.91E-05	2.1E-06	ENSSSCG00000021248	43632	U6	-
ASGA0042433	9	34828091	G/A	0.125	0.5	0.0059030609	1.52668E-07	0.536217685	0.000619035	ENSSSCG00000014975	260634	CNTN5	contactin 5
ALGA0057399	10	18781303	A/G	0.275	0.5833	0.0059326404	1.53433E-07	0.754721997	0.000437263	ENSSSCG00000030502	25825	-	centrosomal protein of 170 kDa-like
ASGA0085087	13	20479372	G/A	0.475	0.04167	0.0059922634	1.54975E-07	0.662634964	0.034011176	ENSSSCG00000024546	17477	-	-
ALGA0098565	18	51975074	A/G	0.075	0.3958	0.0060223068	1.55752E-07	0.294779484	0.000319543	ENSSSCG00000016714	60994	CYCS	ENSSSCG00000016714
H3GA0041574	14	101227861	C/A	0.225	0.5417	0.0061439887	1.58899E-07	0.000459494	1.47E-06	ENSSSCG00000021978	441938	RPL37A	60S ribosomal protein L37a-like
ALGA0065183	12	14208800	G/A	0.275	0.6667	0.0063310935	1.63738E-07	0.030197705	8.55E-07	ENSSSCG00000019526	4199	5S_rRNA	-
MARC0025074	7	16442858	A/G	0.1	0.5833	0.0066224805	1.71274E-07	0.226507305	3.17E-05	ENSSSCG00000001078	13414	MBOAT1	membrane bound O-acyltransferase domain containing 1
ALGA0119911	2	50269665	C/A	0.175	0.5	0.0068241623	1.7649E-07	0.015185304	2.48E-06	ENSSSCG00000028827	23567	PARVA	parvin, alpha
M1GA0024860	2	51003535	G/A	0.025	0.3125	0.0068583818	1.77375E-07	0.268179552	0.010393423	ENSSSCG00000021199	102446	-	-
MARC0005311	2	51067606	G/A	0.025	0.3125	0.0068583818	1.77375E-07	0.268179552	0.010393423	ENSSSCG00000019714	134905	7SK	-
ASGA0031989	7	27389510	A/G	0.2	0.5625	0.0069620453	1.80056E-07	0.019247309	1.73E-06	ENSSSCG00000001386	5146	DPCR1	-
DIAS0000201	13	34182534	G/C	0.05	0.3542	0.0070319921	1.81865E-07	0.436540279	0.007966367	ENSSSCG00000021494	7948	-	M-phase inducer phosphatase 1-like
MARC0055599	6	20274364	G/A	0.175	0.4375	0.0074668299	1.93111E-07	0.965072963	0.005338321	-	-	-	-
ALGA0110805	18	46355992	G/A	0.325	0.625	0.007579696	1.9603E-07	0.007186232	5.66E-07	ENSSSCG00000016672	30774	PAC1	adenylate cyclase activating polypeptide 1 (pituitary) receptor type I
DRGA0003111	2	91003962	C/A	0.4722	0.8696	0.007579696	1.9603E-07	0.028053315	8.82E-05	ENSSSCG00000014124	524	ANKRD34B	ankyrin repeat domain 34B
ALGA0074194	14	873585	G/A	0.15	0.5	0.007849662	2.03012E-07	0.046460898	3.58E-06	ENSSSCG00000009580	9545	S1PR3	sphingosine-1-phosphate receptor 3
ALGA0006042	1	147666164	G/A	0.475	0.8333	0.0080483666	2.08151E-07	0.005896471	1.33E-05	ENSSSCG00000030198	371309	ENV	-
MARC0078284	1	10318726	G/A	0.275	0.7083	0.0080483666	2.08151E-07	0.000220559	5.82E-07	ENSSSCG00000024876	53642	-	transmembrane protein 181
ALGA0103374	2	8162977	G/A	0.175	0.4773	0.0084188255	2.17732E-07	0.05117617	2.95E-05	ENSSSCG00000023571	2911	POAT1-1	organic anion transporter-like
ALGA0080399	14	109595459	G/A	0.4211	0.7708	0.0085460753	2.21023E-07	0.003524689	4.1E-06	ENSSSCG00000023141	7353	RNLS	-
ASGA0103481	6	18273846	A/G	0.025	0.2917	0.0089394245	2.31196E-07	0.701997985	0.039896652	ENSSSCG00000021911	10383	-	NDRG family member 4
MARC0022007	15	152383856	C/A	0.175	0.5625	0.0093977713	2.4305E-07	0.037546642	3.88E-05	ENSSSCG00000016340	1181	ASB1	ankyrin repeat and SOCS box containing 1
MARC0101759	9	34148025	A/G	0.325	0.7292	0.0093977713	2.4305E-07	0.120203163	1.64E-06	-	-	-	-
H3GA0015339	5	5646876	G/A	0.375	0.75	0.0095397915	2.46723E-07	0.012317185	3.03E-06	ENSSSCG00000000079	900	-	family with sequence similarity 83, member F
H3GA0002301	1	104514556	A/G	0.4	0.08333	0.0095876214	2.4796E-07	0.013043032	0.000227104	ENSSSCG00000004490	7555	SETBP1	SET binding protein 1

ALGA0006485	1	168692866	A/G	0.05	0.4167	0.0100289164	2.59373E-07	0.086782023	0.000528491	ENSSSCG00000004878	299284	SOCS6	suppressor of cytokine signaling 6
DIAS0004583	16	22332270	G/A	0.2	0.6458	0.01012968	2.61979E-07	0.01378563	7.15E-07	ENSSSCG00000016833	446	CAPSL	calcyphosine-like
M1GA0010251	7	48279545	A/C	0.45	0.7917	0.0102827974	2.65939E-07	0.007831756	2.29E-05	ENSSSCG00000001724	6632	MEP1A	meprin A, alpha (PABA peptide hydrolase)
ALGA0048001	8	66072630	A/G	0.2	0.5833	0.010438196	2.69958E-07	0.236553029	4.42E-06	ENSSSCG00000024681	211928	TECRL	trans-2,3-enoyl-CoA reductase-like
ALGA0109418	18	56860445	A/G	0.05	0.3542	0.010438196	2.69958E-07	0.270060482	0.002560988	ENSSSCG00000020769	2759	PSMA2	proteasome (prosome, macropain) subunit, alpha type, 2
ASGA0089739	18	56805417	G/A	0.05	0.3542	0.010438196	2.69958E-07	0.270060482	0.002560988	ENSSSCG00000025060	43038	MRPL32	mitochondrial ribosomal protein L32
ALGA0047963	8	63119376	G/A	0.425	0.8125	0.0107024395	2.76792E-07	0.002563961	3.53E-06	-	-	-	-
H3GA0007151	2	91619544	G/A	0.35	0.75	0.0107024395	2.76792E-07	0.004013913	9.07E-07	ENSSSCG00000020126	6744	-	-
INRA0029837	8	62838673	A/G	0.425	0.8125	0.0107024395	2.76792E-07	0.002563961	3.53E-06	-	-	-	-
ALGA0080004	14	101214975	G/A	0.2	0.5417	0.0111950443	2.89532E-07	0.000714719	1.08E-06	ENSSSCG00000021978	429052	RPL37A	60S ribosomal protein L37a-like
DRGA0014282	14	101161138	A/G	0.2	0.5417	0.0111950443	2.89532E-07	0.000714719	1.08E-06	ENSSSCG00000021978	375215	RPL37A	60S ribosomal protein L37a-like
INRA0045792	14	101193342	G/A	0.2	0.5417	0.0111950443	2.89532E-07	0.000714719	1.08E-06	ENSSSCG00000021978	407419	RPL37A	60S ribosomal protein L37a-like
MARC0090892	14	100862271	A/G	0.2	0.5417	0.0111950443	2.89532E-07	0.000714719	1.08E-06	ENSSSCG00000021978	76348	RPL37A	60S ribosomal protein L37a-like
ALGA0048011	8	66465699	G/A	0.275	0.6042	0.0114784661	2.96862E-07	0.000122302	9.36E-07	ENSSSCG00000024681	135717	TECRL	trans-2,3-enoyl-CoA reductase-like
M1GA0001077	1	109731334	A/G	0.475	0.08333	0.0117690411	3.04377E-07	0.005395158	2.42E-05	ENSSSCG00000004518	1476	SKA1	spindle and kinetochore associated complex subunit 1
ASGA0077154	17	50740211	C/A	0.375	0.7708	0.0118873204	3.07436E-07	0.000208756	1.04E-06	ENSSSCG00000027279	26177	PTPRT	protein tyrosine phosphatase, receptor type, T
H3GA0003173	1	181369985	G/A	0.425	0.1042	0.0119469047	3.08977E-07	0.040432421	0.001605131	ENSSSCG00000004938	37140	DENND4A	-
ASGA0091617	14	399610	G/A	0.325	0.6667	0.0121274362	3.13646E-07	0.00185481	8.84E-07	ENSSSCG00000023190	130590	SPIN1	-
ASGA0078111	17	64965732	A/G	0.1	0.4167	0.0122493115	3.16798E-07	0.395106988	0.000482093	ENSSSCG00000007504	3520	RNPC1	RNA binding motif protein 38
ALGA0020243	3	102571269	G/A	0.075	0.4583	0.012372424	3.19982E-07	0.231469085	0.000405941	ENSSSCG00000020378	51726	SNORA25	-
ASGA0047227	10	30012240	C/A	0.175	0.5833	0.0124967739	3.23198E-07	0.008982668	1.55E-06	ENSSSCG00000028111	9215	-	-
M1GA0025366	12	3331012	C/A	0.45	0.9375	0.0128131391	3.3138E-07	0.029387115	9.56E-06	ENSSSCG00000017164	6912	TIMP-2	TIMP metallopeptidase inhibitor 2
ASGA0071786	15	152434684	G/A	0.2	0.625	0.0129418969	3.3471E-07	0.051489424	6.44E-05	ENSSSCG00000016341	3252	TRAF3IP1	-
ALGA0121005	2	50293553	A/G	0.175	0.5	0.0130719693	3.38074E-07	0.012866838	3.16E-06	ENSSSCG00000028827	321	PARVA	parvin, alpha
ALGA0123380	16	27549968	A/C	0.1	0.375	0.0131375082	3.39769E-07	0.125159199	0.000714821	ENSSSCG00000016862	15754	PLCXD3	phosphatidylinositol-specific phospholipase C, X domain containing 3
ASGA0032322	7	33259737	A/G	0.45	0.04167	0.0134700744	3.4837E-07	0.002612594	4.52E-06	ENSSSCG00000001498	13422	BEND6	BEN domain containing 6
ALGA0013074	2	40796080	G/A	0.325	0.75	0.0136054441	3.51871E-07	0.131264501	5.66E-06	ENSSSCG00000019003	266692	U6	-
MARC0048886	16	11583165	A/G	0.125	0.4583	0.0138110699	3.57189E-07	0.28652645	0.000186432	ENSSSCG00000016800	277738	-	-
ASGA0002001	1	33651390	A/G	0.3	0.6458	0.0139498808	3.60779E-07	0.141728282	6.54E-06	ENSSSCG00000004175	52293	-	-
ASGA0028788	6	81775025	G/A	0.375	0.08333	0.0139498808	3.60779E-07	0.222490163	0.000301631	ENSSSCG00000003599	4529	-	-
H3GA0026947	9	34841808	A/G	0.125	0.5	0.0139498808	3.60779E-07	0.536217685	0.000619035	ENSSSCG00000014975	246917	CNTN5	contactin 5
ALGA0070492	13	64898919	G/A	0.125	0.4167	0.0140198276	3.62588E-07	0.525770381	0.000620742	ENSSSCG00000026687	291821	SNORA31	-
INRA0054509	17	61236574	G/A	0.2	0.5625	0.0140198276	3.62588E-07	0.045767949	9.74E-06	ENSSSCG00000007482	7280	TSHZ2	ENSSSCG00000007482
MARC0030292	4	143134967	A/G	0.45	0.0625	0.0143030174	3.69912E-07	0.006989796	0.000378481	ENSSSCG00000006941	36717	DDAH1	dimethylarginine dimethylaminohydrolase 1
DRGA0006607	6	29400244	A/G	0.25	0.6042	0.014519199	3.75503E-07	3.52E-05	2.12E-06	ENSSSCG00000002835	18751	TOX3	TOX high mobility group box family member 3
ASGA0091220	2	121468509	A/G	0.225	0.4167	0.0145919684	3.77385E-07	0.001299559	5.81E-06	ENSSSCG00000022048	95	REEP5	receptor accessory protein 5
INRA0051331	16	26037972	G/A	0.2	0.5625	0.014886758	3.85009E-07	0.112439847	2.78E-05	ENSSSCG00000020028	22766	U2	-
ALGA0105374	8	39467882	A/G	0.475	0.1458	0.0149613834	3.86939E-07	0.011068552	0.000621505	ENSSSCG00000008812	35215	ATP10D	probable phospholipid-transporting ATPase VD-like

ALGA0040040	7	32915748	A/G	0.45	0.04167	0.0151874635	3.92786E-07	0.00168886	2.56E-06	ENSSSCG00000001493	107220	-	primase, DNA, polypeptide 2 (58kDa)
ASGA0020060	4	75563714	G/A	0.475	0.75	0.0153401168	3.96734E-07	0.067707872	2.48E-05	ENSSSCG00000022092	208764	-	25-hydroxycholesterol 7-alpha-hydroxylase-like
ALGA0039343	7	21815831	A/G	0.475	0.0625	0.0154942782	4.00721E-07	0.013902793	0.001160722	ENSSSCG00000001106	2134	SLC17A4	solute carrier family 17, member 4
ALGA0098636	18	53021055	G/A	0.25	0.6042	0.016126622	4.17075E-07	0.078197468	2.04E-05	ENSSSCG00000016718	83488	NPY	neuropeptide Y
ALGA0014130	2	86940572	A/G	0.15	0.5417	0.0162074726	4.19166E-07	0.010447533	2.04E-06	ENSSSCG00000014087	23748	-	-
ALGA0038878	7	16185178	G/A	0.175	0.625	0.0162887098	4.21267E-07	0.009959497	9.77E-07	ENSSSCG00000024312	2187	ID4	inhibitor of DNA binding 4, dominant negative helix-loop-helix protein
ASGA0080381	18	57033078	A/C	0.425	0.8333	0.0163703337	4.23378E-07	0.010568078	3.33E-05	ENSSSCG00000016760	81492	C7orf25	UPF0415 protein C7orf25 homolog
MARC0006637	7	24632323	C/A	0.157 9	0.625	0.0163703337	4.23378E-07	0.02864117	1.98E-05	ENSSSCG00000030790	4979	TMP-CH242-74M17.6	-
ALGA0120391	2	47473825	G/A	0.175	0.5208	0.0166177575	4.29777E-07	0.002549256	2.02E-06	ENSSSCG00000013387	1007	CRSP2	ENSSSCG00000013387
ASGA0082267	2	47485301	A/G	0.175	0.5208	0.0166177575	4.29777E-07	0.002549256	2.02E-06	ENSSSCG00000013387	12483	CRSP2	ENSSSCG00000013387
DRGA0014328	14	104195215	C/A	0.225	0.5417	0.0167847559	4.34096E-07	0.000402069	2E-06	ENSSSCG00000010426	206600	PCDH15	-
DRGA0013313	13	194912413	A/G	0.425	0.125	0.0169534557	4.38459E-07	0.008077612	1.97E-05	ENSSSCG00000027488	375152	U6	-
H3GA0004945	1	301126002	A/C	0.025	0.4167	0.0170384436	4.40657E-07	0.115988907	0.011948812	ENSSSCG00000005606	99138	LMX1B	LIM homeobox transcription factor 1, beta
ALGA0111929	18	7558931	G/A	0.275	0.4792	0.0171238181	4.42865E-07	0.305392619	7.87E-05	ENSSSCG00000016472	3886	KEL	Kell blood group, metallo-endopeptidase
H3GA0021302	7	46433394	A/G	0.375	0.0625	0.0171238181	4.42865E-07	0.270463108	0.029124174	ENSSSCG00000025551	16122	RUNX2	runt-related transcription factor 2-like
ASGA0021413	4	110552282	A/G	0.25	0	0.0172096566	4.45085E-07	0.600920854	0.558565339	ENSSSCG00000029963	13170	U6	-
DIAS0002084	12	15315430	A/G	0.25	0.6875	0.0172096566	4.45085E-07	0.068990357	1.64E-06	ENSSSCG00000017294	4285	DCAF7	DDB1 and CUL4 associated factor 7
ASGA0019958	4	71104715	C/A	0.475	0.1667	0.0174697628	4.51812E-07	0.548152153	0.000301951	ENSSSCG00000006196	3965	SLCO5A1	solute carrier organic anion transporter family, member 5A1
ASGA0068966	15	25098826	C/G	0.5	0.1458	0.0176453451	4.56353E-07	0.001007709	9.55E-06	ENSSSCG00000015711	140232	DPP10	dipeptidyl-peptidase 10 (non-functional)
ASGA0080380	18	57050998	A/G	0.175	0.5208	0.0176453451	4.56353E-07	0.163091357	5.07E-05	ENSSSCG00000016760	99412	C7orf25	UPF0415 protein C7orf25 homolog
DRGA0010767	11	7020997	G/A	0.05	0.3958	0.0177337742	4.5864E-07	0.957016613	0.002205494	ENSSSCG00000009328	15060	-	SUMO-specific isopeptidase USPL1-like
MARC0114192	16	70319503	A/C	0.325	0.75	0.0177337742	4.5864E-07	0.019759091	2.26E-06	ENSSSCG00000017046	96139	-	early B-cell factor 1
ASGA0010854	2	91023501	A/G	0.475	0.125	0.0180920147	4.67905E-07	0.031456667	0.000126978	ENSSSCG00000014124	17474	ANKRD34B	ankyrin repeat domain 34B
ALGA0031066	5	20060561	A/G	0.15	0.4792	0.0181827252	4.70251E-07	0.212943185	0.000210449	ENSSSCG00000024221	10343	SMUG1	single-strand-selective monofunctional uracil-DNA glycosylase 1
ALGA0123968	2	19182541	A/G	0.1	0.3958	0.0184575244	4.77358E-07	0.257313291	0.001993705	ENSSSCG00000029521	67208	CD82	CD82 molecule
ALGA0048118	8	75129526	A/G	0.375	0.7083	0.0186430052	4.82155E-07	0.001265889	1.71E-06	ENSSSCG00000008966	89303	PARM1	prostate androgen-regulated mucin-like protein 1
DRGA0007207	7	17100569	A/T	0.225	0.5625	0.018736461	4.84572E-07	0.004161055	1.98E-06	ENSSSCG00000001080	280685	CDKAL1	CDK5 regulatory subunit associated protein 1-like 1
ALGA0014331	2	90404370	A/G	0.025	0.3542	0.0188303807	4.87001E-07	0.047803723	0.000419544	ENSSSCG00000014116	7668	-	metaxin-3-like
DRGA0010491	10	45903094	G/A	0.15	0.4375	0.0188303807	4.87001E-07	0.040857581	4.57E-05	ENSSSCG00000011022	26355	SVIL	supervillin-like
ASGA0010336	2	48424699	A/G	0.15	0.4792	0.0191149624	4.94361E-07	0.005134363	4.9E-06	ENSSSCG00000013393	184505	SPON1	spondin 1, extracellular matrix protein
MARC0064305	8	63062433	G/A	0.425	0.8125	0.0191149624	4.94361E-07	0.002579963	3.1E-06	-	-	-	-
ALGA0064341	12	2272704	G/A	0.15	0.5833	0.0193070551	4.99329E-07	0.00226895	2.86E-05	ENSSSCG00000017155	5196	EIF4A3	eukaryotic translation initiation factor 4A3
ALGA0007463	1	215320955	A/G	0.275	0.6042	0.0195011198	5.04348E-07	0.591836906	1.1E-05	ENSSSCG00000005109	394954	ER2	estrogen receptor 2 (ER beta)
H3GA0043639	15	3176420	G/A	0.325	0.08333	0.0195011198	5.04348E-07	0.034622201	0.001167155	ENSSSCG00000027211	54962	-	enhancer of polycomb homolog 2 (Drosophila)
MARC0040976	12	15896244	A/G	0.075	0.3125	0.0195988674	5.06876E-07	0.046874627	0.006356222	ENSSSCG00000017299	1437	MARCH10	membrane-associated ring finger (C3HC4) 10, E3 ubiquitin protein ligase
ASGA0031873	7	24149834	G/A	0.375	0.8333	0.019795832	5.1197E-07	0.004479364	2.02E-06	ENSSSCG00000001209	4261	ZSCAN12	zinc finger and SCAN domain containing 12
MARC0089377	2	40201108	A/G	0.4	0.8125	0.019795832	5.1197E-07	0.022225047	1.39E-05	ENSSSCG00000013343	110551	SLC17A6	solute carrier family 17 (vesicular glutamate transporter), member 6

ALGA0070475	13	64170665	A/C	0.125	0.4375	0.0199947686	5.17115E-07	0.753611386	0.000745361	ENSSSCG00000022851	208105	U6	-
ASGA0026945	5	102979194	G/A	0.05	0.4583	0.0199947686	5.17115E-07	0.099563325	0.00179264	ENSSSCG00000021240	488400	-	-
ALGA0084256	15	21648582	A/G	0.45	0.875	0.0201957158	5.22312E-07	0.009339597	3.21E-06	ENSSSCG00000024217	137186	-	uncharacterized LOC100513689
MARC0047493	15	21613488	A/G	0.45	0.875	0.0201957158	5.22312E-07	0.009339597	3.21E-06	ENSSSCG00000024217	102092	-	uncharacterized LOC100513689
ALGA0072005	13	142194326	A/G	0.475	0.1458	0.0211252851	5.46353E-07	0.015314325	0.000168467	ENSSSCG00000024109	53228	-	D-beta-hydroxybutyrate dehydrogenase, mitochondrial-like
ALGA0048032	8	67669311	C/A	0.475	0.1875	0.0218777641	5.65814E-07	0.013253459	3.2E-05	ENSSSCG00000022232	65148	-	-
MARC0000554	8	67026060	A/G	0.475	0.1875	0.0218777641	5.65814E-07	0.013253459	3.2E-05	ENSSSCG00000023488	33014	-	-
MARC0050311	8	67597907	A/G	0.475	0.1875	0.0218777641	5.65814E-07	0.013253459	3.2E-05	ENSSSCG00000022232	6137	-	-
MARC0084543	8	67568200	G/A	0.475	0.1875	0.0218777641	5.65814E-07	0.013253459	3.2E-05	ENSSSCG00000008919	27528	EPHA5	EPH receptor A5
ASGA0068640	15	14909574	G/A	0.15	0.5	0.0223197165	5.77244E-07	0.159492979	5.15E-05	ENSSSCG00000030701	1347	U6	-
ASGA0065507	14	106653236	A/G	0.25	0.5625	0.0229994648	5.94824E-07	0.003393499	1.08E-06	ENSSSCG00000010429	54434	PRKG1	protein kinase, cGMP-dependent, type I
DRGA0014387	14	106744337	A/G	0.25	0.5625	0.0229994648	5.94824E-07	0.003393499	1.08E-06	ENSSSCG00000010429	36667	PRKG1	protein kinase, cGMP-dependent, type I
ALGA0048014	8	66622716	A/G	0.475	0.1875	0.02381872	6.16012E-07	0.012522172	3.19E-05	ENSSSCG00000023486	120220	7SK	-
ASGA0031237	7	12019377	G/A	0.175	0.7083	0.02381872	6.16012E-07	0.002159154	1.72E-06	ENSSSCG00000001061	49323	JARID2	jumonji, AT rich interactive domain 2
ASGA0038870	8	66562625	A/C	0.475	0.1875	0.02381872	6.16012E-07	0.012522172	3.19E-05	ENSSSCG00000023486	180311	7SK	-
ALGA0073889	13	211884452	A/G	0.225	0.5	0.0240581012	6.22203E-07	0.658131288	0.007776926	ENSSSCG00000019614	117744	SNORA72	-
MARC0041991	2	50271279	G/A	0.175	0.5	0.0242998797	6.28456E-07	0.010891054	4.02E-06	ENSSSCG00000028827	21953	PARVA	parvin, alpha
MARC0111039	8	105991646	A/G	0.225	0.6042	0.0250399083	6.47595E-07	0.000735961	1.21E-06	ENSSSCG00000025142	18780	-	protocadherin Fat 4-like
H3GA0041580	14	101532615	A/G	0.225	0.5417	0.0252915853	6.54104E-07	0.000470839	1.83E-06	-	-	-	-
H3GA0007027	2	88090537	C/A	0.05	0.375	0.0256737987	6.63989E-07	0.058661578	0.000279274	ENSSSCG00000014099	46337	OTP	orthopedia homeobox
ASGA0038865	8	64409837	A/G	0.5	0.2292	0.0261924644	6.77403E-07	0.002519733	1.21E-05	ENSSSCG00000008915	234676	-	peptidyl-prolyl cis-trans isomerase C-like
ALGA0039868	7	31011932	A/G	0.2	0.5417	0.0263237355	6.80798E-07	0.065619149	5.31E-05	ENSSSCG00000001484	17870	-	tubulointerstitial nephritis antigen-like
H3GA0042360	14	134792938	G/A	0.475	0.1667	0.0263237355	6.80798E-07	0.009013198	0.000169848	ENSSSCG00000010638	4696	TCF7L2	transcription factor 7-like 2 (T-cell specific, HMG-box)
ALGA0016994	2	157286285	G/A	0.25	0.5208	0.0267215699	6.91087E-07	0.040244379	9E-06	ENSSSCG00000014432	3212	GRPEL2	GrpE-like 2, mitochondrial (E. coli)
H3GA0020102	7	16634919	G/A	0.35	0.7917	0.026855509	6.94551E-07	0.007152777	3.16E-06	ENSSSCG00000001078	95737	MBOAT1	membrane bound O-acyltransferase domain containing 1
H3GA0020846	7	36202231	A/G	0.4	0.08333	0.026990144	6.98033E-07	0.00776225	0.000157004	ENSSSCG00000001539	13029	PPARD	peroxisome proliferator-activated receptor delta
ALGA0115976	15	2798685	C/A	0.3	0.7708	0.027261386	7.05048E-07	0.010052076	1.63E-06	ENSSSCG00000022919	68326	KIF5C	kinesin family member 5C
ASGA0059307	13	179392818	A/G	0.45	0.7917	0.0275353732	7.12134E-07	0.139667876	1.32E-05	ENSSSCG00000011996	60395	POU1F1	POU class 1 homeobox 1
ALGA0088976	16	9124611	A/C	0.125	0.4792	0.0278121058	7.19291E-07	0.065159516	9.96E-05	ENSSSCG00000026842	246936	CDH18	cadherin 18, type 2
ALGA0047992	8	65489064	A/G	0.025	0.25	0.0280916223	7.2652E-07	0.48473301	0.1244458	-	-	-	-
ALGA0047995	8	65627195	A/G	0.025	0.25	0.0280916223	7.2652E-07	0.48473301	0.1244458	-	-	-	-
MARC0058200	8	65571761	G/A	0.025	0.25	0.0280916223	7.2652E-07	0.48473301	0.1244458	-	-	-	-
ALGA0102491	8	69215764	A/G	0.475	0.1875	0.0282324439	7.30162E-07	0.012581935	3.42E-05	ENSSSCG00000020337	176863	SNORA31	-
ASGA0085207	8	69065980	A/G	0.475	0.1875	0.0282324439	7.30162E-07	0.012581935	3.42E-05	ENSSSCG00000020337	326647	SNORA31	-
H3GA0024938	8	69622289	A/G	0.475	0.1875	0.0282324439	7.30162E-07	0.012581935	3.42E-05	ENSSSCG00000008921	36468	-	centromere protein C
MARC0020237	8	69070421	A/G	0.475	0.1875	0.0282324439	7.30162E-07	0.012581935	3.42E-05	ENSSSCG00000020337	322206	SNORA31	-
ALGA0027872	4	117478732	G/A	0.45	0.8333	0.0283739615	7.33822E-07	0.235055465	1.31E-05	ENSSSCG00000006768	27683	-	leucine-rich repeats and immunoglobulin-like domains 2
H3GA0022595	7	100997092	A/G	0.425	0.6042	0.028516175	7.375E-07	0.851023821	4.83E-05	ENSSSCG00000002330	135011	PCNX	pecanex homolog (Drosophila)

INRA0045898	14	103915659	A/G	0.225	0.5417	0.0289471462	7.48646E-07	0.000511206	1.9E-06	ENSSSCG00000010426	486156	PCDH15	-
ASGA0099838	1	106506645	A/G	0.5	0.1458	0.0295318922	7.63769E-07	0.041809521	7.83E-05	ENSSSCG00000004503	2482	-	-
DIAS0001336	7	27128290	A/G	0.5	0.3125	0.0295318922	7.63769E-07	0.317177012	2.45E-05	ENSSSCG00000001384	1670	VARS2	valyl-tRNA synthetase
MARC0027972	8	148158316	A/G	0.025	0.25	0.0296799443	7.67598E-07	0.475944478	0.12391292	ENSSSCG00000009262	43681	CSN1S1	casein alpha s1
ASGA0031822	7	22852252	C/A	0.45	0.1429	0.0301285085	7.79199E-07	0.674317233	0.004998977	ENSSSCG00000029787	6993	-	-
ASGA0031451	7	16354696	G/A	0.15	0.6667	0.030431302	7.8703E-07	0.019410715	2.68E-06	ENSSSCG00000001078	74748	MBOAT1	membrane bound O-acyltransferase domain containing 1
H3GA0024937	8	69466912	G/A	0.425	0.7917	0.0308911954	7.98924E-07	0.001663377	4.06E-06	ENSSSCG00000020337	74153	SNORA31	-
ALGA0122057	16	64790441	A/G	0.3	0.625	0.0310460527	8.02929E-07	0.00320227	4.98E-06	ENSSSCG00000019195	28475	U6	-
DRGA0005129	4	123593419	G/A	0.425	0.625	0.0310460527	8.02929E-07	0.238856454	5.34E-06	ENSSSCG00000006853	125632	-	-
DRGA0016148	16	41222726	A/G	0.4	0.6667	0.031515226	8.15063E-07	0.019277721	1.9E-06	ENSSSCG00000018875	113021	ssc-mir-582	-
MARC0038923	17	14222845	A/C	0.225	0.5417	0.031515226	8.15063E-07	0.009010511	3.54E-06	ENSSSCG00000026893	6096	-	arylamine N-acetyltransferase 1-like
MARC0008576	8	81626170	G/A	1	0.7083	0.0316732152	8.19149E-07	0.212396615	0.01024101	ENSSSCG00000025372	10036	-	glutamyl-tRNA(Gln) amidotransferase subunit B, mitochondrial-like
ALGA0013618	2	48125145	G/A	0.15	0.5	0.0318319778	8.23255E-07	0.004257672	3.57E-06	ENSSSCG00000013393	35989	SPON1	spondin 1, extracellular matrix protein
ALGA0049963	8	141941102	A/G	0.5	0.1458	0.0326378159	8.44096E-07	0.029616086	5.16E-05	ENSSSCG00000009228	56473	MAPK10	mitogen-activated protein kinase 10
H3GA0007135	2	91335286	G/A	0.35	0.7292	0.0326378159	8.44096E-07	0.018882133	3.84E-06	ENSSSCG00000014126	53299	MSH3	-
ASGA0056435	13	19066969	A/G	0.2	0	0.0328014118	8.48327E-07	0.309277616	0.122189263	ENSSSCG00000011224	1672	GADL1	-
MARC0051886	15	2964759	A/G	0.35	0.7917	0.0329658196	8.52579E-07	0.021190806	9.75E-06	ENSSSCG00000027211	15661	-	enhancer of polycomb homolog 2 (Drosophila)
ASGA0039984	8	135860942	G/A	0.35	0.1667	0.0332971486	8.61148E-07	0.000155412	8.62E-05	ENSSSCG00000009200	172927	-	-
MARC0003869	2	50547294	G/A	0.05	0.375	0.0332971486	8.61148E-07	0.092591779	0.000400304	ENSSSCG00000020731	7841	-	-
ALGA0077258	14	46536533	G/A	0.325	0.7292	0.0338003479	8.74162E-07	0.01883273	1.7E-06	ENSSSCG00000009957	24147	MYO18B	myosin XVIIIB
ASGA0058857	13	142153651	A/G	0.45	0.1042	0.0343111644	8.87373E-07	0.04112795	0.001934591	ENSSSCG00000011832	13717	-	D-beta-hydroxybutyrate dehydrogenase, mitochondrial-like
ASGA0036418	7	120617327	A/G	0.425	0.7292	0.0344831508	8.91821E-07	0.012872595	3.6E-06	ENSSSCG00000002450	2829	SLC24A4	solute carrier family 24 (sodium/potassium/calcium exchanger), member 4
ASGA0065470	14	104699566	G/A	0.225	0.5417	0.0344831508	8.91821E-07	0.000472711	2.16E-06	ENSSSCG00000010426	154602	PCDH15	-
DRGA0014325	14	104117355	A/G	0.225	0.5417	0.0344831508	8.91821E-07	0.000472711	2.16E-06	ENSSSCG00000010426	284460	PCDH15	-
ALGA0079389	14	91290555	C/A	0.425	0.1042	0.0353561131	9.14398E-07	0.803748215	0.008678465	ENSSSCG00000020142	28174	SNORA31	-
ASGA0102346	8	142146988	G/A	0.35	0.7917	0.0358904385	9.28217E-07	0.004694382	1.19E-05	ENSSSCG00000009229	1097	ARHGAP24	Rho GTPase activating protein 24
H3GA0001977	1	80667615	A/G	0.425	0.2083	0.0358904385	9.28217E-07	0.139293759	0.139079009	ENSSSCG00000004366	13499	BVES	blood vessel epicardial substance
ALGA0048051	8	68852606	A/G	0.425	0.7917	0.0364328452	9.42245E-07	0.001130341	4.41E-06	ENSSSCG00000021957	409154	5S_rRNA	-
ALGA0045159	7	123301941	A/G	0.325	0.7708	0.0367990122	9.51715E-07	0.001030571	6.05E-06	ENSSSCG00000026899	20264	-	-
M1GA0009653	7	17426267	A/G	0.275	0.6667	0.0373551839	9.66099E-07	0.007232019	1.99E-06	ENSSSCG00000001080	54760	CDKAL1	CDK5 regulatory subunit associated protein 1-like 1
ALGA0113830	13	210457466	A/G	0.4	0.7917	0.0383007996	9.90555E-07	0.127716885	4.28E-06	ENSSSCG00000027179	17484	SIM2	single-minded homolog 2 (Drosophila)
ALGA0070637	13	70431811	A/G	0.325	0.08333	0.038492815	9.95521E-07	0.109313928	0.00078436	ENSSSCG00000023891	158206	GRM7	glutamate receptor, metabotropic 7
ALGA0093127	17	12496334	A/G	0.325	0.5833	0.038492815	9.95521E-07	0.009191325	2.05E-06	ENSSSCG00000007023	9181	KAT6A	-
ASGA0102687	17	52043004	G/A	0.475	0.1667	0.0386857197	1.00051E-06	0.005654513	8.01E-06	ENSSSCG00000007368	850	-	TOX high mobility group box family member 2
ASGA0053324	12	14114528	A/G	0.15	0.3958	0.0390746996	1.01057E-06	0.054732559	0.000958323	ENSSSCG00000017274	9831	PITPNC1	phosphatidylinositol transfer protein, cytoplasmic 1
ALGA0120187	5	67276981	G/A	0.3	0.08333	0.0392703496	1.01563E-06	0.184899355	0.000706989	ENSSSCG00000000713	858	-	anoctamin 2
ALGA0039607	7	26679310	A/G	0.425	0.1042	0.0396651294	1.02584E-06	0.351950693	0.001138691	ENSSSCG00000001341	4948	SLA-11	-
ALGA0037987	7	1571651	G/A	0.225	0.5625	0.0400637759	1.03615E-06	0.473784787	0.000207694	ENSSSCG00000028777	102720	MYLK4	myosin light chain kinase family, member 4

ASGA0060118	13	212649283	G/A	0.45	0.7292	0.0400637759	1.03615E-06	0.000531414	2.5E-06	ENSSSCG00000030314	253974	PSMG1	proteasome (prosome, macropain) assembly chaperone 1
MARC0078025	17	13901451	A/C	0.3	0.625	0.0402644524	1.04134E-06	0.397361911	1.2E-05	ENSSSCG00000027808	11449	-	-
ALGA0020630	3	112743377	G/A	0.225	0.5625	0.0410779851	1.06238E-06	0.099875319	4.27E-05	ENSSSCG00000019040	34076	U2	-
ALGA0116461	6	13929389	G/A	0.375	0.8125	0.0412836882	1.0677E-06	0.779804476	0.000111044	ENSSSCG00000002730	3878	CALB2	calbindin 2
ASGA0097399	13	210534102	G/C	0.5	0.1042	0.0412836882	1.0677E-06	0.057370531	2.7E-05	ENSSSCG00000012059	52402	HLCS	holocarboxylase synthetase (biotin-(proprionyl-CoA-carboxylase (ATP-hydrolysing)) ligase)
DIAS0000346	4	72722972	G/A	0.4	0.7292	0.0412836882	1.0677E-06	0.004849368	3.57E-06	ENSSSCG00000006199	23862	PREX2	phosphatidylinositol-3,4,5-trisphosphate-dependent Rac exchange factor 2
MARC0019610	13	210504370	G/A	0.5	0.1042	0.0412836882	1.0677E-06	0.057370531	2.7E-05	ENSSSCG00000012059	22670	HLCS	holocarboxylase synthetase (biotin-(proprionyl-CoA-carboxylase (ATP-hydrolysing)) ligase)
ALGA0103392	8	69146919	A/C	0.475	0.1875	0.0427545428	1.10574E-06	0.011951629	3.91E-05	ENSSSCG00000020337	245708	SNORA31	-
ASGA0064790	14	88963503	A/G	0.375	0.1042	0.0427545428	1.10574E-06	0.452576071	0.049855156	ENSSSCG00000010341	7597	TSPAN14	tetraspanin 14
MARC0095739	8	69146481	A/G	0.475	0.1875	0.0427545428	1.10574E-06	0.011951629	3.91E-05	ENSSSCG00000020337	246146	SNORA31	-
MARC0104889	8	71042714	A/G	0.475	0.1875	0.0427545428	1.10574E-06	0.011951629	3.91E-05	ENSSSCG00000027194	21578	SULT1B1	sulfotransferase family cytosolic 1B member 1-like
ALGA0122094	8	89927504	A/G	0.375	0.7083	0.0434006517	1.12245E-06	0.026644117	2.73E-06	ENSSSCG00000009049	129483	USP38	ubiquitin carboxyl-terminal hydrolase 38-like
MARC0054558	8	89905287	G/A	0.375	0.7083	0.0434006517	1.12245E-06	0.026644117	2.73E-06	ENSSSCG00000009049	107266	USP38	ubiquitin carboxyl-terminal hydrolase 38-like
ALGA0044331	7	111031594	G/A	0.475	0.25	0.0438368042	1.13373E-06	0.000777781	3.06E-06	ENSSSCG00000026660	359866	-	-
H3GA0046526	16	41348366	G/A	0.425	0.6667	0.0440564271	1.13941E-06	0.046784945	2.8E-06	ENSSSCG00000018875	12522	ssc-mir-582	-
ALGA0109927	13	196079847	G/A	0.325	0.08333	0.0442772099	1.14512E-06	0.105495869	0.026671526	ENSSSCG00000018744	101459	U6	-
H3GA0012886	4	74607888	G/A	0.25	0.4792	0.0442772099	1.14512E-06	0.081566176	1.38E-05	ENSSSCG00000006215	43732	CRH	ENSSSCG00000006215
H3GA0043950	15	22667341	A/G	0.2	0.5417	0.0447222556	1.15663E-06	0.018699175	1.18E-05	ENSSSCG00000015710	57915	ARP3	ARP3 actin-related protein 3 homolog (yeast)
M1GA0011801	8	11263172	A/G	0.425	0.75	0.0451715545	1.16825E-06	0.00053147	1.9E-06	ENSSSCG00000029227	27430	LDB2	LIM domain binding 2
ASGA0060510	14	3691354	A/C	0.15	0.5625	0.04562588	1.18E-06	0.094788015	4.25E-05	ENSSSCG00000009593	48033	ROR2	receptor tyrosine kinase-like orphan receptor 2
INRA0014602	4	73348718	A/G	0.4	0.125	0.0458543961	1.18591E-06	0.189061955	0.000591173	ENSSSCG00000006200	119272	CPA6	carboxypeptidase A6-like
H3GA0032070	11	66469251	G/A	0.45	0.25	0.0470151494	1.21593E-06	0.009040915	6.35E-05	ENSSSCG00000018617	140809	ssc-mir-17	microRNA mir-17
CASI0009035	14	101145595	A/G	0.225	0.5417	0.047251012	1.22203E-06	0.005325052	2.02E-06	ENSSSCG00000021978	359672	RPL37A	60S ribosomal protein L37a-like
ASGA0103427	15	32287086	G/A	0.175	0	0.0482052889	1.24671E-06	0.577918046	0.129416743	ENSSSCG00000015726	483923	CNTNAP5	contactin associated protein-like 5
ALGA0047982	8	64602171	A/G	0.45	0.1667	0.0489337563	1.26555E-06	0.024877613	0.000108736	ENSSSCG00000026129	55722	LPHN3	latrophilin-3-like
ALGA0070606	13	68184094	G/A	0.3	0.6458	0.0496734369	1.28468E-06	0.41982646	1.31E-05	ENSSSCG00000020679	30036	EDEM1	ER degradation-enhancing alpha-mannosidase-like protein 1-like
ASGA0065418	14	101934775	G/A	0.225	0.5625	0.0504243306	1.3041E-06	0.002757201	2.34E-06	-	-	-	-
DRGA0016071	16	33038762	A/C	0.4	0.625	0.0519597575	1.34381E-06	4.34E-05	6.19E-06	ENSSSCG00000016883	479760	ISL1	insulin gene enhancer protein ISL-1-like
H3GA0033606	12	14478438	G/A	0.125	0.375	0.0519597575	1.34381E-06	0.11556375	0.00228713	ENSSSCG00000025451	62211	-	-
DRGA0011634	12	16627750	A/G	0.375	0.8958	0.0522203663	1.35055E-06	0.048357352	1.43E-05	ENSSSCG00000017305	6593	EFCAB13	-
ALGA0105006	12	14344586	A/G	0.125	0.3958	0.0527450639	1.36412E-06	0.065669949	0.000605267	ENSSSCG00000017274	31652	PITPNC1	phosphatidylinositol transfer protein, cytoplasmic 1
ASGA0104369	18	4101623	A/G	0.375	0.02083	0.0527450639	1.36412E-06	0.218565353	0.002383487	ENSSSCG00000016424	58054	-	-
ASGA0073684	16	64584237	G/A	0.15	0.5208	0.0530095394	1.37096E-06	0.19204337	0.00033464	ENSSSCG00000029965	134514	snoU13	-
MARC0034873	1	176159527	A/G	0.35	0.08333	0.0551729021	1.42691E-06	0.041741466	0.01573502	ENSSSCG00000004896	40110	PHLPP1	PH domain and leucine rich repeat protein phosphatase 1
ALGA0013068	2	40735360	A/G	0.342	0.75	0.055449364	1.43406E-06	0.303834765	2.82E-05	ENSSSCG00000019003	205972	U6	-
H3GA0020136	7	17936702	G/A	0.25	0.6875	0.0562872562	1.45573E-06	0.001857152	2.13E-06	ENSSSCG00000001081	82279	SOX4	SRY (sex determining region Y)-box 4
ALGA0045559	7	127959806	A/G	0.5	0.1667	0.0568529398	1.47036E-06	0.24140916	0.000106348	ENSSSCG00000002508	13420	SETD3	SET domain containing 3

DRGA0014300	14	102728645	G/A	0.325	0.7083	0.0568529398	1.47036E-06	0.004088555	1.88E-06	ENSSSCG00000027984	163433	ZWINT	ZW10 interacting kinetochore protein
H3GA0021155	7	41720015	A/G	0.4	0.7708	0.0568529398	1.47036E-06	0.004567366	1.85E-06	ENSSSCG00000001612	13037	-	adenylate cyclase type 10-like
ALGA0097048	18	12846683	A/G	0.1	0.3333	0.0571382948	1.47774E-06	0.042048824	0.002272196	ENSSSCG00000028951	15813	-	-
M1GA0022455	17	64997324	G/A	0.125	0.4792	0.0571382948	1.47774E-06	0.115814906	3.21E-05	ENSSSCG00000007505	27139	CTCFL	CCCTC-binding factor (zinc finger protein)-like
MARC0095015	2	86276937	A/G	0.125	0.375	0.0574244232	1.48514E-06	0.014739347	0.000172363	ENSSSCG00000014083	19612	ANKDD1B	-
MARC0096726	2	86270250	G/A	0.125	0.375	0.0574244232	1.48514E-06	0.014739347	0.000172363	ENSSSCG00000014083	12925	ANKDD1B	-
DRGA0008861	8	136985884	T/A	0.175	0.6042	0.0582924749	1.50759E-06	0.108736867	7.54E-06	-	-	-	-
ASGA0047209	10	30502415	A/G	0.2	0.5417	0.0588782648	1.52274E-06	0.010522175	4.61E-06	ENSSSCG00000010941	50828	ERCC6L2	-
ALGA0066854	12	51261422	A/G	0.475	0.08333	0.0591732864	1.53037E-06	0.087330285	8.95E-05	ENSSSCG00000021855	6564	-	olfactory receptor 1E2-like
ALGA0084968	15	40821463	A/G	0.15	0.3542	0.0591732864	1.53037E-06	0.027451581	0.00030667	-	-	-	-
ASGA0057727	13	60268089	A/G	0.4	0.875	0.0591732864	1.53037E-06	0.12263401	6.46E-05	ENSSSCG00000024410	72798	-	-
MARC0015767	15	40833878	A/C	0.15	0.3542	0.0591732864	1.53037E-06	0.027451581	0.00030667	-	-	-	-
SIRI0000090	15	40857125	A/G	0.15	0.3542	0.0591732864	1.53037E-06	0.027451581	0.00030667	-	-	-	-
DIAS0001261	1	120580759	G/A	0.25	0.6042	0.0597679695	1.54575E-06	9.37E-05	4.91E-06	ENSSSCG00000004571	3655	-	-
H3GA0002664	1	132709605	A/G	0.175	0.6042	0.0603688391	1.56129E-06	0.024408119	3.5E-06	ENSSSCG00000004623	20045	BCL2L10	BCL2-like 10 (apoptosis facilitator)
MARC0025849	1	132779670	A/G	0.175	0.6042	0.0603688391	1.56129E-06	0.024408119	3.5E-06	ENSSSCG00000004624	35788	MAPK6	mitogen-activated protein kinase 6
ASGA0093262	14	100971757	A/C	0.2	0.5	0.0609755087	1.57698E-06	0.001255483	3.44E-06	ENSSSCG00000021978	185834	RPL37A	60S ribosomal protein L37a-like
ASGA0039614	8	117902625	A/G	0.35	0.04167	0.0631473779	1.63315E-06	0.067739522	0.000354887	ENSSSCG00000009126	88811	CALM3	calmodulin-like
H3GA0020057	7	14524296	A/G	0.275	0.5417	0.0631473779	1.63315E-06	0.003208961	5.69E-06	ENSSSCG00000001072	34462	-	-
ALGA0004067	1	75540483	G/A	0.375	0.125	0.0641016548	1.65783E-06	0.286175314	0.178650146	ENSSSCG00000004356	23224	MCHR2	melanin-concentrating hormone receptor 2
H3GA0055422	12	13901807	A/T	0.175	0.4167	0.0641016548	1.65783E-06	0.010051248	0.000278157	ENSSSCG00000017272	38932	HELZ	helicase with zinc finger
MARC0113018	12	14013189	A/G	0.175	0.4167	0.0641016548	1.65783E-06	0.010051248	0.000278157	ENSSSCG00000019207	3833	U6	-
ALGA0025556	4	72955795	G/A	0.4	0.7292	0.0644229692	1.66614E-06	0.002740822	6.91E-06	ENSSSCG00000006199	34510	PREX2	phosphatidylinositol-3,4,5-trisphosphate-dependent Rac exchange factor 2
ALGA0122772	13	196698222	A/C	0.45	0.125	0.0644229692	1.66614E-06	0.007255856	0.000128159	-	-	-	-
ALGA0013636	2	48332744	C/G	0.175	0.5	0.0650706247	1.68289E-06	0.003299648	5.05E-06	ENSSSCG00000013393	243588	SPON1	spondin 1, extracellular matrix protein
ASGA0010781	2	89499964	A/G	0.175	0.5	0.0653965791	1.69132E-06	0.001489698	8.93E-06	ENSSSCG00000014110	24172	DMGDH	-
ALGA0080038	14	102136878	C/A	0.225	0.5625	0.0660539011	1.70832E-06	0.002262321	2.74E-06	ENSSSCG00000027984	423883	ZWINT	ZW10 interacting kinetochore protein
ALGA0080040	14	102310696	A/C	0.225	0.5625	0.0660539011	1.70832E-06	0.002262321	2.74E-06	ENSSSCG00000027984	250065	ZWINT	ZW10 interacting kinetochore protein
ASGA0101253	14	102244897	A/G	0.225	0.5625	0.0660539011	1.70832E-06	0.002262321	2.74E-06	ENSSSCG00000027984	315864	ZWINT	ZW10 interacting kinetochore protein
INRA0045843	14	102798434	A/G	0.225	0.5625	0.0660539011	1.70832E-06	0.002262321	2.74E-06	ENSSSCG00000027984	233222	ZWINT	ZW10 interacting kinetochore protein
ALGA0062775	11	67199186	G/A	0.475	0.7917	0.0663848821	1.71688E-06	0.4249429	0.000101834	-	-	-	-
ALGA0038863	7	16062680	A/G	0.25	0.6458	0.0670522572	1.73414E-06	0.288767954	2.02E-05	ENSSSCG00000024312	115409	ID4	inhibitor of DNA binding 4, dominant negative helix-loop-helix protein
ALGA0064392	12	4992763	A/C	0.425	0.1458	0.0680656931	1.76035E-06	0.042114171	9.14E-05	ENSSSCG00000029901	6052	UBE2O	ubiquitin-conjugating enzyme E2O
ALGA0097073	18	13610548	G/A	0.1	0.3333	0.0690942087	1.78695E-06	0.040322161	0.002286049	ENSSSCG00000025324	248490	ssc-mir-490-1	microRNA mir-490
MARC0078052	18	13598735	G/A	0.1	0.3333	0.0690942087	1.78695E-06	0.040322161	0.002286049	ENSSSCG00000025324	236677	ssc-mir-490-1	microRNA mir-490
ALGA0108082	18	57514307	A/G	0.45	0.7917	0.0701385774	1.81396E-06	0.000947689	2.11E-05	ENSSSCG00000016762	77623	-	-
H3GA0004899	1	300929374	C/A	0.025	0.25	0.0704900513	1.82305E-06	0.744436918	0.71684143	ENSSSCG00000027126	58293	MVB12B	multivesicular body subunit 12B
ALGA0084119	15	19392860	C/A	0.25	0	0.0708434585	1.83219E-06	0.379021267	0.007305295	ENSSSCG00000015696	9552	MAP3K19	mitogen-activated protein kinase kinase kinase 19
ALGA0111719	1	300568479	A/G	0.025	0.25	0.0708434585	1.83219E-06	0.744436918	0.71684143	ENSSSCG00000005604	134776	PBX3	pre-B-cell leukemia homeobox 3

ASGA0068760	15	19345084	A/G	0.25	0	0.0708434585	1.83219E-06	0.379021267	0.007305295	ENSSSCG00000015696	38224	MAP3K19	mitogen-activated protein kinase kinase kinase 19
MARC0058372	7	70901866	A/G	0.2	0.4375	0.0708434585	1.83219E-06	0.722006083	0.010391136	ENSSSCG00000001964	69065	-	neuronal PAS domain protein 3
MARC0087013	15	19187327	G/A	0.25	0	0.0708434585	1.83219E-06	0.379021267	0.007305295	ENSSSCG00000015694	4982	ZRANB3	zinc finger, RAN-binding domain containing 3
DRGA0007180	7	16157280	G/A	0.3	0.75	0.0719141201	1.85988E-06	0.000868584	2.61E-06	ENSSSCG00000024312	20809	ID4	inhibitor of DNA binding 4, dominant negative helix-loop-helix protein
DRGA0007802	7	76463729	A/C	0.125	0.4565	0.0722744872	1.8692E-06	0.008520302	4.8E-05	-	-	-	-
MARC0112120	18	55230564	A/G	0.45	0.8125	0.0722744872	1.8692E-06	0.085525147	8.3E-06	ENSSSCG00000016734	3314	CCM2	cerebral cavernous malformation 2
ASGA0068934	15	22821739	G/A	0.025	0.25	0.0733668017	1.89745E-06	0.933025997	0.208118991	ENSSSCG00000015710	33681	ARP3	ARP3 actin-related protein 3 homolog (yeast)
MARC0032054	8	82173582	A/G	1	0.7292	0.0744757426	1.92613E-06	0.386030478	0.039405858	ENSSSCG00000020717	19660	FAM160A1	protein FAM160A1-like
ALGA0047974	8	63475663	A/C	0.45	0.8125	0.0748488695	1.93578E-06	0.003545811	1.19E-05	-	-	-	-
H3GA0036617	13	65922076	C/A	0.475	0.08333	0.0763610967	1.97489E-06	0.072090259	0.00022832	ENSSSCG00000011527	13944	CNTN4	contactin 4
ALGA0110225	1	300646084	A/G	0.5	0.875	0.0810829887	2.09701E-06	0.012091428	2.19E-05	ENSSSCG00000005605	111467	-	-
MARC0069976	1	300686863	A/G	0.5	0.875	0.0810829887	2.09701E-06	0.012091428	2.19E-05	ENSSSCG00000005605	70688	-	-
ALGA0097126	18	14440945	G/A	0.1	0.2917	0.0823083142	2.1287E-06	0.090738198	0.013277353	ENSSSCG00000030386	14079	-	nucleoporin 205kDa
ALGA0124172	18	54644105	G/A	0.15	0.4792	0.0823083142	2.1287E-06	0.190985685	0.00196795	ENSSSCG00000016729	207399	IGFBP3	insulin-like growth factor binding protein 3
MARC0042080	8	81902637	G/A	1	0.6875	0.0823083142	2.1287E-06	0.250621873	0.008721222	ENSSSCG00000009015	1829	-	PET112 homolog (yeast)
ASGA0096934	14	107513215	G/A	0.4	0.7917	0.0831353799	2.15009E-06	0.018113666	4.26E-06	ENSSSCG00000029076	13610	-	-
ASGA0031526	7	17868173	A/G	0.25	0.6667	0.0839709522	2.1717E-06	0.002757401	3.03E-06	ENSSSCG00000001081	13750	SOX4	SRY (sex determining region Y)-box 4
ALGA0033883	5	103240594	A/G	0.05	0.3125	0.0843920249	2.18259E-06	0.434078466	0.073178094	ENSSSCG00000021240	227000	-	-
MARC0026696	6	26404310	G/A	0.375	0.7083	0.0843920249	2.18259E-06	0.001726487	2.92E-06	ENSSSCG00000024305	7486	MT3	-
MARC0082315	9	141625060	G/A	0.25	0.5208	0.0843920249	2.18259E-06	0.238918139	1.9E-05	ENSSSCG00000015581	72144	CENPF	centromere protein F, 350/400kDa
MARC0084603	14	134848431	G/A	0.45	0.7292	0.0865283214	2.23784E-06	0.0022638	9.59E-06	ENSSSCG00000010638	60189	TCF7L2	transcription factor 7-like 2 (T-cell specific, HMG-box)
M1GA0018476	14	30526642	G/A	0.125	0.4167	0.0873979198	2.26033E-06	0.236334562	0.000889601	ENSSSCG00000020756	7739	FAM101A	family with sequence similarity 101, member A
MARC0006578	13	208119290	A/G	0.5	0.125	0.0878360056	2.27166E-06	0.360173369	0.003728221	ENSSSCG00000025700	2383	U6	-
ALGA0047962	8	63000451	A/G	0.45	0.8125	0.0891634093	2.30599E-06	0.004791678	1.37E-05	-	-	-	-
DRGA0013292	13	192974850	A/G	0.475	0.1667	0.0891634093	2.30599E-06	0.014666031	1.98E-05	ENSSSCG00000012018	3055	CHODL	chondrolectin
ALGA0026100	4	84090680	A/G	0.475	0.2083	0.0896103883	2.31755E-06	0.883213943	0.008572447	ENSSSCG00000030578	51139	MRPL15	mitochondrial ribosomal protein L15
ALGA0039010	7	17812401	A/G	0.25	0.6667	0.0896103883	2.31755E-06	0.00288187	3.09E-06	ENSSSCG00000001081	40589	SOX4	SRY (sex determining region Y)-box 4
DRGA0007226	7	17837943	A/G	0.25	0.6667	0.0896103883	2.31755E-06	0.00288187	3.09E-06	ENSSSCG00000001081	15047	SOX4	SRY (sex determining region Y)-box 4
ASGA0037550	8	5065847	G/A	0.375	0.0625	0.0900596872	2.32917E-06	0.312650412	0.001140575	ENSSSCG00000022562	90669	STK32B	serine/threonine kinase 32B
ALGA0038408	7	8438208	G/A	0.2	0.4792	0.0905109194	2.34084E-06	0.447457931	0.001104084	ENSSSCG00000001046	123303	NEDD9	enhancer of filamentation 1-like
ASGA0082882	10	32207568	G/A	0.5	0.2917	0.0909648583	2.35258E-06	6.1E-05	2.97E-05	ENSSSCG00000026279	178757	DAPK1	death-associated protein kinase 1-like
H3GA0043933	15	22454707	A/G	0.375	0.7292	0.0909648583	2.35258E-06	0.000144033	6.57E-06	ENSSSCG00000015707	41471	GPR39	G protein-coupled receptor 39
MARC0075909	1	176700629	G/A	0.4	0.1042	0.0914207304	2.36437E-06	0.042579042	0.000866343	ENSSSCG00000004899	28547	KIAA1468	KIAA1468 ortholog
ALGA0046580	8	15696295	A/G	0.2	0.3958	0.0918789225	2.37622E-06	0.077271843	7.25E-05	ENSSSCG00000023934	87137	KCNIP4	Kv channel interacting protein 4
ALGA0056199	10	1570425	G/A	0.425	0.5833	0.0918789225	2.37622E-06	0.057575654	1.29E-05	-	-	-	-
ASGA0003886	1	104685944	G/A	0.425	0.1042	0.0951516128	2.46086E-06	0.036808306	0.00010966	ENSSSCG00000004490	178943	SETBP1	SET binding protein 1
MARC0032253	8	93360395	A/C	0.325	0.8333	0.0956283645	2.47319E-06	0.001968671	1.08E-05	ENSSSCG00000009062	24591	OSAP	ovary-specific acidic protein
ALGA0015051	2	117422430	A/C	0.125	0.4792	0.0970736996	2.51057E-06	0.036904355	2.09E-05	ENSSSCG00000029552	2083	-	-
H3GA0024861	8	50329649	G/A	0.15	0.4792	0.0970736996	2.51057E-06	0.081632739	6E-06	ENSSSCG00000021470	24299	-	-

ALGA0046922	8	22180534	A/G	0.225	0.02083	0.0975605046	2.52316E-06	0.381092058	0.701182055	-	-	-	-
ALGA0116876	6	17189607	C/A	0.375	0.6875	0.0975605046	2.52316E-06	0.001038908	3.41E-06	ENSSSCG00000002820	6115	RSPRY1	ring finger and SPRY domain containing 1
DRGA0007201	7	16929997	C/A	0.2	0.413	0.0985406877	2.54851E-06	0.070214076	0.00057691	ENSSSCG00000001080	110113	CDKAL1	CDK5 regulatory subunit associated protein 1-like 1
ALGA0044485	7	114553936	G/A	0.375	0.6458	0.0995313106	2.57413E-06	0.000143777	6.06E-06	ENSSSCG00000002418	194749	FLRT2	fibronectin leucine rich transmembrane protein 2
MARC0083540	6	24135910	A/G	0.3	0.5208	0.0995313106	2.57413E-06	0.024047876	4.57E-06	ENSSSCG00000002795	320780	CDH11	cadherin 11, type 2, OB-cadherin (osteoblast)

Table S4. Haploblocks including single nucleotide polymorphisms (SNPs) with $P_{Bonferroni}<0.10$.

Total no. of SNPs	No. of SNPs with $P_{Bonferroni}<0.10$	Haploblock size	Haploblock coordinates	SNPs
4	1	171699	1:10295470-10467169	ALGA0000778, MARC0078284, H3GA0000686
6	4	204504	1:104481440-104685944	H3GA0002300, ALGA0005010, H3GA0002301, ALGA0005012, ASGA0003886
6	2	780617	1:106360939-107141556	ASGA0106270, ASGA0099838, ASGA0102231, ALGA0005150, ASGA0003986
3	1	274406	1:120580759-120855165	DIAS0001261, MARC0094747
5	2	473903	1:121782234-122256137	H3GA0002540, ASGA0004239, MARC0035967, ASGA0004240
4	2	122072	1:132657598-132779670	ASGA0004378, H3GA0002664, MARC0025849
3	1	187401	1:138070556-138257957	ASGA0004417, ALGA0005834
6	2	580236	1:147390123-147970359	MARC0020010, MARC0056240, MARC0056241, ALGA0006042, ALGA0006048
4	1	103613	1:176128278-176231891	ALGA0006582, MARC0034873, INRA0004895
3	1	164343	1:181369929-181534272	H3GA0003173, ALGA0006736
3	1	20930	1:19798005-19818935	ALGA0001463, M1GA0000832
3	1	28745	1:215320955-215349700	ALGA0007463, MARC0087576
3	1	22650	1:33651390-33674040	ASGA0002001, ASGA0002002
4	1	68614	1:48539442-48608056	ASGA0002521, H3GA0001561, H3GA0001563
5	4	190966	1:74096746-74287712	INRA0002726, ASGA0003244, MARC0075306, ALGA0003981
6	4	429513	1:74410201-74839714	ALGA0003995, ALGA0004000, ALGA0004002, ALGA0004005, ALGA0119247
4	1	129927	1:75472274-75602201	ALGA0004061, ALGA0004067, ASGA0003288

4	1	154120	1:80634441-80788561	H3GA0001976, H3GA0001977, ALGA0004223
4	1	279922	2:121283845-121563767	MARC0097037, ASGA0091220, DRGA0003393
4	1	280633	2:157110989-157391622	MARC0039166, ALGA0016994, MARC0006554
5	2	88411	2:18717421-18805832	ALGA0111487, ALGA0108373, ASGA0096260, ASGA0009403
3	1	218984	2:39982124-40201108	ASGA0083049, MARC0089377
6	3	375062	2:40545801-40920863	MARC0036243, ALGA0013068, MARC0064216, ALGA0013074, MARC0018628
4	2	383857	2:41614382-41998239	ASGA0009972, MARC0064720, ALGA0013197
10	8	1007135	2:47117943-48125078	ALGA0013585, MARC0077792, ALGA0120391, ASGA0082267, ALGA0013592, ALGA0013597, ALGA0013598, ALGA0013605, ALGA0013618
4	2	247159	2:48332744-48579903	ALGA0013636, ASGA0010336, H3GA0006743
5	2	147715	2:48657856-48805571	ALGA0013642, ALGA0013646, H3GA0006751, ALGA0013650
6	2	721327	2:49336890-50058217	H3GA0006754, ALGA0109366, MARC0104903, ASGA0097814, ASGA0010359
10	9	494487	2:50226694-50721181	MARC0095075, ALGA0119911, MARC0041991, ALGA0121005, ASGA0095968, ASGA0094162, MARC0003869, H3GA0006772, ASGA0010360
6	4	373257	2:51003535-51376792	M1GA0024860, MARC0005311, ALGA0013695, MARC0079844, ALGA0013690
3	1	37342	2:8162976-8200318	ALGA0103374, MARC0088806
3	1	144182	2:83953431-84097613	ALGA0014045, INRA0008879
4	2	320611	2:85956326-86276937	H3GA0006971, MARC0096726, MARC0095015
3	1	287457	2:86940572-87228029	ALGA0014130, H3GA0006992
4	2	583410	2:87507127-88090537	M1GA0002981, ASGA0010692, H3GA0007027
7	5	759915	2:88913171-89673086	MARC0017959, ASGA0010757, MARC0035674, ALGA0108178, ASGA0010781, M1GA0003004
3	1	351065	2:90053305-90404370	ALGA0014302, ALGA0014331
4	2	39470	2:90444137-90483607	ALGA0014325, ALGA0104357, H3GA0007101
3	1	53437	2:90892618-90946055	ALGA0014351, DIAS0004400
4	2	52051	2:90971450-91023501	H3GA0007129, DRGA0003111, ASGA0010854
3	2	18079	2:91619544-91637623	H3GA0007151, ASGA0010866
4	1	378246	3:112674884-113053130	ASGA0015742, ALGA0020630, M1GA0004568
4	1	544114	3:120816964-121361078	ASGA0093257, MARC0043998, DRGA0004220

5	2	426387	4:101720321-102146708	M1GA0006114, ALGA0026861, ASGA0085258, ALGA0026875
3	1	21451	4:117457281-117478732	ALGA0027868, ALGA0027872
3	1	161941	4:123593410-123755351	DRGA0005129, ASGA0022327
3	1	214162	4:142920805-143134967	ALGA0029749, MARC0030292
4	2	366546	4:72722997-73089543	DIAS0000346, ALGA0025556, ALGA0025563
4	1	45645	4:73326039-73371684	ALGA0025577, INRA0014602, ASGA0020008
4	1	336345	4:74357690-74694035	ALGA0025611, H3GA0012886, ALGA0025638
5	2	121029	4:75519547-75640576	ASGA0020054, ASGA0020060, ALGA0025668, INRA0014665
5	3	177219	5:102816448-102993667	ALGA0033847, DRGA0006307, ASGA0026945, ALGA0033856
6	5	245796	5:103046152-103291948	ALGA0033870, MARC0004117, ALGA0033883, MARC0035348, ASGA0026958
3	1	55135	5:10866974-10922109	ALGA0030510, ALGA0030515
4	2	634017	5:20060561-20694578	ALGA0031066, INRA0018824, ASGA0025008
4	1	715352	5:5357091-6072443	DRGA0005428, H3GA0015339, ALGA0030145
4	1	6420	5:67275781-67282201	MARC0032012, ALGA0120187, MARC0074551
4	1	202254	6:10339260-10541514	ALGA0117024, MARC0083378, MARC0094765
4	1	77403	6:13899770-13977173	ALGA0119967, ALGA0116461, ALGA0118884
3	1	366770	6:20274364-20641134	MARC0055599, H3GA0017685
3	1	485434	6:22591223-23076657	MARC0006685, MARC0107808
4	3	456620	6:23679290-24135910	ALGA0034926, ASGA0027882, MARC0083540
4	1	30629	6:26402360-26432989	ASGA0098408, MARC0026696, ASGA0100103
3	1	57096	6:81717915-81775011	ALGA0108289, ASGA0028788
3	1	42071	6:9448971-9491042	ASGA0089652, ALGA0109170
4	1	61516	7:100960569-101022085	DRGA0008010, H3GA0022595, H3GA0022598
3	1	33119	7:111031594-111064713	ALGA0044331, ALGA0044333
4	2	230108	7:115056487-115286595	ALGA0044519, ALGA0044524, MARC0022069
3	1	25926	7:123276015-123301941	H3GA0023301, ALGA0045159
4	1	404124	7:14288108-14692232	ASGA0031366, H3GA0020057, ALGA0038804
4	2	79672	7:1532038-1611710	ASGA0030577, ALGA0037987, M1GA0009342
4	1	234276	7:15854705-16088981	H3GA0020089, ALGA0038863, MARC0067026
4	3	61320	7:16157280-16218600	DRGA0007180, ALGA0038878, ALGA0038890
4	2	109510	7:16354696-16464206	ASGA0031451, MARC0025074, ASGA0031469
5	3	248032	7:16532734-16780766	H3GA0020101, CASI0009829, ASGA0031485, ALGA0038950
3	1	26543	7:17100569-17127112	DRGA0007207, MARC0008851

4	1	106347	7:17366634-17472981	MARC0053893, M1GA0009653, ASGA0031511
3	2	30230	7:17837943-17868173	DRGA0007226, ASGA0031526
10	6	509413	7:17912283-18421696	DRGA0007227, H3GA0020136, ALGA0039015, DRGA0007230, ALGA0039018, MARC0024047, ASGA0031545, ASGA0031546, ALGA0039056
4	1	47668	7:21798615-21846283	H3GA0020279, ALGA0039343, ALGA0039345
3	2	59025	7:22888270-22947295	ALGA0039405, H3GA0020313
3	1	32943	7:24116868-24149811	H3GA0020334, ASGA0031873
3	2	21876	7:31157864-31179740	ALGA0039910, ALGA0039917
10	6	533238	7:32499186-33032424	DRGA0007462, ASGA0032266, ASGA0032282, H3GA0020623, ALGA0040040, MARC0063300, ASGA0032302, ASGA0032304, ALGA0040052
3	1	66283	7:33115876-33182159	ASGA0032316, ALGA0040066
4	1	93685	7:33205440-33299125	ASGA0032320, ASGA0032322, ALGA0040076
3	1	141661	7:35422882-35564543	ALGA0040284, ALGA0040291
4	1	449484	7:35880196-36329680	M1GA0010006, H3GA0020846, INRA0024809
3	1	103474	7:46433394-46536868	H3GA0021302, ALGA0040963
3	2	159424	7:48120121-48279545	ASGA0033314, M1GA0010251
3	1	94805	7:52184508-52279313	H3GA0021445, ALGA0041464
3	1	279238	8:105991646-106270884	MARC0111039, ALGA0103920
4	1	737231	8:135507570-136244801	INRA0030641, ASGA0039984, ASGA0097249
3	1	180541	8:141760561-141941102	MARC0044074, ALGA0049963
3	1	13687	8:142146987-142160674	ASGA0102346, ASGA0083677
5	3	29229	8:3607452-3636681	MARC0003410, ASGA0085562, ASGA0088966, ALGA0106161
6	5	636981	8:62838673-63475654	INRA0029837, ALGA0047962, MARC0064305, ALGA0047963, ALGA0047974
4	2	617563	8:64409837-65027400	ASGA0038865, ALGA0047982, H3GA0024905
4	3	138131	8:65489064-65627195	ALGA0047992, MARC0058200, ALGA0047995
5	3	550086	8:66072630-66622716	ALGA0048001, DRGA0008601, ALGA0048011, ALGA0048014
7	5	862789	8:67026060-67888849	MARC0000554, ALGA0048019, MARC0084543, MARC0050311, ALGA0048032, CASI0009341
3	2	43830	8:68852606-68896436	ALGA0048051, ALGA0048053
9	6	379396	8:69039573-69418969	ALGA0108174, ASGA0085207, MARC0020237, MARC0095739, ALGA0103392, ALGA0102491, ASGA0038884, MARC0069589

5	4	566446	8:69466912-70033358	H3GA0024937, H3GA0024938, DRGA0008622, ASGA0038898
3	2	207799	8:74321384-74529183	MARC0006179, ALGA0117771
3	2	36939	8:75092587-75129526	H3GA0024948, ALGA0048118
3	1	56480	8:75190777-75247257	ASGA0038938, ALGA0121925
6	3	793080	8:81585411-82378491	ALGA0117366, MARC0008576, MARC0042080, MARC0032054, ASGA0039088
5	2	202898	8:89778559-89981457	ALGA0111293, MARC0054558, ALGA0122094, ALGA0048454
4	1	80181	9:152870127-152950308	ALGA0056042, MARC0024811, MARC0094086
4	3	505806	9:34828091-35333897	ASGA0042433, H3GA0026947, ALGA0052304
3	1	300262	10:18781303-19081565	ALGA0057399, ASGA0090773
4	1	45067	10:30481283-30526350	ASGA0047214, ASGA0047209, ASGA0047207
6	4	698697	11:64041222-64739919	ALGA0062575, MARC0009324, ASGA0098788, SIRI0000656, ALGA0062604
3	1	225400	11:66469251-66694651	H3GA0032070, SIRI0000627
3	1	72696	11:67126490-67199186	H3GA0032097, ALGA0062775
4	1	76375	11:6988731-7065106	MARC0064230, DRGA0010767, MARC0068782
5	2	202815	12:13872843-14075658	H3GA0033589, H3GA0055422, MARC0113018, DIAS0003482
3	1	47120	12:14161680-14208800	MARC0075799, ALGA0065183
5	2	335913	12:14221473-14557386	ALGA0065186, ALGA0105006, H3GA0033606, MARC0070276
4	2	59417	12:14901690-14961107	MARC0032480, MARC0112630, ALGA0065273
3	1	4829	12:15315430-15320259	DIAS0002084, MARC0006983
4	1	81214	12:16590416-16671630	ALGA0065337, DRGA0011634, ALGA0065344
4	1	116403	12:3329592-3445995	ASGA0084858, M1GA0025366, M1GA0024830
3	1	193918	12:4992763-5186681	ALGA0064392, M1GA0015856
4	1	407854	12:50892615-51300469	DRGA0011794, ALGA0066854, DRGA0011803
3	1	317101	12:59563353-59880454	H3GA0034945, MARC0034728
3	1	106925	13:13619831-13726756	ALGA0068042, MARC0010390
4	2	58962	13:142153651-142212613	ASGA0058857, ALGA0072005, ASGA0058862
4	1	500131	13:150609181-151109312	ALGA0072233, MARC0036292, ALGA0072255
3	2	17603	13:177964129-177981732	MARC0065723, ALGA0072779
4	1	297854	13:179169750-179467604	ASGA0059304, ASGA0059307, CAHM0000041
4	1	803434	13:18620897-19424331	ASGA0056431, ASGA0056435, ALGA0068598
4	2	85763	13:196079847-196165610	ALGA0109927, ALGA0121208, DRGA0013338
3	1	21240	13:198288173-198309413	ASGA0059621, H3GA0037784

4	2	59471	13:210457466-210516937	ALGA0113830, MARC0019610, ASGA0089689
3	1	3009	13:210531092-210534101	ASGA0089950, ASGA0097399
4	1	98442	13:32330620-32429062	ALGA0069344, MARC0048245, ASGA0095010
7	3	262619	13:60176850-60439469	MARC0021871, ALGA0070336, MARC0097301, ALGA0070329, ASGA0057740, ALGA0070362
4	1	178245	13:60498666-60676911	ALGA0070351, DRGA0012439, ALGA0070372
4	1	46920	13:60697983-60744903	DRGA0012442, MARC0030786, ALGA0070377
3	1	211700	13:60768558-60980258	DRGA0012446, DRGA0012449
3	1	74307	13:64824612-64898919	ALGA0070489, ALGA0070492
3	1	377057	13:65211841-65588898	ASGA0105748, DRGA0012492
3	1	57117	13:65864959-65922076	ALGA0070517, H3GA0036617
3	1	120938	13:66227544-66348482	ALGA0070531, ALGA0070545
4	1	96230	13:67822598-67918828	ALGA0070583, ASGA0057875, DRGA0012516
3	1	23080	13:68184094-68207174	ALGA0070606, ALGA0070608
4	1	214390	13:70391963-70606353	H3GA0036671, ALGA0070637, ALGA0070646
11	7	887346	14:100494372-101381718	INRA0045783, MARC0090892, MARC0011234, ASGA0093262, CASI0009035, DRGA0014282, INRA0045792, ALGA0080004, H3GA0041574, DRGA0014283
10	5	852242	14:101477646-102329888	DRGA0014288, H3GA0041580, ASGA0065418, ALGA0080034, ALGA0080038, SIRI0000335, ASGA0101253, ALGA0080040, INRA0045833
4	1	208672	14:102629597-102838269	ALGA0080043, DRGA0014300, DRGA0014301
4	1	35793	14:103900532-103936325	ALGA0080094, INRA0045898, MARC0070934
4	1	70783	14:104083021-104153804	ALGA0080103, DRGA0014325, DRGA0014327
4	1	51420	14:104179804-104231224	ALGA0080105, DRGA0014328, ALGA0080107
4	1	98122	14:104661126-104759248	ALGA0080136, ASGA0065470, INRA0045948
5	2	168834	14:106587223-106756057	INRA0046003, ASGA0065507, DRGA0014387, H3GA0041633
4	1	209159	14:107457286-107666445	H3GA0041644, ASGA0096934, DRGA0014395
3	1	193934	14:109401525-109595459	ASGA0065593, ALGA0080399
3	1	343974	14:29514762-29858736	ASGA0062418, M1GA0018460
4	1	234982	14:30898763-31133745	MARC0064671, ALGA0076653, ASGA0062483
3	1	20500	14:3670854-3691354	ALGA0074358, ASGA0060510
3	1	324985	14:46211548-46536533	H3GA0040005, ALGA0077258
3	1	31568	14:85857890-85889458	ALGA0079160, SIRI0000058

4	1	117079	14:88941176-89058255	H3GA0041197, ASGA0064790, MARC0057510
4	1	186563	14:91200984-91387547	INRA0045347, ALGA0079389, DRGA0014191
4	1	108882	14:93481573-93590455	CASI0000717, H3GA0041266, MARC0034469
3	1	161639	15:14909574-15071213	ASGA0068640, INRA0048844
5	3	175323	15:152354557-152529880	M1GA0020805, MARC0022007, ASGA0071786, ALGA0112129
6	2	318423	15:19122717-19441140	ALGA0084105, MARC0087013, ASGA0068758, ALGA0084119, DRGA0014980
5	2	262598	15:21414820-21677418	ALGA0084241, MARC0047493, ALGA0084256, ALGA0084262
5	2	145553	15:22635811-22781364	ASGA0068926, MARC0102789, H3GA0043950, DRGA0014998
3	1	20029	15:22879717-22899746	CASI0010169, INRA0048907
3	1	12555	15:2964759-2977314	MARC0051886, ASGA0098801
4	1	198304	15:3017289-3215593	ALGA0102725, H3GA0043639, ASGA0068326
3	1	291838	15:31995248-32287086	ALGA0123902, ASGA0103427
4	3	35662	15:40821463-40857125	ALGA0084968, MARC0015767, SIRI0000090
4	1	319575	16:11509437-11829012	DRGA0015844, MARC0048886, MARC0083542
3	1	16952	16:24398034-24414986	MARC0039041, ALGA0089682
4	1	78180	16:25999742-26077922	ALGA0089813, INRA0051331, ASGA0072736
3	1	219075	16:27330893-27549968	ALGA0089896, ALGA0123380
4	1	187634	16:32914269-33101903	MARC0020433, DRGA0016071, DRGA0016070
3	2	125640	16:41222726-41348366	DRGA0016148, H3GA0046526
4	2	399194	16:64185043-64584237	ALGA0091071, ALGA0091060, ASGA0073684
3	1	13166	16:64777275-64790441	ALGA0117046, ALGA0122057
4	1	111514	16:70292057-70403571	H3GA0046923, MARC0114192, MARC0077437
3	1	171042	16:7300334-7471376	DRGA0015754, ALGA0088909
3	1	1106	16:81251422-81252528	MARC0041144, ALGA0109040
4	1	394671	17:12403467-12798138	H3GA0047804, ALGA0093127, ASGA0075297
3	2	321394	17:13901451-14222845	MARC0078025, MARC0038923
4	1	201834	17:50594610-50796444	MARC0089594, ASGA0077154, ALGA0095334
3	1	2892	17:52040112-52043004	ASGA0105240, ASGA0102687
3	1	15583	17:61236574-61252157	INRA0054509, ALGA0095986
4	2	56985	17:64925361-64982346	ALGA0096302, ASGA0078111, H3GA0049787
3	2	11813	18:13598735-13610548	MARC0078052, ALGA0097073
4	2	118524	18:4711435-4829959	ASGA0097549, ALGA0112743, MARC0029189

10	8	888809	18:51896691-52785500	MARC0071707, ALGA0098565, ASGA0080153, ALGA0098588, ALGA0098596, ALGA0098601, ALGA0098607, ASGA0097161, MARC0010291
8	6	1003202	18:52862142-53865344	ALGA0120564, DRGA0017057, ALGA0098636, ASGA0080235, MARC0051792, MARC0055353, ALGA0098680
4	1	17982	18:55225408-55243390	DIAS0003387, MARC0112120, M1GA0023373
3	2	212810	18:56551972-56764782	ASGA0080358, ASGA0090395
5	2	202082	18:56770029-56972111	ASGA0103963, ASGA0089739, ALGA0109418, H3GA0051209
3	1	66133	18:57033078-57099211	ASGA0080381, MARC0100363
3	1	72579	18:57514307-57586886	ALGA0108082, MARC0024065
4	1	176476	18:7383148-7559624	MARC0019932, ALGA0111929, ASGA0098693

Table S5. Results of the Gene Ontology (GO) enrichment analysis including annotated genes within regions around (±500 kbp) SNPs with $P_{Bonferroni}$ <0.10 or in enriched windows of SNPs with $P_{Bonferroni}$<0.10 (reported in Table S2).

David clusters obtained with annotations in ±500 kbp regions				
GO Category	**GO Term**	**Count**	**%**	**P value**
Annotation Cluster 1	**Enrichment Score: 2.6794643381665373**			
GOTERM_BP_FAT	GO:0006333~chromatin assembly or disassembly	21	1.860053	9.654E-5
GOTERM_BP_FAT	GO:0065004~protein-DNA complex assembly	16	1.417183	4.235E-4
GOTERM_BP_FAT	GO:0034728~nucleosome organization	16	1.417183	5.379E-4
GOTERM_BP_FAT	GO:0006334~nucleosome assembly	15	1.328609	5.812E-4
GOTERM_BP_FAT	GO:0006323~DNA packaging	18	1.594331	8.181E-4
GOTERM_BP_FAT	GO:0031497~chromatin assembly	15	1.328609	8.343E-4

David clusters obtained with annotations in enriched windows				
GO Category	**GO Term**	**Count**	**%**	**P value**
Annotation Cluster 1	**Enrichment Score: 3.7622400606355844**			
GOTERM_BP_FAT	GO:0006334~nucleosome assembly	13	2.783726	1.726E-6
GOTERM_BP_FAT	GO:0031497~chromatin assembly	13	2.783726	2.525E-6
GOTERM_BP_FAT	GO:0065004~protein-DNA complex assembly	13	2.783726	4.091E-6
GOTERM_BP_FAT	GO:0034728~nucleosome organization	13	2.783726	5.154E-6
GOTERM_BP_FAT	GO:0006333~chromatin assembly or disassembly	14	2.997859	2.717E-5
GOTERM_BP_FAT	GO:0006323~DNA packaging	13	2.783726	5.415E-5
Annotation Cluster 2	**Enrichment Score: 3.0114814064047395**			
GOTERM_MF_FAT	GO:0050785~advanced glycation end-product receptor activity	4	0.856531	7.55E-5
GOTERM_BP_FAT	GO:0048661~positive regulation of smooth muscle cell proliferation	7	1.498929	1.282E-4
GOTERM_BP_FAT	GO:0048660~regulation of smooth muscle cell proliferation	8	1.713062	1.729E-4
Annotation Cluster 3	**Enrichment Score: 2.0963014261964985**			
GOTERM_MF_FAT	GO:0042288~MHC class I protein binding	5	1.070664	7.229E-4
Annotation Cluster 4	**Enrichment Score: 1.9349276165885048**			
GOTERM_BP_FAT	GO:0003002~regionalization	16	3.426124	2.082E-4
GOTERM_BP_FAT	GO:0009952~anterior/posterior pattern formation	13	2.783726	3.015E-4
GOTERM_BP_FAT	GO:0007389~pattern specification process	18	3.85439	6.611E-4

Supplementary Figures

Figure S1. Estimated breeding values (EBVs) in the Italian Large White boar population over the 1992-2012 period for several traits (ADG = average daily gain, g; BFT = back fat thickness, mm; FGR = feed gain ratio; HW = ham weight, kg; LC = lean meat cuts, kg).

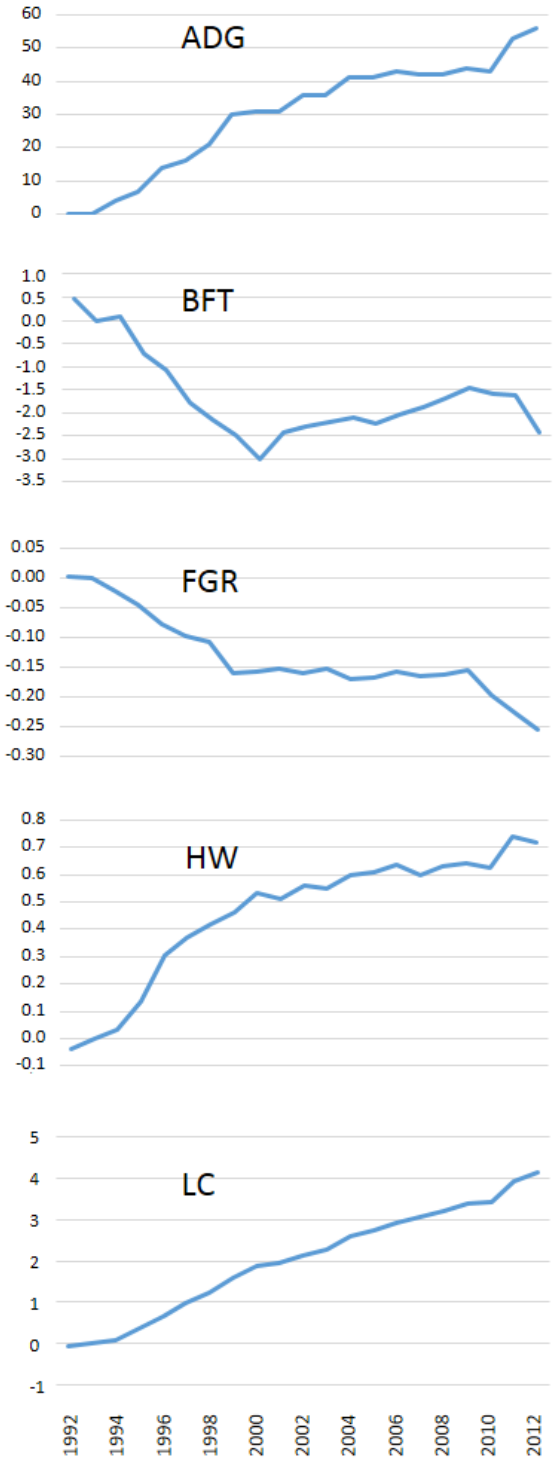
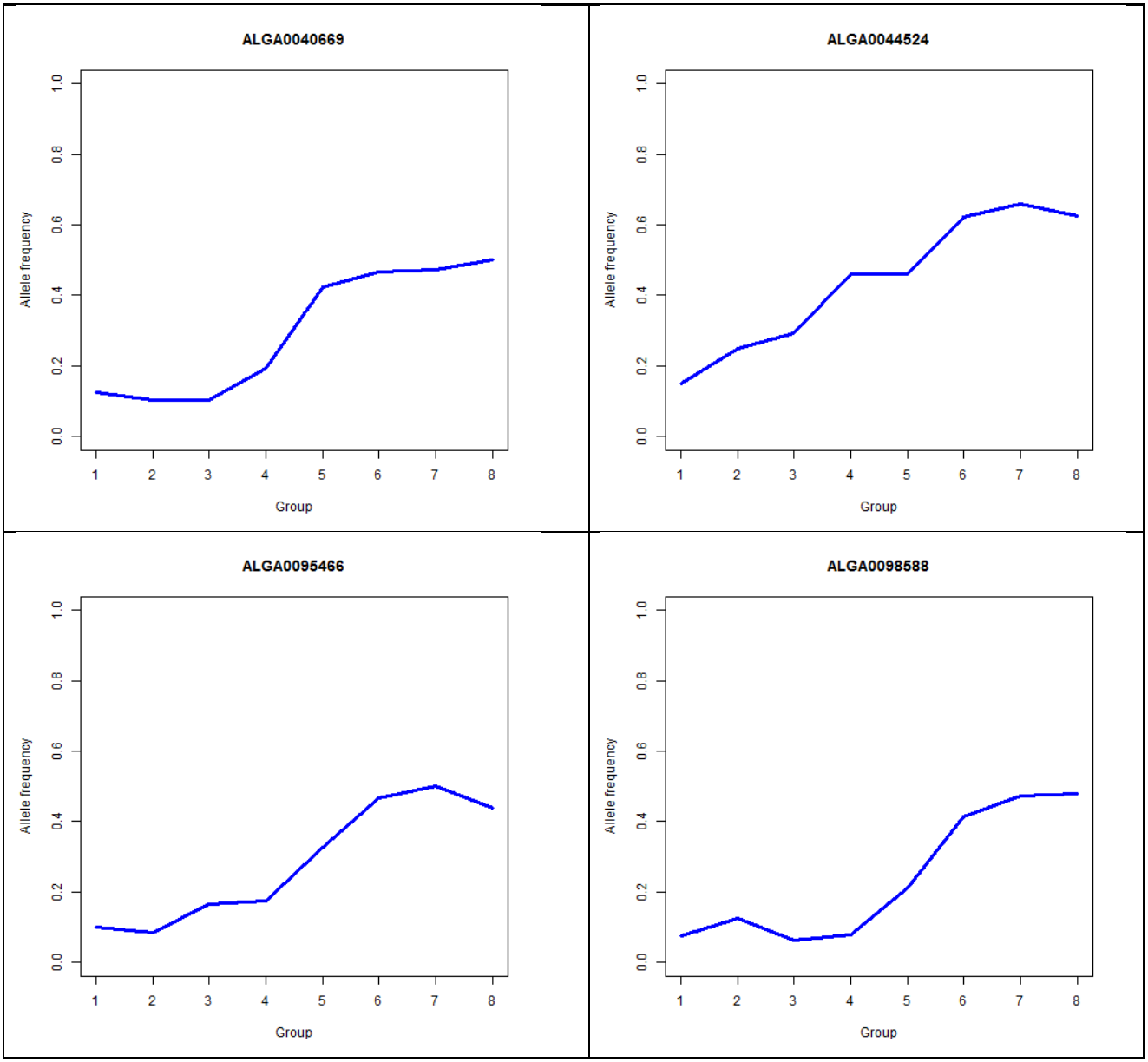
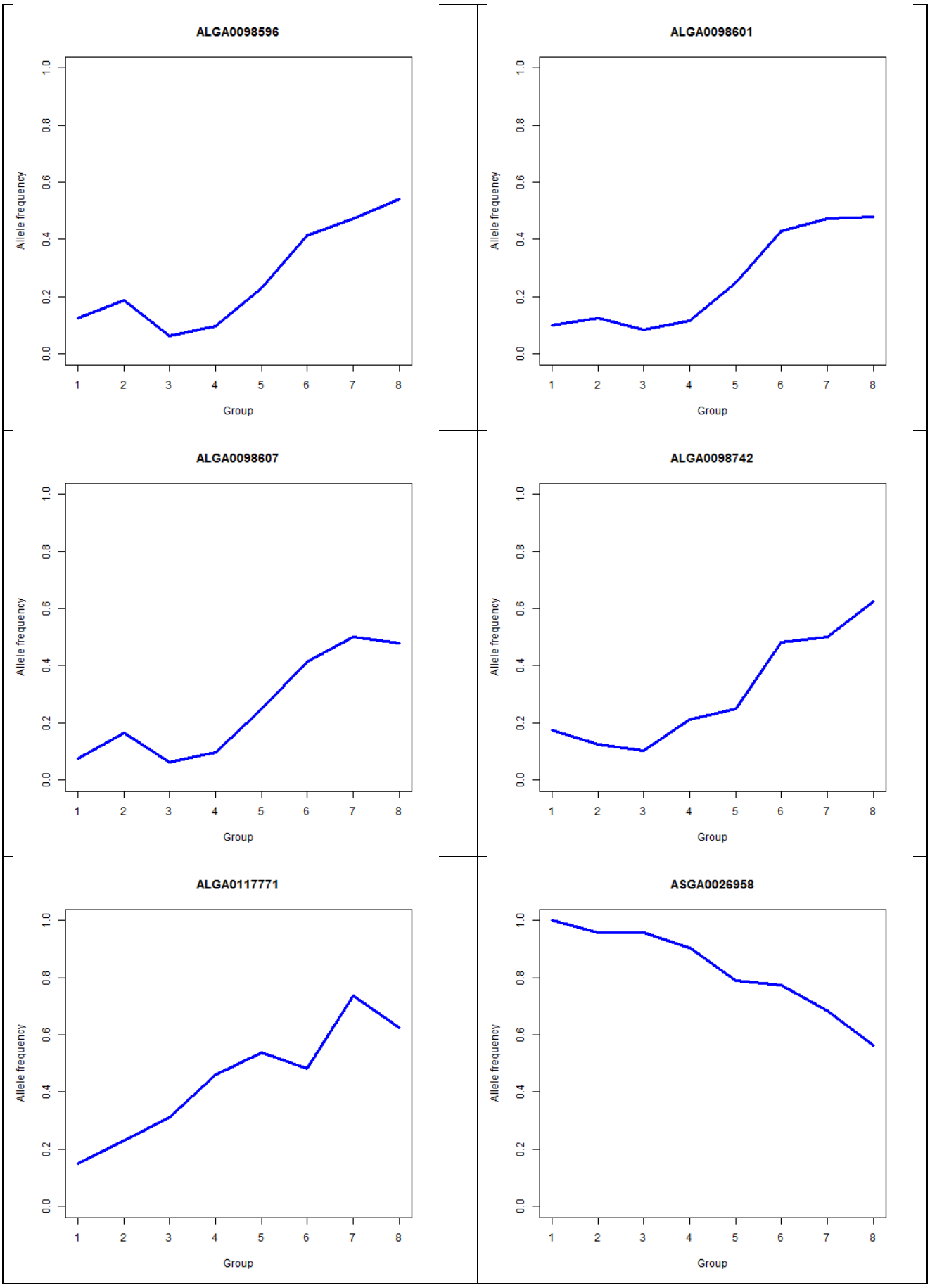
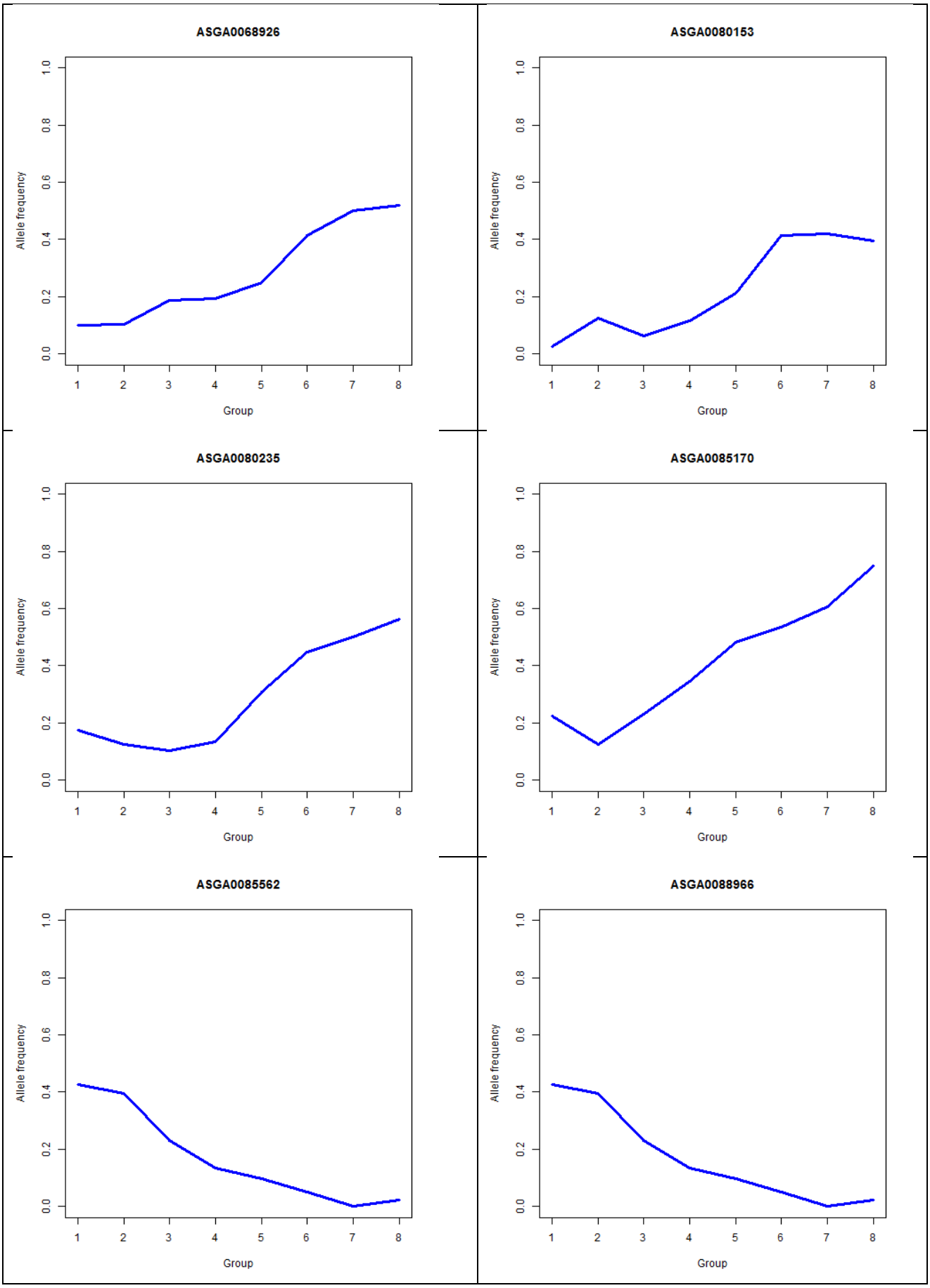
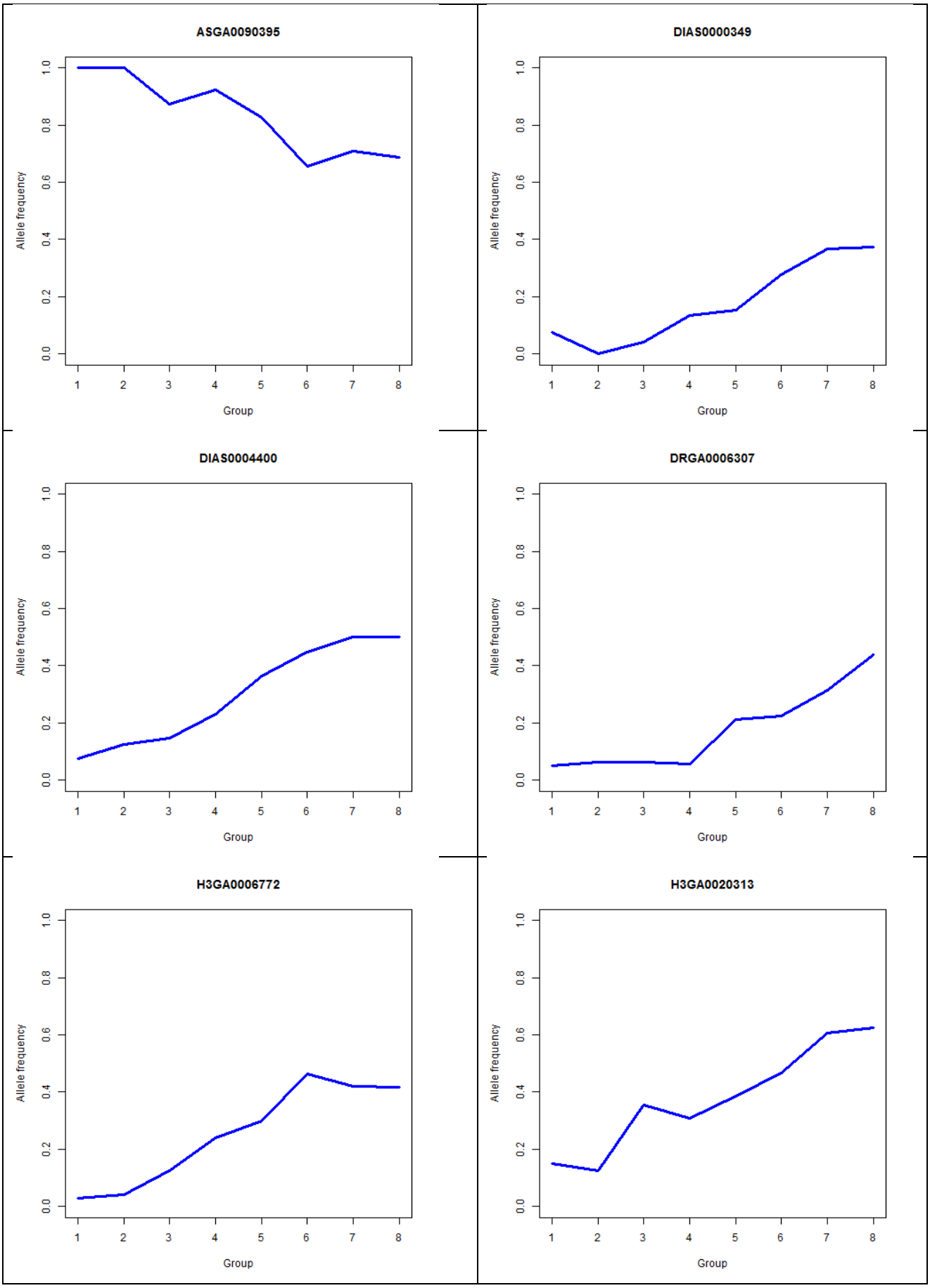


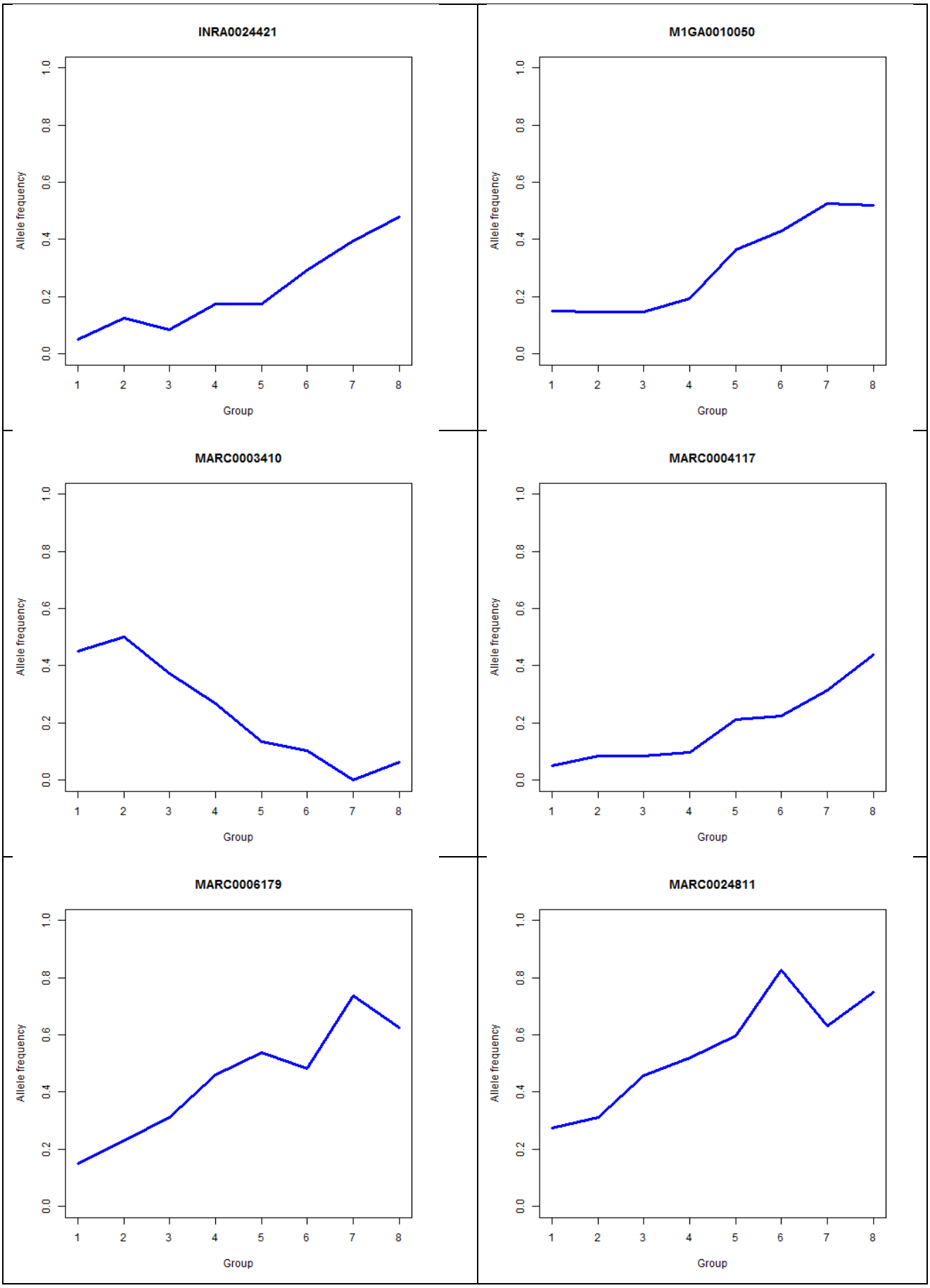
Figure S2. Change of allele frequencies during the considered time period (defined in eight time windows according to Table 1) for the most significant single nucleotide polymorphisms listed in Table 2 ($P_{\text{nominal value}}$ <1.0E-10). Frequencies are obtained for the first allele presented in Table 2 by counting alleles in the eight time windows.

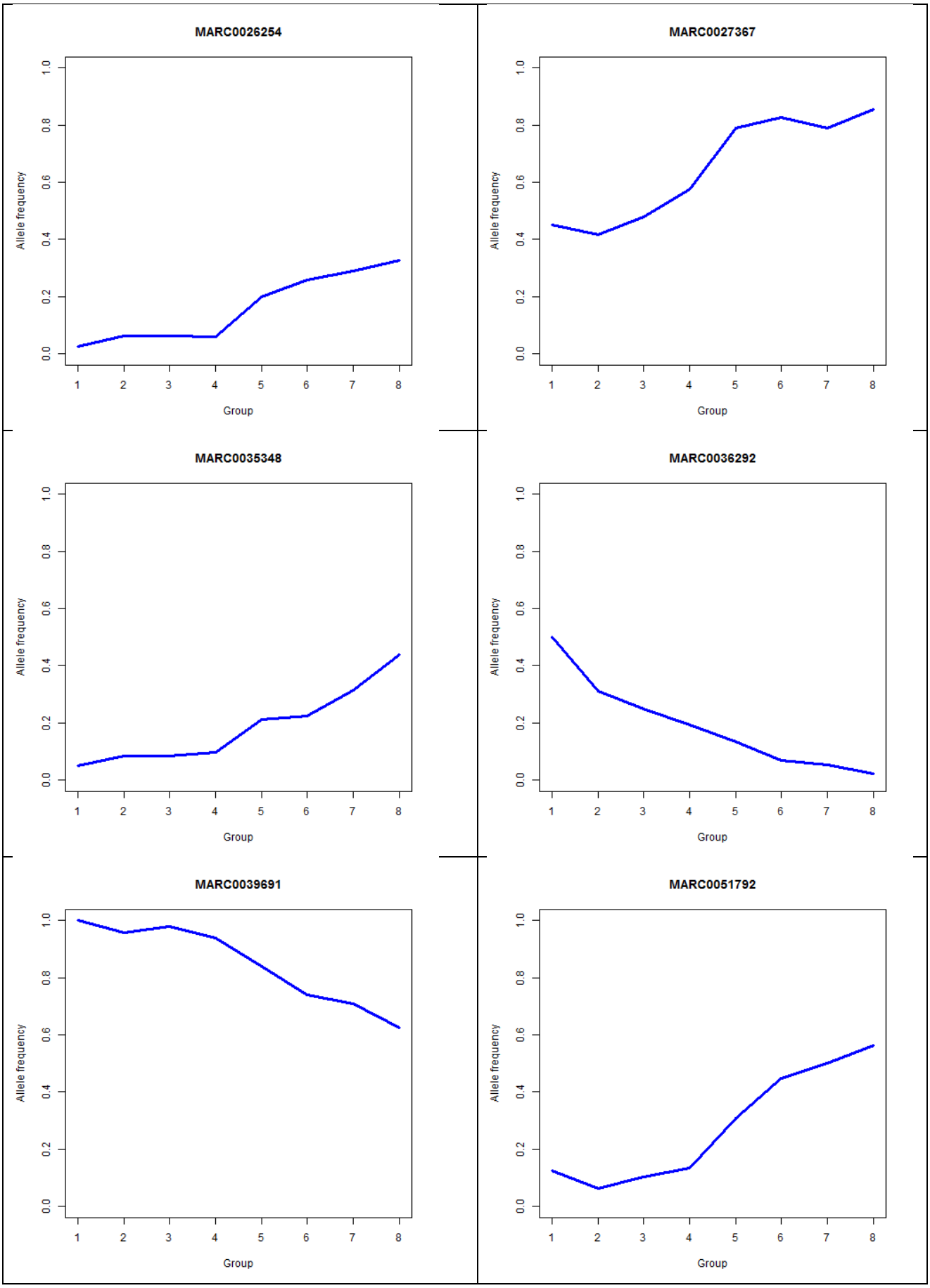


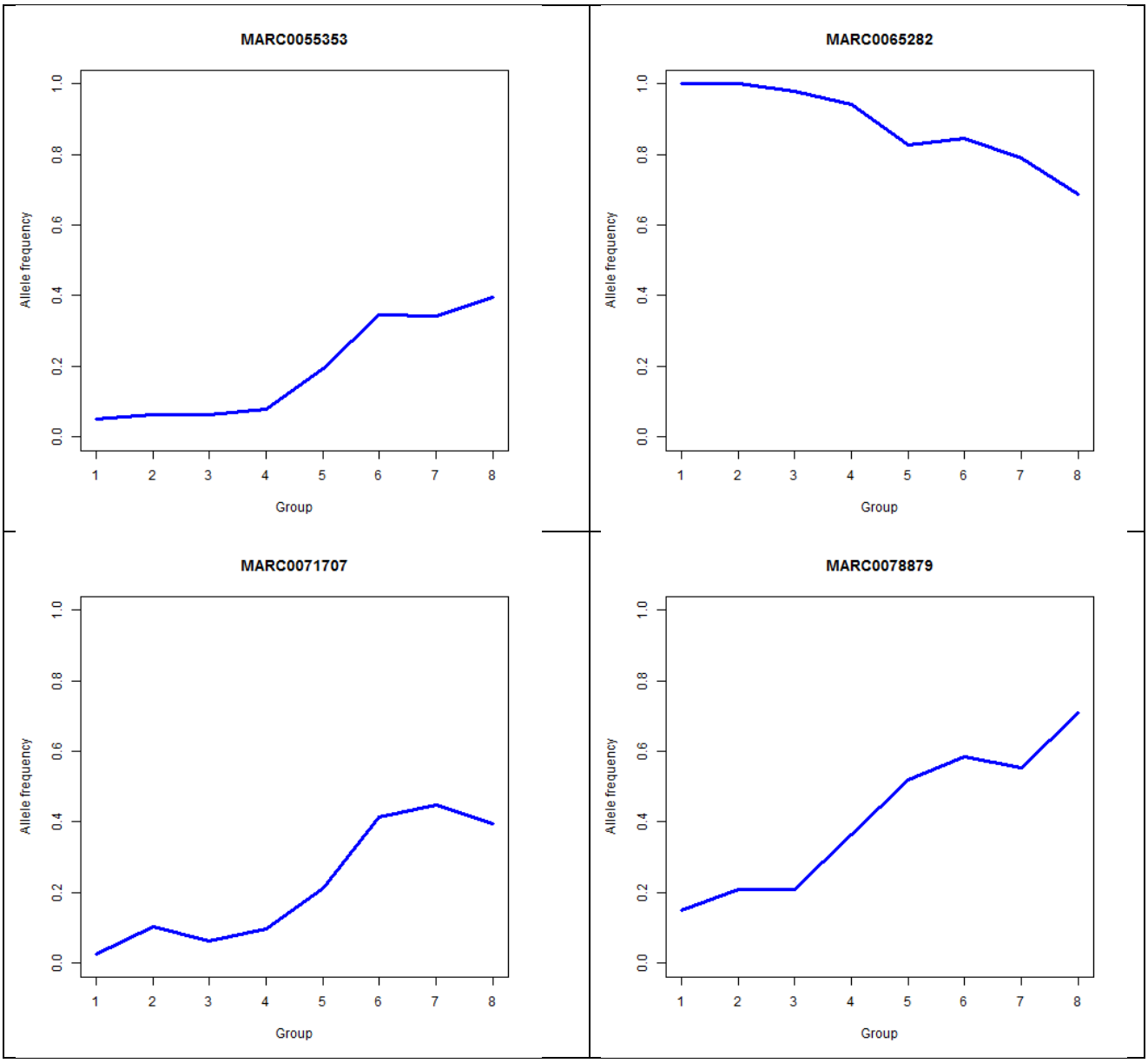

ALGA0098596
Allele frequency
Group
ALGA0098601
Allele frequency
Group
ALGA0098607
Allele frequency
Group
ALGA0098742
Allele frequency
Group
ALGA0117771
Allele frequency
Group
ASGA0026958
Allele frequency
Group


ASGA0068926
ASGA0080153
ASGA0080235
ASGA0085170
ASGA0085562
ASGA0088966
Allele frequency
Group
0.0
0.2
0.4
0.6
0.8
1.0
1
2
3
4
5
6
7
8


ASGA0090395
DIAS0000349
DIAS0004400
DRGA0006307
H3GA0006772
H3GA0020313
Allele frequency
Group
0.0
0.2
0.4
0.6
0.8
1.0
1
2
3
4
5
6
7
8


INRA0024421
M1GA0010050
MARC0003410
MARC0004117
MARC0006179
MARC0024811
Allele frequency
Group
0.0
0.2
0.4
0.6
0.8
1.0
1
2
3
4
5
6
7
8


MARC0026254
MARC0027367
MARC0035348
MARC0036292
MARC0039691
MARC0051792
Allele frequency
Group
0.0
0.2
0.4
0.6
0.8
1.0
1
2
3
4
5
6
7
8


MARC0055353
Allele frequency
0.0
0.2
0.4
0.6
0.8
1.0
1
2
3
4
5
6
7
8
Group
MARC0065282
Allele frequency
0.0
0.2
0.4
0.6
0.8
1.0
1
2
3
4
5
6
7
8
Group
MARC0071707
Allele frequency
0.0
0.2
0.4
0.6
0.8
1.0
1
2
3
4
5
6
7
8
Group
MARC0078879
Allele frequency
0.0
0.2
0.4
0.6
0.8
1.0
1
2
3
4
5
6
7
8
Group

CHAPTER 6

Reduced representation libraries from DNA pools analysed with next generation semiconductor based-sequencing to identify SNPs in extreme and divergent pigs for back fat thickness

Samuele Bovo,[1*] Francesca Bertolini,[1*] Giuseppina Schiavo,[1*] Gianluca Mazzoni,[1] Stefania Dall'Olio,[1] and Luca Fontanesi[1]

[1] Department of Agricultural and Food Sciences (DISTAL), Division of Animal Sciences, University of Bologna, Viale Fanin 46, 40127 Bologna, Italy

* These authors contributed equally to this work.

Published in International Journal of Genomics, 2015 Feb, http://dx.doi.org/10.1155/2015/950737

Abstract

The aim of this study was to identify single nucleotide polymorphisms (SNPs) that could be associated with back fat thickness (BFT) in pigs. To achieve this goal, we evaluated the potential and limits of an experimental design that combined several methodologies. DNA from two groups of Italian Large White pigs with divergent Estimating Breeding Value (EBV) for BFT were separately pooled and sequenced, after preparation of Reduced Representation Libraries (RRLs), on the Ion Torrent technology. Taking advantage from SNAPE for SNPs calling in sequenced DNA pools, 39,165 SNPs were identified, 1/4 of them were novel variants not reported in dbSNP. Combining sequencing data with Illumina PorcineSNP60 BeadChip genotyping results on the same animals, 661 genomic positions overlapped with a good approximation of minor allele frequency estimation. A total of 54 SNPs showing enriched alleles in one or in the other RRLs might be

potential markers associated with BFT. Some of these SNPs were close to genes involved in obesity related phenotypes.

Key words: Fat deposition; Heavy pigs; Ion Torrent semiconductor sequencing; Obesity; SNP

1. Introduction

The pig (*Sus scrofa*) is the most relevant agricultural meat species as well as an important animal model for its numerous physiological and morphological similarities to the human [1]. A parameter that is important for both aspects (meat production and animal model) is the level of fat deposition [2]. This is a complex phenotype that can be evaluated considering different traits. For example, back fat thickness (BFT) is a trait that affects ham and carcass values and, indirectly, correlates with production efficiency. For these reasons, breeding programs in most pig breeds and lines are designed to reduce BFT and increase lean meat content. In a few pig lines, an excessive reduction of the level of BFT could create problems to the meat processing industries as in the case of heavy pigs whose legs are cured for the production of dry-cured hams, and, for this reason, animals are selected to maintain an optimized fat thickness [3]. This trait is also an interesting phenotype to consider the pig as a model for human obesity [4, 5], that is one of the major health problems in both developed and developing countries.

To understand the biological mechanisms affecting BFT in pigs, we recently carried out several studies to elucidate the genetic factors involved in the definition of this trait and to obtain a systems biology comparative picture of human and pig obesity related traits [6]. In a whole genome candidate gene approach, we reported that polymorphisms in genes already shown to affect fat deposition in humans and mice are associated with BFT or correlated traits in commercial pigs and in the Italian Large White heavy pig breed [7-10]. In addition, a genome wide association (GWA) study we carried out in the same breed using a selective genotyping approach and the Illumina PorcineSNP60 BeadChip [11] showed quite a large number of markers associated with BFT (each with a small effect that could not explain the whole genetic variability for this trait), with a limited overlap with other GWA studies that investigated the same or similar traits in other breeds and pig populations [12]. This could be due to different experimental designs and incomplete power in the different studies as well as different linkage disequilibrium structures of the investigated

populations that could not be captured completely by the genotyping tool (Illumina PorcineSNP60 BeadChip).

Taking advantage from the sequenced genome of the pig and its reference assembly (Sscrofa10.2) [13], it is now possible to use next generation sequencing (NGS) platforms to further investigate the level and extent of genetic variability in different breeds and populations (i.e. [14]). The Ion Torrent technology is a cheap promising NGS platform that is based on a semiconductor detection of pH variation during the sequencing process that can be applied in different experimental approaches in which a medium-high throughput is needed [15]. We already evaluated the Ion Torrent platform to analyse a mammalian genome by sequencing reduced representation libraries (RRLs) obtained from rabbit genomic DNA and identified thousands of new single nucleotide polymorphisms (SNPs) in this species [16].

In this study, with the final aim to identify SNPs that could be useful to evaluate the peculiarities of the Italian Large White heavy pig breed and explain, at least in part, the missed genetic variability for the BFT trait not completely captured by our previous association works, we tested the potential and limits of an experimental design in which we combined the Ion Torrent sequencing technology to sequence RRLs. Reduced representation libraries were obtained by enzymatically digest DNA pools obtained constructed from divergent Italian Large White pigs with extreme estimated breeding value (EBV) for BFT. In addition, we used Illumina PorcineSNP60 BeadChip genotyping data already generated from the same animals to obtain a comparative analysis and validation of the sequencing information.

2. Materials and Methods

2.1. Animals and Genomic DNA

A subset of the Italian Large White pigs that were previously used in a GWA study carried out to identify markers associated with BFT EBV [12], were used to constitute two genomic DNA pools. The selected animals were from two groups, each of 50 pigs, of two-generations unrelated

gilts with extreme and divergent BFT EBV (50 with the most negative BFT EBV and 50 with the most positive BFT EBV), selected among about 12,000 pigs individually performance tested at the Central Test Station of the National Pig Breeder Association (ANAS) for the sib-testing evaluation of candidate boars within the national selection program of the Italian Large White breed [7, 9, 12]. Average and standard deviation of BFT EBV of the pigs in the negative and positive tails was: –9.40 ± 1.60 mm and +8.00 ± 5.95 mm, respectively. Estimated breeding values for this trait was calculated by a BLUP-multiple trait animal model including the fixed factors of batch, age at the beginning of test, date of slaughtering, inbreeding coefficient, body weight at slaughter and age at slaughter, besides the random factors of animal and litter.

Genomic DNA was extracted from blood using the Wizard® Genomic DNA Purification kit (Promega Corporation, Madison, WI, USA). Extracted DNA was quantified using a NanoPhotometer P-330 instrument (Implen GmbH, München, Germany) and pooled at equimolar concentration to constitute two DNA pools, one including DNA from the 50 Italian Large White pigs with the lowest BFT EBV and a second including DNA from the 50 Italian Large White pigs with the highest BFT EBV.

2.2. Genotyping

The investigated animals were previously genotyped with the Illumina PorcineSNP60 BeadChip (Illumina Inc., San Diego, CA, USA), interrogating 62,163 SNPs [11]. No filter was applied and all samples and genomic positions were retained for subsequent evaluation and comparison with sequencing data (see below).

2.3. Reduced Representation Libraries

Ten µg of DNA from each of the two pools were digested overnight with 50 U of *Hae*III restriction enzyme and the digested products were loaded in a 0.8% agarose gel. *Hae*III was selected as it did not produce visible patterns that could be ascribed to repetitive elements in the

range of 500-700 bp (data not shown). DNA fragments from this range obtained from *Hae*III digestion were purified from the agarose gel with the QIAquick Gel Extraction Kit (Qiagen, Hilden, Germany) according to the manufacturer instructions. Obtained DNA was used for library preparation and sequencing with the Ion Torrent PGM (Life Technologies, Carlsbad, CA, USA).

2.4. Ion Torrent sequencing

Sequencing of the two RRLs was obtained using 200 ng of DNA that was purified by agarose gel electrophoresis as described above, enzymatically sheared, end repaired and adapter ligated using the Ion Xpress™ Plus Fragment Library kit (Life Technologies). Obtained DNA material was size selected using the e-gel system (Invitrogen, Carlsbad, CA, USA) and bands corresponding to 100 bp of inserts were collected and quantified by qPCR using a StepOnePlus™ Real-Time PCR System (Life Technologies). Selected fragments were clonally amplified, purified and sequenced using the Ion One Touch™ 100 Template kit and the Ion PGM™ Sequencing kit with two Ion 318 chips (Life Technologies), for the two RRLs.

2.5. Sequence data analyses

Obtained sequencing reads were filtered and trimmed using the Ion Torrent suite v.2.2 (Life Technologies) which i) eliminated polyclonal sequences and sequences of low quality and ii) trimmed adapters and low quality 3'-ends. Then data were inspected with FastQC v.0.11.22 [17]. Sequenced reads were trimmed and filtered using PRINSEQ lite v.0.20.4 [18] as follow: i) trimming at the 3'-end up to 140 bp, ii) trimming of the 5'-end and 3'-end for poly A/T sequences > 5, iii) trimming the 5'-end and 3'-end up to reach a base with a quality score > 20, iv) exclusion of reads having average quality < 20 and v) exclusion of reads shorter than 20 bp. PCR duplicates were removed from each library using Picard v. 1.107 [19]. After the PCR duplicates removing step, reads were merged, processed and aligned on the Sscrofa10.2 genome version using BWA v.0.7.7 [18]. Reads aligning in only one place of the genome and with mapping quality score (Qm) > 20

were retained. SNP calling was obtained using SNAPE [20], setting divergence to 0.01, prior nucleotide diversity (θ) of 0.001, folded spectrum and filtering by a posterior probability of segregation > 0.90. SNAPE input files (PILEUP format) were obtained using Samtools v.0.1.4 [21, 22]. SNAPE filters were applied to consider only positions with minimum depth of 3×, to avoid INDELs (as indel calling algorithm are not specific for pools [23]). For each putative SNP, we identified if it was already included in the dbSNP or if it was new using the Ensembl Biomart data mining tool [24], interrogating the Ensembl Variation 77 database (October 2014) for Sscrofa10.2 short variations and indels (based on dbSNP build 140). All the SNPs that did not match with those reported on dbSNPs were also analyzed with the Samtools mpileup function [21, 22]. Variant Effect Predictor (VEP) tool (http://www.ensembl.org/tools.html; [25]) was used to map gene positions and to predict the effect of each substitution. SIFT [26] was used to evaluate if missense mutations could have deleterious effects on the translated proteins.

In order to evaluate differences in allele frequency derived by the number of alternative reads between the two RRLs, Fisher's exact test was computed for each alternative genomic position covered by a minimum depth of 3×. All the positions with $P_{Fisher} < 0.05$ were also visually inspected with IGV (Integrative Genomics Viewer) software [27]

3. Results

3.1. Sequencing data and identification of SNPs

A total of 3,390,796 and 3,731,776 sequenced reads were obtained from the two RRLs produced using the positive and negative BFT EBV DNA pools, respectively (Table 1). After cleaning the datasets for duplicated reads, the number of unique reads was 2,692,605 and 2,885,815, respectively (Table 1). A total of 1,449,838 (positive BFT EBV RRL) and 1,476,125 (negative BFT EBV RRL) reads were mapped with high confidence to the Sscrofa10.2 assembly of the pig genome. The merged dataset had an average read depth (RD) of 1.28× (range from 1 to 426×). Table S1 reports the number of reads and nucleotides mapped on the different pig chromosomes.

Sequence data obtained from the two RRLs have been submitted to the European Nucleotide Archive database (EMBL, http://www.ebi.ac.uk/ena/) and are indexed with the accession number ERP009239.

Using sequencing data, a total of 39,165 putative SNPs were called with high confidence by SNAPE [20]. Of these SNPs 24,560 (62.5%) were polymorphic carrying two alleles within the sequenced reads and 14,605 (37.58%) were monomorphic for an alternative form than that of the reference genome. We detected 9,680 new putative SNPs not yet reported in dbSNP (24.72% of the called SNPs) while the major part of identified variations (29,485; 75.28%) were already present in dbSNP. The Transition/Transversion ratio considering all the detected SNPs is 2.08, comparable to other mammalian genomes [28]. In addition, 6,324 of the newly detected SNPs were also detected using Samtools and 3,964 of these SNPs had score ≥20. Table 2 reports the summary of the annotations of the identified SNPs. Most of the SNPs were in intergenic (56.1%) or in intronic (28.9%) regions. The list of SNPs included in transcribed regions is reported in Table S2. Among the putative SNPs predicted in coding regions, 217 were synonymous mutations, 159 were missense mutations, two were stop-gained mutations (in the novel gene ENSSSCG00000028324 and in the NUT family member 2D gene, known as *NUTM2D*) and one was a stop lost variation (in the putative pleckstrin and Sec7 domain containing 2 gene; *PS2D*). Among the missense mutations, 37 were considered deleterious by SIFT (Table S2). Several genes with deleterious missense mutations [e.g., NADH dehydrogenase (ubiquinone) 1, subcomplex unknown, 1, 6kDa (*NDUFC1*); parathyroid hormone 1 receptor (*PTH1R*); glycerol-3-phosphate acyltransferase 2, mitochondrial (*GPAT2*); several olfactory receptor like genes] play important roles in different biochemical and physiological cellular mechanisms.

3.2. Sequencing vs PorcineSNP60 Beadchip genotyping data

To validate some of the called SNPs we took advantage from the Illumina PorcineSNP60 Beadchip genotyping data obtained on the same animals used to constructed the two RRLs.

Considering SNP positions covered by a minimum of three reads, 661 out of 62,163 SNPs of the chip (1.1%) were identified from the 13,596,939 sequenced positions (0.45% of the porcine genome). SNAPE analysis over these positions reported that i) 3 positions were discarded and 8 had read depth < 3 (for further features of SNAPE in addition to the general criteria adopted), ii) 257 were identified as SNPs (152 polymorphic SNPs carrying two alleles while 105 SNPs were monomorphic for an alternative form from that of the reference genome), iii) 375 positions showed only the sequence of the reference genome.

Of the overlapping 653 positions (661 – 8 = 653), i) for 28 of them the chip genotype data of the individual pigs were not possible to retrieve (probably due to problems in the design of the chip probes that could prevent the genotyping) and ii) for 63 DNA positions having all individuals homozygous for only one genotype, 59 of these base positions matched with the genotype inferred by NGS, whereas 2 were called as heterozygous and 2 were called as homozygous for a non-complementary nucleotide by sequencing data (Table S3). If we go into much details for the 28 SNPs that failed to report reliable genotyping data from the PorcineSNP60 Beadchip, for 12 out of 28 both alleles were present in the NGS reads, 15 out of 28 showed only one allele and one was an erroneous SNP.

In addition to these overlaps between NGS sequencing and genotyping data, we wanted to evaluate if estimated allele frequencies derived by NGS in RRLs obtained from DNA pools could match the true allele frequencies at the same positions obtained by using the PorcineSNP60 Beadchip. Starting from 559 SNPs (derived by the subsequent filtering steps of the 661 SNPs reported above), 262 (145 called as SNPs by SNAPE) had the same type of substitution. Excluding the transversions GC ↔ CG, AT ↔ TA, for each one of the remaining 258 SNPs, we compared the minor allele frequency (MAF) of the genotyping data against the frequency of the same allele derived by the sequencing. Results of the regression analysis are reported in Table 3 and in Figure 1**.** As expected, a low correlation from these two data was observed when considering all 258 SNPs due to the low coverage depth (3×) that was not enough for a reliable allele frequency estimation

from NGS data. This value increased up to 3 times setting a coverage depth equal or higher to the double of the minimum coverage depth (≥6). When adding data coming from monomorphic allele, correlation increased up to 0.70. These data indicate that even using a coverage depth ≥ 6 the MAF of these SNPs can be estimated with good approximation.

3.3. Sequencing derived SNPs: differences between the two libraries

For each of the two initial pileups we filtered out genomic positions having depth < 3x and then we used SNAPE to extract the allele frequency of each genomic position taking the advantage of the filters implemented in it. Polymorphic positions were compared among the 237,969 positions that were in common between the two RRLs (Table 1). Among these nucleotides, 67 genomic positions (filtered to 54 when tested by SNAPE and inspected with IGV) showed a $P_{Fisher} < 0.05$ comparing alternative reads observed in the two RRLs generated from DNA of pigs with extreme and divergent BFT EBV (Table S4). Only one of these SNPs showed a $P_{Fisher} < 0.01$. However, no one remained significant after Bonferroni correction. These SNPs were located in several autosomal chromosomes (SSC1, SSC3, SSC6, SSC8, SSC9, SSC10, SSC12, SSC15, SSC16, SSC17 and SSC18). These variants (only 12 of which already deposited in dbSNP) were localized as follows: 63% were intergenic variants, 21% were in introns, 11% were upstream gene variants and the 5% were downstream gene variants. Intronic variants were located in four genes of which only two were annotated with a known function: 1) Dysbindin (dystrobrevin binding protein 1) domain containing 1 (*DBNDD1*); 2) Phosphatidic acid phosphatase type 2A (*PPAP2A*).

3.4. Comparison with genome wide association results

In order to evaluate if the 54 SNPs that showed differences in number of alternative reads between the two RRLs were located in chromosome regions associated with BFT in Italian Large White pigs (listed in Table S4), we compared their positions on the basis of our previous GWA study carried out in the same breed [12]. We considered a window spanning ± 0.5 Mbp from each

marker having nominal P value < 0.05 in our previous study [12]. The top P_{Fisher} for each identified regions is reported in Table 4 (the complete list is reported in Table S5). The most significant marker (M1GA0008302; P=1.65E-06) is located 72,572 bp downstream SNP SSC6:859837 (P_{Fisher}=0.038) and 85,796 bp downstream the 6th top SNP SSC6:873061 (P_{Fisher}=0.012) obtained from the list of the 54 SNPs. In this region there is the Acyl-CoA synthetase family member 3 (*ACSF3*) gene that belongs to a family of enzymes that activate fatty acids. In the same region we previously showed that other markers (M1GA0008329, SSC6:996248, P=9.35E-05; and M1GA0008318, SSC6:945991; P=4.41E-04) were associated with BFT in the same breed. Within Table 4, the second most significant marker as reported previously (ALGA0000014, P=1.74E-05 [10]) is located close to the SNP SSC1:68514 (P_{Fisher}=0.029) identified in the present study (Table 4). In this region there is another marker associated with BFT in the previous GWA study (ALGA0000009, P=2.75E-03; [12]). An interesting gene located in this part of the pig genome [12], Delta-like 1 (Drosophila) (*DLL1*), seems associated to type 1 diabetes in humans. For marker DRGA009307 (SSC9:17138159, P= 8.66E-04) there is no annotated gene in a ± 500 kbp region. DIAS0000309 (SSC12:48865200, P=9.96E-04) is near the Active breakpoint cluster-related (*ABR*) gene and ENSSSCG00000017808 gene, orthologous of the Acyl-CoA-binding protein (*DBI*) gene. *ABR* gene is annotated with two interesting Gene Ontology (GO) terms: phospholipid binding and brain development. *DBI* gene functions as intracellular carrier of acyl-CoA esters and it seems that could act as a neuropeptide modulating the action of the GABA receptor. It is annotated with the GO terms: long-chain fatty acyl-CoA binding, transport, phosphatidylcholine acyl-chain remodeling and triglyceride metabolic process that might suggest a potential role in fat metabolism and deposition.

4. Discussion

Next generation sequencing is changing the way to identify markers associated with production traits in livestock species. Several applications and strategies have been designed mainly

using Illumina platforms (i.e. [14]). To our knowledge, this study applied for the first time the Ion Torrent technology to identify DNA polymorphisms in the pig genome. The experimental design was quite simple as, at this stage, we wanted to test this NGS technology to identify markers that could be useful for subsequent association studies in the Italian Large White pig breed. The identification of polymorphisms was based on the construction and sequencing of two RRLs generated from DNA pools of pigs with extreme and divergent BFT EBV. This approach was tested to set up a strategy for the identification of polymorphisms at a reduced fraction of the cost required for individual sequencing. In this way, we could also identify variants that might be enriched in one pool compared to the other one. To call SNPs we used SNAPE that is a software package that implemented a Bayesian approach for SNP identification and MAF estimation in sequenced pools [20]. The validation of identified SNPs was obtained by comparing the genotyping data generated with the Illumina PorcineSNP60 BeadChip on the same animals. As we sequenced DNA in pools and genotyping data were obtained on individual animals, we evaluated how allele frequency correlated between the two approaches varying the depth of sequencing. This approach was able to define an interesting procedure to validate SNPs identified from DNA pools.

Reduced representation libraries were generated as a simple strategy to reduce the complexity of mammalian genomes and to obtain information from a small part of it that can be sampled after restriction fragment digestion [29]. Several studies have already applied this strategy in farm animals for SNP discovery [16, 30-32]. For example, in pigs Wiedmann et al. [31] and Ramos et al. [11] sequenced RRLs for the identification of SNPs that were used to construct the Illumina PorcineSNP60 BeadChip genotyping platform. In our study, we identified about 40k SNPs in the pig genome. This is a quite large number of SNPs, considering the limited throughput of the benchtop Ion Torrent technology (compared to Illumina platforms [33]) and the stringent criteria that we used to call SNPs. As the technology is prone to errors in case of homopolymeric regions [34], indels were not considered in this study. That means that we could probably have discovered other short variants but we did not consider them to guarantee a high quality of the called

polymorphisms. In addition, other bioinformatics tools should be developed to obtain a reliable MAF estimation of indels from sequencing data generated from DNA pools [22].

Among the 159 SNPs causing missense mutations, 37 were predicted to affect the function of the encoded protein (Table S2). These polymorphisms will be prioritized to evaluate their association with several production traits together with SNPs whose alleles were differentially enriched in the two RRLs (Table 4, Table S4 and Table S5). The identification of these latter SNPs was based on allele frequency generated by mapping alternative reads in the two extreme groups of pigs with divergent BFT EBV. The low coverage of many SNP positions in both RRLs limited the possibility to identify markers associated with this trait. This problem is also due to the incomplete overlapping of read-coverage between the two RRLs. However, a comparative analysis of the nominally significant SNPs with our previous GWA study for BFT obtained using the same animals analyzed in this study [12], indirectly supported, to some extent, the identified association results. Some of these markers were located close to genes already shown in humans and mouse to be involved in obesity related phenotypes and pathologies a potential effect of these polymorphisms on BFT and fat deposition in Italian Large White pigs. These indications should be supported by association studies with fat deposition traits in the investigated breed or in other pig populations.

4. Conclusion

Several methodological approaches were tested in this study for the first time: i) partial sequencing obtained with Ion Torrent technology of the pig genome from DNA pools by using RRLs; ii) the application of SNP calling and MAF estimation on Ion Torrent low coverage sequencing data from DNA pools; iii) the validation of SNP called in DNA pools using individual genotyping data from the same animals of the pools; iv) the possibility to identify enriched alleles in the two sequenced RRLs representing two extremes for an important phenotypes (BFT). All these approaches were implemented in a case study that tried to identify additional markers associated with BFT in the Italian Large White pig breed. The purpose was to set up a strategy that could

reduce as much as possible the sequencing cost and that could produce data useful to identify novel markers for the targeted trait. Association studies will be carried out to evaluate the effects of the 54 selected markers.

Ion Torrent can be successfully applied for SNP discovery even if its limited throughput reduced the possibilities to obtain reliable allele frequencies in the two DNA pools. Other reductionist approaches, like genotyping-by-sequencing or genotyping by genome reducing and sequencing [35, 36], might be used to validate the most interesting called SNPs.

Conflict of Interests

The authors declare that they have no financial and personal relationships with other people or organizations that can inappropriately influence their work.

Authors' Contribution

Samuele Bovo, Francesca Bertolini and Giuseppina Schiavo contributed equally to this work.

Acknowledgements

We thank ANAS for providing data and samples, Sara De Fanti (BiGEA Dept.) and Emilio Scotti (DISTAL) of the University of Bologna for technical assistance, members of the Centre for Genome Biology for their support and Rita Casadio and Pier Luigi Martelli (Biocomputing Group, University of Bologna) for their advises on data analysis. This study was supported by Italian MiPAAF (INNOVAGEN project) and AGER-HEPIGET (Grant no. 2011-0279) funds.

References

[1] K. N. Kuzmuk, and L. B. Schook, "Pigs as a Model for Biomedical Sciences." In: M.F. Rothschild and A. Ruvinsky: The Genetics of the Pig, 2nd Edition. CAB International, Wallingford, UK, pp. 426-444, 2011.

[2] M. Switonski, M. Stachowiak, J. Cieslak, M. Bartz, and M. Grzes, “Genetics of fat tissue accumulation in pigs: a comparative approach.” Journal of Applied Genetics, vol. 51, no.2, pp. 153-168, 2010.

[3] P. Bosi and V. Russo, “The production of the heavy pig for high quality processed products.” Italian Journal of Animal Science, vol. 3, no. 4, pp. 309-321, 2004.

[4] L. J. Kogelman, S. Cirera, D. V. Zhernakova, M. Fredholm, L. Franke, and H. N. Kadarmideen, “Identification of co-expression gene networks, regulatory genes and pathways for obesity based on adipose tissue RNA Sequencing in a porcine model.” BMC Medical Genomics, vol. 1, article 57, 2014.

[5] L. J. Kogelman, S. D. Pant, M. Fredholm, and H. N. Kadarmideen, “Systems genetics of obesity in an F2 pig model by genome-wide association, genetic network, and pathway analyses.” Frontiers in Genetics, vol. 5, article 214, 2014.

[6] P. L. Martelli, L. Fontanesi, D. Piovesan, P. Fariselli, R. Casadio, “Mapping and annotating obesity-related genes in pig and human genomes.” Protein & Peptide Letters, vol. 21, no. 8, pp. 840-846, 2014.

[7] L. Fontanesi, C. Speroni, L. Buttazzoni, E. Scotti, S. Dall'Olio, L. Nanni Costa, R. Davoli, and V. Russo, “The insulin-like growth factor 2 (*IGF2*) gene intron3-g.3072G>A polymorphism is not the only *Sus scrofa* chromosome 2p mutation affecting meat production and carcass traits in pigs: evidence from the effects of a cathepsin D (*CTSD*) gene polymorphism.” Journal of Animal Science, vol. 88, no. 7, pp. 2235-2245, 2010.

[8] L. Fontanesi, E. Scotti, L. Buttazzoni, S. Dall'Olio, A. Bagnato, D. P. Lo Fiego, R. Davoli, and V. Russo, “Confirmed association between a single nucleotide polymorphism in the FTO gene and obesity-related traits in heavy pigs.” Molecular Biology Reports, vol. 37, no. 1, pp. 461-466, 2010.

[9] L. Fontanesi, G. Galimberti, D. G. Calò, R. Fronza, P. L. Martelli, E. Scotti, M. Colombo, G. Schiavo, R. Casadio, L. Buttazzoni, and V. Russo, “Identification and association analysis of

several hundred single nucleotide polymorphisms within candidate genes for back fat thickness in Italian Large White pigs using a selective genotyping approach." Journal of Animal Science, vol. 90, no. 8, pp. 2450-2464, 2012.

[10] L. Fontanesi, L. Buttazzoni, G. Galimberti, D. G. Calò, E. Scotti, V. Russo, "Association between melanocortin 4 receptor (*MC4R*) gene haplotypes and carcass and production traits in Italian Large White pigs evaluated with a selective genotyping approach." Livestock Science, vol. 157, no. 1, pp. 48-56, 2013.

[11] M. Ramos, R. P. Crooijmans, N. A. Affara, A. J. Amaral, A. L. Archibald, J. E. Beever, C. Bendixen, C. Churcher , R. Clark, P. Dehais, M. S. Hansen, J. Hedegaard , Z. L. Hu, H. H. Kerstens, A. S. Law, H. J. Megens, D. Milan, D. J. Nonneman, G. A. Rohrer, M. F. Rothschild, T. P. Smith, R. D. Schnabel, C. P. Van Tassell, J. F. Taylor, R. T. Wiedmann, L. B. Schook, and M. A. Groenen, "Design of a High Density SNP Genotyping Assay in the Pig Using SNPs Identified and Characterized by Next Generation Sequencing Technology." PLoS ONE, vol. 4, no. 8, article e6524, 2009.

[12] L. Fontanesi, G. Schiavo, G. Galimberti, D. G. Calò, E. Scotti, P. L. Martelli, L. Buttazzoni, R. Casadio, and V. Russo, "A genome wide association study for backfat thickness in Italian Large White pigs highlights new regions affecting fat deposition including neuronal genes." BMC Genomics, vol. 13, article 583, 2012.

[13] M. A. Groenen, A. L. Archibald, H. Uenishi, C. K. Tuggle, Y. Takeuchi, M. F. Rothschild, C. Rogel-Gaillard, C. Park, D. Milan, H. J. Megens, S. Li, D. M. Larkin, H. Kim, L. A. Frantz, M. Caccamo, H. Ahn, B. L. Aken, A. Anselmo, C. Anthon, L. Auvil, B. Badaoui, C. W. Beattie, C. Bendixen, D. Berman, F. Blecha, J. Blomberg, L. Bolund, M. Bosse, S. Botti, Z. Bujie, M. Bystrom, B. Capitanu, D. Carvalho-Silva, P. Chardon, C. Chen, R. Cheng, S. H. Choi, W. Chow, R. C. Clark, C. Clee, R. P. Crooijmans, H. D. Dawson, P. Dehais, F. De Sapio, B. Dibbits, N. Drou, Z. Q. Du, K. Eversole, J. Fadista, S. Fairley, T. Faraut, G. J. Faulkner, K. E. Fowler, M. Fredholm, E. Fritz, J. G. Gilbert, E. Giuffra, J. Gorodkin, D. K.

Griffin, J. L. Harrow, A. Hayward, K. Howe, Z. L. Hu, S. J. Humphray, T. Hunt, H. Hornshøj, J. T. Jeon, P. Jern, M. Jones, J. Jurka, H. Kanamori, R. Kapetanovic, J. Kim, J. H. Kim, K. W. Kim, T. H. Kim, G. Larson, K. Lee, K. T. Lee, R. Leggett, H. A. Lewin, Y. Li, W. Liu, J. E. Loveland, Y. Lu, J. K. Lunney, J. Ma, O. Madsen, K. Mann, L. Matthews, S. McLaren, T. Morozumi, M. P. Murtaugh, J. Narayan, D. T. Nguyen, P. Ni, S. J. Oh, S. Onteru, F. Panitz, E. W. Park, H. S. Park, G. Pascal, Y. Paudel, M. Perez-Enciso, R. Ramirez-Gonzalez, J. M. Reecy, S. Rodriguez-Zas, G. A. Rohrer, L. Rund, Y. Sang, K. Schachtschneider, J. G. Schraiber, J. Schwartz, L. Scobie, C. Scott, S. Searle, B. Servin, B. R. Southey, G. Sperber, P. Stadler, J. V. Sweedler, H. Tafer, B. Thomsen, R. Wali, J. Wang, J. Wang, S. White, X. Xu, M. Yerle, G. Zhang, J. Zhang, J. Zhang, S. Zhao, J. Rogers, C. Churcher, and L. B. Schook, "Analyses of pig genomes provide insight into porcine demography and evolution." Nature, vol. 491, no. 7424, pp. 393-398, 2012.

[14] M. Bosse, H. J. Megens, L. A. Frantz, O. Madsen, G. Larson, Y. Paudel, N. Duijvesteijn, B. Harlizius, Y. Hagemeijer, R. P. Crooijmans, and M. A. Groenen, "Genomic analysis reveals selection for Asian genes in European pigs following human-mediated introgression." Nature Communication, vol. 5, article 4392, 2014.

[15] J. M. Rothberg, W. Hinz, T. M. Rearick, J. Schultz, W. Mileski, M. Davey, J. H. Leamon, K. Johnson, M. J. Milgrew, M. Edwards, J. Hoon, J. F. Simons, D. Marran, J. W. Myers, J. F. Davidson, A. Branting, J. R. Nobile, B. P. Puc, D. Light, T. A. Clark, M. Huber, J. T. Branciforte, I. B. Stoner, S. E. Cawley, M. Lyons, Y. Fu, N. Homer, M. Sedova, X. Miao, B. Reed, J. Sabina, E. Feierstein, M. Schorn, M. Alanjary, E. Dimalanta, D. Dressman, R. Kasinskas, T. Sokolsky, J. A. Fidanza, E. Namsaraev, K. J. McKernan, A. Williams, G. T. Roth, and J. Bustillo, "An integrated semiconductor device enabling non-optical genome sequencing." Nature, vol. 475, no. 7356, pp. 348-352, 2011.

[16] F. Bertolini, G. Schiavo, E. Scotti, A. Ribani, P. L. Martelli, R. Casadio, and L. Fontanesi, "High-throughput SNP discovery in the rabbit (*Oryctolagus cuniculus*) genome by next-

generation semiconductor-based sequencing.” Animal Genetics, vol. 45, no. 2, pp. 304-307, 2014.

[17] S. Andrews, “FastQC: a quality control tool for high throughput sequence data.” 2010 Available at http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/

[18] H. Li and R. Durbin, “Fast and accurate short read alignment with Burrows-Wheeler transform.” Bioinformatics, vol. 25, no. 14, pp. 1754-1760, 2009.

[19] R. Schmieder and R. Edwards, “Quality control and preprocessing of metagenomic datasets.” Bioinformatics, vol. 27, no. 6, pp. 863-864, 2011.

[20] E. Raineri, L. Ferretti, A. Esteve-Codina, B. Nevado, S. Heath, and M. Pérez-Enciso, “SNP calling by sequencing pooled samples.” BMC Bioinformatics, vol. 13, article 239, 2012.

[21] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, and R. Durbin, “The Sequence Alignment/Map format and SAMtools.” Bioinformatics, vol. 27, no. 6, pp. 863-864, 2011.

[22] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, and 1000 Genome Project Data Processing Subgroup, "The Sequence alignment/map (SAM) format and SAMtools." Bioinformatics, vol. 25, pp. 2078-2079, 2009.

[23] A. Esteve-Codina, Y. Paudel, L. Ferretti, E. Raineri, H. J. Megens, L. Silió, M. C. Rodríguez, M. A. Groenen, S. E. Ramos-Onsins, and M. Pérez-Enciso, “Dissecting structural and nucleotide genome-wide variation in inbred Iberian pigs.” BMC Genomics, vol. 15, article 148, 2013.

[24] R. J. Kinsella, A. Kähäri, S. Haider, J. Zamora, G. Proctor, G. Spudich, J. Almeida-King, D. Staines, P. Derwent, A. Kerhornou, P. Kersey, and P. Flicek, “Ensembl BioMarts: a hub for data retrieval across taxonomic space.” Database (Oxford), vol. 2011, article bar030, 2011.

[25] W. McLaren, B. Pritchard, D. Rios, Y. Chen, P. Flicek, and F. Cunningham, “Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor.” Bioinformatics, vol. 26, no. 16, pp. 2069-2070, 2010.

[26] P. C. Ng and S. Henikoff, “Predicting deleterious amino acid substitutions.” Genome Research, vol. 11, no. 5, pp. 863-874, 2011.

[27] J. T. Robinson, H. Thorvaldsdóttir, W. Winckler, M. Guttman, E. S. Lander, G. Getz, and J. P. Mesirov, "Integrative Genomics Viewer." Nature Biotechnology, vol. 29, pp. 24–26, 2011.

[28] M. DePristo, E. Banks, R. Poplin, K.V. Garimella, J.R. Maguire, C. Hartl, A.A. Philippakis, G. del Angel, M.A. Ricas, M. Hanna, A. McKenna, T.J. Fennel, A.M. Kernytsky, A.Y. Sicachenko, K. Cibulskis, S.B. Gabriel, D. Altshuler, and M.J. Daly, “**A framework for variation discovery and genotyping using next-generation DNA sequencing data.”** ***Nature Genetics, vol. 43, no. 5, pp.*** 491-498, 2011.

[29] D. Altshuler, V. J. Pollara, C. R. Cowles, W.J. van Etten, J. Baldwin, L. Linton, and E. S. Lander, “An SNP map of the human genome generated by reduced representation shotgun sequencing.” Nature, vol. 407, no. 6803, pp. 513-516, 2000.

[30] C. P. Van Tassell, T. P. Smith, L. K. Matukumalli, J. F. Taylor, R. D. Schnabel, C. T. Lawley, C. D. Haudenschild, S. S. Moore, W. C. Warren, and T. S. Sonstegard, “SNP discovery and allele frequency estimation by deep sequencing of reduced representation libraries.” Nature Methods, vol. 5, no. 3, pp. 247-252, 2008.

[31] R. T. Wiedmann, T. P. Smith, and D.J. Nonneman, “SNP discovery in swine by reduced representation and high throughput pyrosequencing.” BMC Genetics, vol. 9, article 81, 2008.

[32] M. L. Aslam, J. W. Bastiaansen, M. G. Elferink, H. J. Megens, R. P. Crooijmans, le A. Blomberg, R. C. Fleischer, C. P. Van Tassell, T. S. Sonstegard, S. G. Schroeder, M. A. Groenen, and J. A. Long, “Whole genome SNP discovery and analysis of genetic diversity in Turkey (*Meleagris gallopavo*).” BMC Genomics, vol. 13, article 391, 2012.

[33] M. A. Quail, M. Smith, P. Coupland, T. D. Otto, S. R. Harris, T. R. Connor, A. Bertoni, H. P. Swerdlow, and Y. Gu, “A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers.” BMC Genomics, vol. 13, article 341, 2012.

[34] L. M. Bragg, G. Stone, M. K. Butler, P. Hugenholtz, and G. W. Tyson, “Shining a light on dark sequencing: characterizing errors in Ion Torrent PGM data.” PLoS Computational Biology, vol. 9, article e1003031, 2013.

[35] M. Li, S. Tian, C. K. Yeung, X. Meng, Q. Tang, L. Niu, X. Wang, L. Jin, J. Ma, K. Long, C. Zhou, Y. Cao, L. Zhu, L. Bai, G. Tang, Y. Gu, A. Jiang, X. Li, and R. Li, “Whole-genome sequencing of Berkshire (European native pig) provides insights into its origin and domestication.” Scientific Reports, vol. 4, article 4678, 2014.

[36] M. De Donato, S. O. Peters, S. E. Mitchell, T. Hussain, and I. G. Imumorin, “Genotyping-by-sequencing (GBS): a novel, efficient and cost-effective genotyping method for cattle using next-generation sequencing.” PLoS One, vol. 8, no. 5, article e62137, 2013.

[37] Z. Wang, Q. Chen, Y. Yang, H. Yang, P. He, Z. Zhang, Z. Chen, R. Liao, Y. Tu, X. Zhang, Q. Wang, and Y. Pan, “A genome-wide scan for selection signatures in Yorkshire and Landrace pigs based on sequencing data.” Animal Genetics, vol. 45, no. 6, pp. 808-816, 2014.

Tables

Table 1. Summary of sequencing data obtained from the two reduced representation libraries (RRLs) of the positive (Pos_HaeIII) and negative (Neg_HaeIII) back fat thickness estimated breeding value DNA pools.

Information[1]	**Pos_HaeIII**	**Neg_HaeIII**	**Pos+Neg HaeIII**
Sequenced reads	3,581,496	3,887,066	7,468,562
Reads after pre-processing	3,390,796	3,731,776	7,122,572
Removed duplicates	698,191	845,961	1,544,152
Mapped reads (Qm>20; Rdup)	1,449,838	1,476,125	2,925,963
Sequenced bases (Qm>20; Rdup)	137,429,598	145,859,611	256,880,473
Mean and Max depth of coverage (Qm>20; Rdup)	1.18; 209	1.16; 217	1.29; 426
Sequenced bases (Qm>20; RD≥3; Rdup)	3,394,898	3,057,171	3,942,266
Sequenced bases retained by SNAPE (Qm>20; RD≥3; Rdup)	3,369,555	3,034,731	237,969 (in common)
SNPs (Qm>20; RD≥3; Rdup)	10,694	10,339	39,165

[1]Qm= Mapping quality; RD=Read depth; Rdup=Removed duplicates

Table 2. Summary of the SNP annotation results obtained using the Variant Effect Predictor (VEP) tool.

Gene position or SNP effect	**No. of SNPs**
3 prime UTR variant	203
3 prime UTR variant, NMD transcript variant	1
5 prime UTR variant	58
downstream gene variant	2710
intergenic variant	24414
intron variant	12591
intron variant, NMD transcript variant	126
intron variant, non coding transcript variant	306
missense variant	159
missense variant, splice region variant	8
non coding transcript exon variant, non coding transcript variant	29
splice acceptor variant	2
splice donor variant	1
splice region variant, 3 prime UTR variant	1
splice region variant, intron variant	25
splice region variant, synonymous variant	12
stop gained	2
stop lost	1
stop retained variant	1
synonymous variant	217
synonymous variant, NMD transcript variant	3
upstream gene variant	2675
Total*	43545

*The sum includes 39,165 variations, 4,380 of which have multiple annotations, for a total of 43,545 SNP annotations.

Table 3. Summary of regression analysis between allele frequency estimated by Ion Torrent sequencing and the allele frequency obtained by genotyping with the Illumina PorcineSNP60 Beadchip.

RD	Polymorphic sites		Polymorphic and monomorphic sites	
	R^2	Positions	R^2	Positions
≥3	0.1199	258	0.6882	317 (258+59)
≥4	0.1601	99	0.6399	119 (99+20)
≥5	0.1611	36	0.5868	41 (36+5)
≥6	0.3866	11	0.7006	13 (11+2)

RD = Read depth; R^2 = Regression coefficient; Positions: Number of genomic sites analyzed.

Table 4. Overlapping results between the SNPs associated with back fat thickness as identify with the Ion Torrent sequencing data (P_{Fisher} <0.05) and those obtained in the genome wide association study (GWAS) reported by Fontanesi *et al.* [10] (P<0.05, window = ± 0.5 Mbp for each marker)

Chr	**Marker**	**Pos_M**	**P_{GWAS}**	**Pos_{SNP}**	***P_{Fisher}****
1	ALGA0000009	52,297	2.75E-03	68,514	2.86E-02
1	ALGA0000014	79,763	1.74E-05	68,514	2.86E-02
6	M1GA0008302	787,265	1.65E-06	873,061	1.28E-02
6	M1GA0008318	945,991	4.41E-04	873,061	1.28E-02
6	M1GA0008329	996,248	9.35E-05	873,061	1.28E-02
9	DRGA0009307	17,138,159	8.66E-04	16,885,924	2.81E-02
12	DIAS0000309	48,865,200	9.96E-04	48,937,212	2.63E-02

*Only the top $P_{Fischer}$ for each marker is listed. All other data are presented in Table S5.

Chr = Chromosome; Marker = Marker in the Illumina PorcineSNP60 Beadchip; Pos_M = Nucleotide position of the marker on the Sscrofa10.2 reference genome; P_{GWAS} = P-value of association in the GWAS; Pos_{SNP} = Nucleotide position on the Sscrofa10.2 reference genome of the SNP having $P_{Fischer}$< 0.05; $P_{Fischer}$ = P-value of the Fisher's test.

Figure 1. Scatter plot of allele frequency estimated by Ion Torrent sequencing data for SNPs called by at least 6 reads (Allele Frequency NGS) and obtained by genotyping data (MAF Genotyping) for the same SNPs.

Supplementary material legends

Table S1. Number of sequenced reads and nucleotides mapped on the different pig chromosomes.

Chr: Chromosome; Chr_length: Chromosome length based on Sscrofa10.2 reference genome; No. of mapped reads: Number of mapped reads with Qm>20 and without duplicates; % of mapped reads: Percentage of mapped reads with Qm>20. No. of sequenced bases: Number of sequenced bases with mapping quality >20 and without duplicates; Coverage: Percentage of the chromosome covered by reads.

Chr	Chr_length	N° of mapped reads	N° of sequenced bases
1	315321322	297876	26219709
2	162569375	187980	16615356
3	144787322	173835	15112285
4	143465943	148210	13162827
5	111506441	115019	10269186
6	157765593	207669	17738070
7	134764511	157139	13849488
8	148491826	139877	12574379
9	153670197	159996	14206866
10	79102373	96060	8226788
11	87690581	81623	7326217
12	63588571	91115	7818911
13	218635234	224566	20093779
14	153851969	192611	16904674
15	157681621	148053	12938114
16	86898991	91481	7643367
17	69701581	86199	7446599
18	61220071	70894	6016242
X	144288218	141898	12662327
Y	1637716	36	3650
SCAFFOLDS	200558837	113826	10051639

Location	Allele	Gene	Feature	Consequence	cDNA position	CDS position	Protein position	Amino acids	Codons	Existing variation	SYMBOL	SYMBOL SOURCE	HGNC ID	TREMBL	UNIPARC	SIFT	EXON	INTRON
1:159336	C	ENSSSCG00000030218	ENSSSCT00000026690	missense	2111	2111	704	L/R	cTg/cGg	rs336235415	-	-	-	I3LA70	UPI00025DF4AC	tol(0.06)	8/8	-
1:159364	T	ENSSSCG00000030218	ENSSSCT00000026690	missense	2083	2083	695	E/K	Gaa/Aaa	-	-	-	-	I3LA70	UPI00025DF4AC	tol(0.12)	8/8	-
1:159386	C	ENSSSCG00000030218	ENSSSCT00000026690	missense	2061	2061	687	I/M	atT/atG	rs331515993	-	-	-	I3LA70	UPI00025DF4AC	tol(0.29)	8/8	-
1:159418	G	ENSSSCG00000030218	ENSSSCT00000026690	missense	2029	2029	677	K/Q	Aaa/Caa	rs344334807	-	-	-	I3LA70	UPI00025DF4AC	tol(0.43)	8/8	-
1:164506	T	ENSSSCG00000030218	ENSSSCT00000026690	missense	115	115	39	P/T	Cca/Aca	-	-	-	-	I3LA70	UPI00025DF4AC	tol(0.11)	2/8	-
1:16221198	T	ENSSSCG00000004081	ENSSSCT00000030512	5'UTR	12	-	-	-	-	rs80846666	-	-	-	I3L589	UPI00025E1670	-	1/15	-
1:16221198	T	ENSSSCG00000004081	ENSSSCT00000004510	5'UTR	12	-	-	-	-	rs80846666	-	-	-	F1S7X1	UPI00025E166F	-	1/59	-
1:28741587	G	ENSSSCG00000004147	ENSSSCT00000023305	missense	647	647	216	V/A	gTg/gCg	rs80971879	ECT2L	HGNC	21118	I3LQ46	UPI00025DFEFA	tol(0.92)	4/16	-
1:28741587	G	ENSSSCG00000004147	ENSSSCT00000004583	missense	854	854	285	V/A	gTg/gCg	rs80971879	ECT2L	HGNC	21118	F1S6Y1	UPI00025DFEB9	tol(0.85)	6/20	-
1:34651624	C	ENSSSCG00000004187	ENSSSCT00000004626	missense	470	470	157	I/T	aTc/aCc	rs321977739	-	-	-	F1S3Q1	UPI00025DE782	tol(1)	1/1	-
1:34651680	T	ENSSSCG00000004187	ENSSSCT00000004626	missense	526	526	176	M/L	Atg/Ttg	-	-	-	-	F1S3Q1	UPI00025DE782	tol(0.66)	1/1	-
1:48351495	G	ENSSSCG00000004247	ENSSSCT00000004694	3'UTR	1910	-	-	-	-	rs345907689	ASF1A	HGNC	20995	F1SF37	UPI0001C95ED6	-	4/4	-
1:48351531	A	ENSSSCG00000004247	ENSSSCT00000004694	3'UTR	1874	-	-	-	-	rs325592320	ASF1A	HGNC	20995	F1SF37	UPI0001C95ED6	-	4/4	-
1:105554231	T	ENSSSCG00000004494	ENSSSCT00000004967	3'UTR	6696	-	-	-	-	rs343127038	-	-	-	F1RPT2	UPI00025E0B66	-	15/15	-
1:105554231	T	ENSSSCG00000004494	ENSSSCT00000004965	3'UTR	6699	-	-	-	-	rs343127038	-	-	-	F1RPU3	UPI00025E0B65	-	16/16	-
1:105554265	T	ENSSSCG00000004494	ENSSSCT00000004967	3'UTR	6662	-	-	-	-	rs338171243	-	-	-	F1RPT2	UPI00025E0B66	-	15/15	-
1:105554265	T	ENSSSCG00000004494	ENSSSCT00000004965	3'UTR	6665	-	-	-	-	rs338171243	-	-	-	F1RPU3	UPI00025E0B65	-	16/16	-
1:117313417	G	ENSSSCG00000004539	ENSSSCT00000005014	5'UTR	249	-	-	-	-	rs344934204	-	-	-	F1S1X6	UPI00017EFBFB	-	1/3	-
1:284996895	T	ENSSSCG00000005481	ENSSSCT00000006027	missense	110	110	37	V/D	gTc/gAc	rs337049123	RGS3	HGNC	9999	F1SN85	UPI0001C9612D	deleterious(0)	1/5	-
1:292734283	G	ENSSSCG00000005506	ENSSSCT00000006055	3'UTR	5412	-	-	-	-	rs335619736	MEGF9	HGNC	3234	F1SME8	UPI0001C9617C	-	6/6	-
1:296000953	T	ENSSSCG00000027068	ENSSSCT00000024301	missense	1016	1016	339	R/H	cGc/cAc	rs340362660	-	-	-	I3LHT3	UPI00025DF960	tol(0.06)	6/6	-
1:296055209	C	ENSSSCG00000005554	ENSSSCT00000006105	missense	670	670	224	I/V	Atc/Gtc	rs336080164	-	-	-	F1SLK6	UPI0001C96212	tol(0.54)	1/1	-
1:301554453	C	ENSSSCG00000005608	ENSSSCT00000006167	3'UTR	3551	-	-	-	-	-	ANGPTL2	Uniprot_gn	-	A8BV05	UPI00015D6779	-	5/5	-
1:301554454	A	ENSSSCG00000005608	ENSSSCT00000006167	3'UTR	3550	-	-	-	-	-	ANGPTL2	Uniprot_gn	-	A8BV05	UPI00015D6779	-	5/5	-
1:302314403	A	ENSSSCG00000005628	ENSSSCT00000006188	3'UTR	1488	-	-	-	-	rs319142132	ST6GALNAC2	Uniprot_gn	-	B2ZCZ7	UPI000174700F	-	6/6	-
1:303779881	A	ENSSSCG00000005688	ENSSSCT00000006255	3'UTR	946	-	-	-	-	-	PTGES	Uniprot_gn	-	Q2TJZ7,I3LE38,D0G7E4	UPI0000674A1D	-	3/3	-
1:303779983	T	ENSSSCG00000005688	ENSSSCT00000006255	3'UTR	844	-	-	-	-	rs328715874	PTGES	Uniprot_gn	-	Q2TJZ7,I3LE38,D0G7E4	UPI0000674A1D	-	3/3	-

1:305456644	C	ENSSSCG00000005717	ENSSSCT00000006287	3'UTR	2104	-	-	-	-	rs331135342	UCK1	HGNC	14859	F1S0W2	UPI00025DE8E7	-	7/7	-
1:305456662	T	ENSSSCG00000005717	ENSSSCT00000006287	3'UTR	2086	-	-	-	-	rs340127354	UCK1	HGNC	14859	F1S0W2	UPI00025DE8E7	-	7/7	-
1:305503917	T	ENSSSCG00000005719	ENSSSCT00000006289	3'UTR	4463	-	-	-	-	rs343779298	RAPGEF1	HGNC	4568	F1S0W0	UPI00025DE8EC	-	26/26	-
1:305503968	T	ENSSSCG00000005719	ENSSSCT00000006289	3'UTR	4412	-	-	-	-	rs332368026	RAPGEF1	HGNC	4568	F1S0W0	UPI00025DE8EC	-	26/26	-
1:306137129	A	ENSSSCG00000005723	ENSSSCT00000006293	3'UTR	2777	-	-	-	-	rs336520617	NTNG2	HGNC	14288	F1S0U7	UPI00025DE904	-	8/8	-
1:306730247	A	ENSSSCG00000005739	ENSSSCT00000006310	3'UTR	916	-	-	-	-	rs319925329	GTF3C5	HGNC	4668	F1S0S0	UPI00025DE92D	-	7/7	-
1:308989145	G	ENSSSCG00000005765	ENSSSCT00000006341	missense	431	431	144	N/S	aAt/aGt	-	-	-	-	F1RZZ5	UPI00025DE7A8	tol(0.05)	1/1	-
2:346785	A	ENSSSCG00000030137	ENSSSCT00000032261	missense	209	155	52	P/Q	cCa/cAa	-	LMNTD2	HGNC	28561	I3LKQ7	UPI00025E0B15	deleterious(0.03)	2/14	-
2:519552	T	ENSSSCG00000012857	ENSSSCT00000014053	missense	503	503	168	R/K	aGg/aAg	rs322010820	CARS	HGNC	1493	F1RYA3	UPI00025E0DB6	tol(0.61)	4/23	-
2:5567788	G	ENSSSCG00000012971	ENSSSCT00000014178	3'UTR	2981	-	-	-	-	rs327895371	EFEMP2	HGNC	3219	F1RU22	UPI0001C97AD8	-	11/11	-
2:5567788	G	ENSSSCG00000012973	ENSSSCT00000014180	3'UTR	1796	-	-	-	-	rs327895371	MUS81	HGNC	29814	F1RU20	UPI000210508E	-	16/16	-
2:6130174	G	ENSSSCG00000012999	ENSSSCT00000014210	3'UTR	2390	-	-	-	-	rs339408248	CAPN1	Uniprot_gn	-	Q9MZ23,Q9MZ22,Q38PP0,Q38PN9,Q38PN8,Q2EF32,Q0Q4H3,F1B3B2	UPI0000126E92	-	22/22	-
2:7815128	G	ENSSSCG00000027453	ENSSSCT00000028828	3'UTR	1266	-	-	-	-	rs344132045	-	-	-	I3LPA0	UPI00025E0CB7	-	4/4	-
2:7815161	T	ENSSSCG00000027453	ENSSSCT00000028828	3'UTR	1299	-	-	-	-	rs324957515	-	-	-	I3LPA0	UPI00025E0CB7	-	4/4	-
2:11430551	G	ENSSSCG00000013127	ENSSSCT00000014347	missense	277	277	93	S/P	Tcc/Ccc	-	OR4D9	HGNC	15178	F1RMI7	UPI0001C97D32	deleterious(0.01)	1/1	-
2:11430604	A	ENSSSCG00000013127	ENSSSCT00000014347	missense	224	224	75	S/F	tCc/tTc	rs336522689	OR4D9	HGNC	15178	F1RMI7	UPI0001C97D32	deleterious(0.05)	1/1	-
2:11465240	T	ENSSSCG00000013131	ENSSSCT00000014351	missense	283	283	95	D/N	Gat/Aat	-	OR4D6	HGNC	15175	F1RMI3	UPI0001C97D36	tol(1)	1/1	-
2:11815611	A	ENSSSCG00000013145	ENSSSCT00000014365	3'UTR	2478	-	-	-	-	rs346080101	DTX4	HGNC	29151	F1RMG1	UPI00025E0E73	-	9/9	-
2:12874417	G	ENSSSCG00000013177	ENSSSCT00000014397	5'UTR	113	-	-	-	-	rs343307087	MED19	HGNC	29600	F1SJY0	UPI0001C97DAC	-	1/5	-
2:13404635	C	ENSSSCG00000028527	ENSSSCT00000032273	missense	1310	1084	362	T/P	Aca/Cca	rs337398845	TNKS1BP1	HGNC	19081	I3LGJ5	UPI00025DF718	tol(1)	5/13	-
2:14031395	G	ENSSSCG00000013201	ENSSSCT00000014421	missense	526	526	176	T/A	Acc/Gcc	rs81254850	OR9G4	HGNC	15322	F1SJ64	UPI00025DFB58	tol(0.72)	3/3	-
2:14031448	C	ENSSSCG00000013201	ENSSSCT00000014421	missense	579	579	193	L/F	ttA/ttC	rs81254856	OR9G4	HGNC	15322	F1SJ64	UPI00025DFB58	tol(1)	3/3	-
2:14766381	A	ENSSSCG00000021435	ENSSSCT00000015881	missense	89	89	30	S/F	tCt/tTt	rs334377390	-	-	-	F1RH27	UPI0001C98067	tol(0.13)	1/1	-
2:15044430	G	ENSSSCG00000030330	ENSSSCT00000023161	missense	494	494	165	H/P	cAc/cCc	-	-	-	-	I3LLS4	UPI00025DFC09	tol(1)	1/2	-
2:15595310	T	ENSSSCG00000029415	ENSSSCT00000031270	missense	430	430	144	A/T	Gct/Act	rs345583511	-	-	-	I3LEU2	UPI0001E864A8	tol(0.1)	1/1	-
2:15595318	T	ENSSSCG00000029415	ENSSSCT00000031270	missense	422	422	141	R/H	cGt/cAt	rs327909965	-	-	-	I3LEU2	UPI0001E864A8	tol(0.07)	1/1	-
2:28831939	A	ENSSSCG00000013301	ENSSSCT00000014527	missense	159	71	24	T/N	aCc/aAc	rs81227061	ELF5	HGNC	3320	F1SGT0	UPI0001E86563	deleterious(0.01)	2/7	-
2:28831941	A	ENSSSCG00000013301	ENSSSCT00000014527	missense	161	73	25	D/N	Gac/Aac	rs334046745	ELF5	HGNC	3320	F1SGT0	UPI0001E86563	deleterious(0.02)	2/7	-

2:30902389	T	ENSSSCG00000013316	ENSSSCT00000027177	5'UTR	197	-	-	-	-	rs339188231	WT1	Uniprot_gn	-	Q9TSX0,O18760,I3L5W6	UPI00025E0456	-	1/11	-
2:30902389	T	ENSSSCG00000013316	ENSSSCT00000014542	5'UTR	197	-	-	-	-	rs339188231	WT1	Uniprot_gn	-	Q9TSX0,O18760,F1SGQ9	UPI00025E0457	-	1/10	-
2:60107341	C	ENSSSCG00000013873	ENSSSCT00000015156	5'UTR	65	-	-	-	-	-	NR2F6	HGNC	7977	F1S962	UPI0001C97EDA	-	1/4	-
2:62006657	G	ENSSSCG00000013829	ENSSSCT00000015107	3'UTR	4767	-	-	-	-	-	SYDE1	HGNC	25824	F1SAN7	UPI0001C97C9E	-	8/8	-
2:68169776	C	ENSSSCG00000013690	ENSSSCT00000014955	missense	101	101	34	L/P	cTg/cCg	rs331864115	-	-	-	F1S2R2	UPI0001C97F7D	deleterious(0)	2/2	-
2:72174240	C	ENSSSCG00000013569	ENSSSCT00000014821	missense	494	463	155	W/R	Tgg/Cgg	rs344841648	PEX11G	HGNC	20208	F1SCJ5	UPI00025E11A2	tol(0.39)	4/4	-
2:87188490	A	ENSSSCG00000014089	ENSSSCT00000015390	3'UTR	2644	-	-	-	-	-	F2RL2	HGNC	3539	F1S2H6	UPI0001C98573	-	2/2	-
2:91281262	C	ENSSSCG00000014126	ENSSSCT00000015430	3'UTR	3862	-	-	-	-	rs336131301	MSH3	HGNC	7326	F1RF09	UPI00025DF5A7	-	30/30	-
2:100130296	A	ENSSSCG00000024741	ENSSSCT00000026135	missense	143	25	9	G/R	Ggg/Agg	-	-	-	-	I3LF13	UPI00025DF864	tol(0.3)	3/6	-
2:124691703	G	ENSSSCG00000014218	ENSSSCT00000015536	5'UTR	196	-	-	-	-	rs329134304	TMED7	HGNC	24253	F1RLE5	UPI0001C981F1	-	1/3	-
2:147540285	A	ENSSSCG00000029541	ENSSSCT00000030535	stop_lost	1064	946	316	*/K	Tag/Aag	rs321613979	-	-	-	I3LAZ5	UPI00025DEA14	-	10/11	-
2:148345545	C	ENSSSCG00000014376	ENSSSCT00000015706	5'UTR	724	-	-	-	-	rs325004341	HARS2	HGNC	4817	F1RGD8	UPI0001C98467	-	1/13	-
2:149253466	A	ENSSSCG00000021051	ENSSSCT00000022662	missense	1696	1696	566	A/T	Gcc/Acc	rs339436969	PCDHGA10	HGNC	8697	I3LHJ4	UPI00025DE6E8	tol(0.67)	1/1	-
2:149494394	A	ENSSSCG00000014390	ENSSSCT00000015727	3'UTR	2496	-	-	-	-	rs319481778	FCHSD1	HGNC	25463	F1RMS3	UPI00025DEDC0	-	20/20	-
2:149494394	A	ENSSSCG00000014390	ENSSSCT00000028864	3'UTR	2523	-	-	-	-	rs319481778	FCHSD1	HGNC	25463	I3LGQ3	UPI00025DEDBF	-	21/21	-
2:158629967	G	ENSSSCG00000014457	ENSSSCT00000015805	missense	527	527	176	M/R	aTg/aGg	rs324531515	-	-	-	F1RL59	UPI0001C985D6	deleterious(0)	1/1	-
2:159217172	T	ENSSSCG00000028482	ENSSSCT00000015830	missense	617	617	206	T/M	aCg/aTg	rs327702647	-	-	-	F1RHZ0	UPI00025DE73F	tol(0.11)	1/1	-
3:4281335	T	ENSSSCG00000007577	ENSSSCT00000008316	missense	115	115	39	P/S	Ccc/Tcc	rs343193772	AP5Z1	HGNC	22197	F1RI49	UPI00025DF71A	tol(0.71)	1/16	-
3:4520547	A	ENSSSCG00000007580	ENSSSCT00000008319	3'UTR	4635	-	-	-	-	rs345364277	WIPI2	Uniprot_gn	-	D7RA31	UPI0001D6282A	-	12/12	-
3:4520570	G	ENSSSCG00000007580	ENSSSCT00000008319	3'UTR	4658	-	-	-	-	rs333942033	WIPI2	Uniprot_gn	-	D7RA31	UPI0001D6282A	-	12/12	-
3:4520580	C	ENSSSCG00000007580	ENSSSCT00000008319	3'UTR	4668	-	-	-	-	rs345269678	WIPI2	Uniprot_gn	-	D7RA31	UPI0001D6282A	-	12/12	-
3:5169777	G	ENSSSCG00000024245	ENSSSCT00000026197	3'UTR	2185	-	-	-	-	rs341535542	AIMP2	HGNC	20609	F1RFM7	UPI0001C96921	-	4/4	-
3:10281349	C	ENSSSCG00000007700	ENSSSCT00000008443	3'UTR	5211	-	-	-	-	rs325214457	HIP1	HGNC	4913	F1RKC0	UPI0001C966D7	-	31/31	-
3:10598642	A	ENSSSCG00000007710	ENSSSCT00000008453	3'UTR	3003	-	-	-	-	rs322446395	CHREBP	Uniprot_gn	-	F1RJN3,Q0QHM0	UPI0001C966E1	-	17/17	-
3:18801041	A	ENSSSCG00000007808	ENSSSCT00000008556	missense	571	502	168	A/T	Gct/Act	rs331373717	NFATC2IP	HGNC	25906	F1RFH6	UPI0001E868D5	tol(0.86)	3/8	-
3:18801042	A	ENSSSCG00000007808	ENSSSCT00000008556	missense	572	503	168	A/D	gCt/gAt	rs339629291	NFATC2IP	HGNC	25906	F1RFH6	UPI0001E868D5	tol(0.53)	3/8	-
3:25972329	T	ENSSSCG00000007854	ENSSSCT00000008604	3'UTR	2444	-	-	-	-	rs336219303	DCUN1D3	HGNC	28734	F1RPB5	UPI0001C96A25	-	2/2	-
3:25972334	T	ENSSSCG00000007854	ENSSSCT00000008604	3'UTR	2449	-	-	-	-	rs319797113	DCUN1D3	HGNC	28734	F1RPB5	UPI0001C96A25	-	2/2	-

3:26966004	T	ENSSSCG00000007866	ENSSSCT00000008620	3'UTR	3410	-	-	-	-	rs322813307	TMC7	HGNC	23000	F1RP99	UPI00025DF09A	-	17/17	-
3:45094200	A	ENSSSCG00000026202	ENSSSCT00000026543	missense	2020	2020	674	R/W	Cgg/Tgg	rs322413506	-	-	-	I3L9D6	UPI00025DF661	tol(0.13)	6/7	-
3:45183485	A	ENSSSCG00000008087	ENSSSCT00000008860	missense	440	440	147	T/K	aCa/aAa	-	-	-	-	F1SUB9	UPI00025DF6E2	tol(0.09)	4/6	-
3:46313337	G	ENSSSCG00000008105	ENSSSCT00000008879	missense	4640	4606	1536	R/G	Cgg/Ggg	rs345263769	-	-	-	F1SU85	UPI00025E0079	tol(0.73)	38/42	-
3:48488144	T	ENSSSCG00000008121	ENSSSCT00000008897	missense	1964	1964	655	F/Y	tTc/tAc	rs325075532	GPAT2	HGNC	27168	F1SU50	UPI00025E0407	tol(0.13)	18/20	-
3:48488144	T	ENSSSCG00000008121	ENSSSCT00000032292	missense	2195	2195	732	F/Y	tTc/tAc	rs325075532	GPAT2	HGNC	27168	I3LP96	UPI0002105481	tol(0.39)	19/21	-
3:48488145	G	ENSSSCG00000008121	ENSSSCT00000008897	missense	1963	1963	655	F/L	Ttc/Ctc	rs331000241	GPAT2	HGNC	27168	F1SU50	UPI00025E0407	deleterious(0.01)	18/20	-
3:48488145	G	ENSSSCG00000008121	ENSSSCT00000032292	missense	2194	2194	732	F/L	Ttc/Ctc	rs331000241	GPAT2	HGNC	27168	I3LP96	UPI0002105481	deleterious(0.02)	19/21	-
3:75027583	T	ENSSSCG00000008314	ENSSSCT00000009105	3'UTR	1665	-	-	-	-	rs329252845	ATP6V1B1	HGNC	853	F1SLE5	UPI00025E1662	-	12/12	-
3:75027636	G	ENSSSCG00000008314	ENSSSCT00000009105	3'UTR	1718	-	-	-	-	rs340146516	ATP6V1B1	HGNC	853	F1SLE5	UPI00025E1662	-	12/12	-
3:75029260	T	ENSSSCG00000008314	ENSSSCT00000009105	3'UTR	3342	-	-	-	-	rs326994041	ATP6V1B1	HGNC	853	F1SLE5	UPI00025E1662	-	12/12	-
3:88261076	A	ENSSSCG00000028692	ENSSSCT00000026975	missense	752	752	251	R/L	cGt/cTt	rs341025593	LLGL1	HGNC	6628	I3LRY3	UPI00025DFDA3	tol(0.54)	4/4	-
3:118976525	C	ENSSSCG00000027409	ENSSSCT00000029034	3'UTR	3322	-	-	-	-	rs346253177	TCF23	HGNC	18602	F1SED6	UPI0001C96E34	-	3/3	-
3:118976546	C	ENSSSCG00000027409	ENSSSCT00000029034	3'UTR	3301	-	-	-	-	rs325386201	TCF23	HGNC	18602	F1SED6	UPI0001C96E34	-	3/3	-
3:121200467	G	ENSSSCG00000008579	ENSSSCT00000009388	3'UTR	1609	-	-	-	-	rs320022170	CENPO	HGNC	28152	F1SDL1	UPI0001C96E96	-	7/7	-
3:121200467	G	ENSSSCG00000008578	ENSSSCT00000009387	3'UTR	4718	-	-	-	-	rs320022170	ADCY3	HGNC	234	F1SDL2	UPI0001C96E95	-	22/22	-
3:121200478	G	ENSSSCG00000008579	ENSSSCT00000009388	3'UTR	1598	-	-	-	-	rs329571664	CENPO	HGNC	28152	F1SDL1	UPI0001C96E96	-	7/7	-
3:121200478	G	ENSSSCG00000008578	ENSSSCT00000009387	3'UTR	4729	-	-	-	-	rs329571664	ADCY3	HGNC	234	F1SDL2	UPI0001C96E95	-	22/22	-
3:121870156	A	ENSSSCG00000008586	ENSSSCT00000009396	3'UTR	1108	-	-	-	-	rs344669802	TP53I3	HGNC	19373	F1SDJ3	UPI0001C96EE7	-	5/5	-
3:128141914	G	ENSSSCG00000023060	ENSSSCT00000029392	missense	351	351	117	W/C	tgG/tgC	rs323581520	-	-	-	I3LL75	UPI00025DF666	deleterious(0)	1/6	-
3:130417968	C	ENSSSCG00000027247	ENSSSCT00000022524	missense	212	204	68	C/W	tgC/tgG	rs341854935	-	-	-	I3LMU8	UPI00025DFB6D	tol(1)	3/15	-
3:130423337	A	ENSSSCG00000027247	ENSSSCT00000022524	missense	106	98	33	P/L	cCc/cTc	rs328031012	-	-	-	I3LMU8	UPI00025DFB6D	tol(0.17)	1/15	-
3:130423395	A	ENSSSCG00000027247	ENSSSCT00000022524	missense	48	40	14	T/S	Aca/Tca	rs345860614	-	-	-	I3LMU8	UPI00025DFB6D	tol(0.69)	1/15	-
4:11112938	A	ENSSSCG00000005962	ENSSSCT00000006545	missense	1793	1259	420	R/H	cGc/cAc	-	GSDMC	HGNC	7151	F1RRS0	UPI00025E0BB5	deleterious(0)	12/12	-
4:38262416	A	ENSSSCG00000006059	ENSSSCT00000036111	3'UTR	1004	-	-	-	-	-	NCALD	HGNC	7655	F2Z561,K7GMM2	UPI0000004090	-	4/4	-
4:38262416	A	ENSSSCG00000006059	ENSSSCT00000006648	3'UTR	1066	-	-	-	-	-	NCALD	HGNC	7655	F2Z561,K7GMM2	UPI0000004090	-	6/6	-
4:38262416	A	ENSSSCG00000006059	ENSSSCT00000035357	3'UTR	1056	-	-	-	-	-	NCALD	HGNC	7655	F2Z561,K7GMM2	UPI0000004090	-	6/6	-
4:38262416	A	ENSSSCG00000006059	ENSSSCT00000033437	3'UTR	1073	-	-	-	-	-	NCALD	HGNC	7655	F2Z561,K7GMM2	UPI0000004090	-	4/4	-

4:38262416	A	ENSSSCG00000006059	ENSSSCT00000033719	3'UTR	1021	-	-	-	-	-	NCALD	HGNC	7655	F2Z561,K7GMM2	UPI0000004090	-	5/5	-
4:38262416	A	ENSSSCG00000006059	ENSSSCT00000033597	3'UTR	1015	-	-	-	-	-	NCALD	HGNC	7655	F2Z561,K7GMM2	UPI0000004090	-	4/4	-
4:38262416	A	ENSSSCG00000006059	ENSSSCT00000035298	3'UTR	969	-	-	-	-	-	NCALD	HGNC	7655	F2Z561,K7GMM2	UPI0000004090	-	5/5	-
4:87370244	G	ENSSSCG00000006276	ENSSSCT00000006875	3'UTR	1191	-	-	-	-	rs80835237	CEBPD	HGNC	1835	F1RSE5	UPI00017F01DE	-	1/1	-
4:87370244	G	ENSSSCG00000006277	ENSSSCT00000006876	3'UTR	3012	-	-	-	-	rs80835237	SPIDR	HGNC	28971	F1RSE4	UPI00025E0E91	-	16/16	-
4:98833523	G	ENSSSCG00000006398	ENSSSCT00000022301	5'UTR	147	-	-	-	-	rs326117550	SLAMF8	HGNC	21391	I3LD70	UPI00028F44C0	-	1/6	-
4:98833523	G	ENSSSCG00000006398	ENSSSCT00000007011	5'UTR	147	-	-	-	-	rs326117550	SLAMF8	HGNC	21391	F1RJ81	UPI0001C966F3	-	1/6	-
4:98833523	G	ENSSSCG00000006398	ENSSSCT00000035784	5'UTR	587	-	-	-	-	rs326117550	SLAMF8	HGNC	21391	F1RJ81,K7GRB7	UPI0001E86E89	-	1/6	-
4:98833526	C	ENSSSCG00000006398	ENSSSCT00000022301	5'UTR	144	-	-	-	-	rs335687032	SLAMF8	HGNC	21391	I3LD70	UPI00028F44C0	-	1/6	-
4:98833526	C	ENSSSCG00000006398	ENSSSCT00000007011	5'UTR	144	-	-	-	-	rs335687032	SLAMF8	HGNC	21391	F1RJ81	UPI0001C966F3	-	1/6	-
4:98833526	C	ENSSSCG00000006398	ENSSSCT00000035784	5'UTR	584	-	-	-	-	rs335687032	SLAMF8	HGNC	21391	F1RJ81,K7GRB7	UPI0001E86E89	-	1/6	-
4:98833535	G	ENSSSCG00000006398	ENSSSCT00000022301	5'UTR	135	-	-	-	-	rs318889431	SLAMF8	HGNC	21391	I3LD70	UPI00028F44C0	-	1/6	-
4:98833535	G	ENSSSCG00000006398	ENSSSCT00000007011	5'UTR	135	-	-	-	-	rs318889431	SLAMF8	HGNC	21391	F1RJ81	UPI0001C966F3	-	1/6	-
4:98833535	G	ENSSSCG00000006398	ENSSSCT00000035784	5'UTR	575	-	-	-	-	rs318889431	SLAMF8	HGNC	21391	F1RJ81,K7GRB7	UPI0001E86E89	-	1/6	-
4:99148865	T	ENSSSCG00000021859	ENSSSCT00000007023	missense	610	610	204	V/I	Gtc/Atc	rs332302417	-	-	-	F1RJ49	UPI00025DF822	tol(0.91)	2/2	-
4:101824581	G	ENSSSCG00000006465	ENSSSCT00000007082	missense	2317	2317	773	K/E	Aag/Gag	rs342268344	INSRR	HGNC	6093	F1RHJ5	UPI00025E01E6	tol(1)	12/22	-
4:101937500	G	ENSSSCG00000006470	ENSSSCT00000007087	5'UTR	165	-	-	-	-	rs325201309	RRNAD1	HGNC	24273	F1RHJ0	UPI00017F0BFA	-	1/8	-
4:116033868	C	ENSSSCG00000006752	ENSSSCT00000007397	3'UTR	3583	-	-	-	-	rs81220452	CSDE1	HGNC	29905	F1SBS1	UPI00025E15B2	-	20/20	-
4:116874075	T	ENSSSCG00000006762	ENSSSCT00000007408	3'UTR	3606	-	-	-	-	rs325582737	AP4B1	HGNC	572	F1SBR2	UPI0001C96651	-	11/11	-
4:136410091	G	ENSSSCG00000006903	ENSSSCT00000007559	missense	930	878	293	L/S	tTa/tCa	rs318993943	RPAP2	HGNC	25791	F1S4F9	UPI0001C96858	tol(0.54)	8/13	-
5:7646046	T	ENSSSCG00000000125	ENSSSCT00000000131	5'UTR	176	-	-	-	-	-	LGALS1	HGNC	6561	F1SKM9	UPI0001C95092	-	1/4	-
5:8916992	T	ENSSSCG00000000142	ENSSSCT00000000148	5'UTR	708	-	-	-	-	rs344189063	FOXRED2	HGNC	26264	F1SKJ4	UPI0001C950FE	-	1/8	-
5:13127879	T	ENSSSCG00000000164	ENSSSCT00000000173	5'UTR	98	-	-	-	-	rs332025706	CRY1	HGNC	2384	F1SPQ5	UPI0001C9513B	-	1/12	-
5:13127890	C	ENSSSCG00000000164	ENSSSCT00000000173	5'UTR	109	-	-	-	-	rs341046373	CRY1	HGNC	2384	F1SPQ5	UPI0001C9513B	-	1/12	-
5:13127921	G	ENSSSCG00000000164	ENSSSCT00000000173	5'UTR	140	-	-	-	-	rs322352916	CRY1	HGNC	2384	F1SPQ5	UPI0001C9513B	-	1/12	-
5:13127952	A	ENSSSCG00000000164	ENSSSCT00000000173	5'UTR	171	-	-	-	-	rs335331524	CRY1	HGNC	2384	F1SPQ5	UPI0001C9513B	-	1/12	-
5:15931750	C	ENSSSCG00000023744	ENSSSCT00000024572	3'UTR	2372	-	-	-	-	rs327777866	-	-	-	I3LBQ7	UPI00025E12D1	-	4/4	-
5:15931796	T	ENSSSCG00000023744	ENSSSCT00000024572	3'UTR	2326	-	-	-	-	rs344987989	-	-	-	I3LBQ7	UPI00025E12D1	-	4/4	-

5:24125051	C	ENSSSCG00000000415	ENSSSCT00000000449	missense	292	292	98	S/G	Agc/Ggc	rs346262620	SDR9C7	HGNC	29958	F1SL80	UPI00021058C3	tol(1)	1/4	-
5:66039732	T	ENSSSCG00000025680	ENSSSCT00000031826	missense	842	20	7	R/H	cGc/cAc	-	-	-	-	I3LNX9	UPI00025E1955	tol(0.32)	3/12	-
5:66039732	T	ENSSSCG00000025680	ENSSSCT00000029996	missense	848	830	277	R/H	cGc/cAc	-	-	-	-	I3LLY1	UPI00025E1956	tol(1)	5/9	-
5:66649964	T	ENSSSCG00000000704	ENSSSCT00000000766	3'UTR	2652	-	-	-	-	rs339569618	TAPBPL	HGNC	30683	F1SL31	UPI00025DFE52	-	7/7	-
5:71682645	C	ENSSSCG00000000763	ENSSSCT00000000827	3'UTR	1455	-	-	-	-	rs330552835	-	-	-	F1SHR7	UPI00025DFB6C	-	11/11	-
5:84049077	C	ENSSSCG00000000847	ENSSSCT00000000924	3'UTR	2181	-	-	-	-	rs334513521	TDG	HGNC	11700	F1SRK8	UPI0001E8707A	-	10/10	-
5:84049137	T	ENSSSCG00000000847	ENSSSCT00000000924	3'UTR	2121	-	-	-	-	rs325273087	TDG	HGNC	11700	F1SRK8	UPI0001E8707A	-	10/10	-
5:97777634	C	ENSSSCG00000000921	ENSSSCT00000001005	3'UTR	3032	-	-	-	-	rs330554244	-	-	-	F1SPZ5	UPI0001C953BF	-	8/8	-
5:110348841	C	ENSSSCG00000000959	ENSSSCT00000001048	3'UTR	1980	-	-	-	-	rs342717391	-	-	-	F1RYI1	UPI00025E13FE	-	9/9	-
6:4665101	T	ENSSSCG00000002671	ENSSSCT00000002962	3'UTR	3409	-	-	-	-	rs341975209	ATP2C2	HGNC	29103	F1S5U9	UPI0001E872AA	-	27/27	-
6:4665104	G	ENSSSCG00000002671	ENSSSCT00000002962	3'UTR	3406	-	-	-	-	rs322907100	ATP2C2	HGNC	29103	F1S5U9	UPI0001E872AA	-	27/27	-
6:4665105	A	ENSSSCG00000002671	ENSSSCT00000002962	3'UTR	3405	-	-	-	-	-	ATP2C2	HGNC	29103	F1S5U9	UPI0001E872AA	-	27/27	-
6:4665137	T	ENSSSCG00000002671	ENSSSCT00000002962	missense	3373	2817	939	N/K	aaC/aaA	rs346337353	ATP2C2	HGNC	29103	F1S5U9	UPI0001E872AA	tol(1)	27/27	-
6:31827020	C	ENSSSCG00000023391	ENSSSCT00000031189	5'UTR	151	-	-	-	-	rs343752257	-	-	-	I3LRA1	UPI00025DFCC1	-	1/3	-
6:45952385	C	ENSSSCG00000003046	ENSSSCT00000003383	missense	299	163	55	S/P	Tcg/Ccg	-	RABAC1	HGNC	9794	F1RMX2	UPI0001C9586D	tol(0.09)	2/5	-
6:50215189	A	ENSSSCG00000003172	ENSSSCT00000003522	missense	554	476	159	R/H	cGc/cAc	rs340159490	CCDC155	HGNC	26520	F1RHZ5	UPI0001C95AA1	tol(0.15)	5/18	-
6:51664668	G	ENSSSCG00000003228	ENSSSCT00000003587	missense	2024	2024	675	L/P	cTg/cCg	-	VSIG10L	HGNC	27111	F1RP94	UPI00025E13EE	deleterious(0)	7/9	-
6:52015820	G	ENSSSCG00000003243	ENSSSCT00000003603	missense	436	250	84	K/E	Aaa/Gaa	rs343218645	-	-	-	F1RNL1	UPI00025DF33C	tol(1)	4/11	-
6:65727909	T	ENSSSCG00000003423	ENSSSCT00000003797	3'UTR	2643	-	-	-	-	-	DRAXIN	HGNC	25054	F1RF90	UPI0001C95E27	-	7/7	-
6:65727909	T	ENSSSCG00000003423	ENSSSCT00000003798	3'UTR	2634	-	-	-	-	-	DRAXIN	HGNC	25054	F1RF89	UPI0001E87596	-	6/6	-
6:65822920	C	ENSSSCG00000003429	ENSSSCT00000003807	3'UTR	3037	-	-	-	-	rs340847061	CLCN6	HGNC	2024	F1RF80	UPI0001C95E31	-	23/23	-
6:65823200	A	ENSSSCG00000003429	ENSSSCT00000003807	3'UTR	3317	-	-	-	-	rs332133750	CLCN6	HGNC	2024	F1RF80	UPI0001C95E31	-	23/23	-
6:65823205	G	ENSSSCG00000003429	ENSSSCT00000003807	3'UTR	3322	-	-	-	-	rs344042775	CLCN6	HGNC	2024	F1RF80	UPI0001C95E31	-	23/23	-
6:65823209	A	ENSSSCG00000003429	ENSSSCT00000003807	3'UTR	3326	-	-	-	-	rs327799109	CLCN6	HGNC	2024	F1RF80	UPI0001C95E31	-	23/23	-
6:65830603	G	ENSSSCG00000003430	ENSSSCT00000003808	missense	159	149	50	M/T	aTg/aCg	rs335475074	NPPA	HGNC	7939	F1RF79	UPI00025E00FC	deleterious(0.01)	2/2	-
6:81864340	G	ENSSSCG00000003601	ENSSSCT00000003996	3'UTR	1615	-	-	-	-	rs340567561	HCRTR1	HGNC	4848	F1SVA1	UPI0001C95B20	-	8/8	-
6:86735510	G	ENSSSCG00000003643	ENSSSCT00000004039	missense	1061	842	281	R/P	cGa/cCa	-	SF3A3	HGNC	10767	F1SV40	UPI00004A5705	deleterious(0)	11/17	-
6:119372599	T	ENSSSCG00000003756	ENSSSCT00000004158	3'UTR	1396	-	-	-	-	rs45435478	EDG7	Uniprot_gn	-	C5G5X7	UPI0001A7B400	-	3/3	-

6:135380929	G	ENSSSCG00000003806	ENSSSCT00000004212	3'UTR	969	-	-	-	-	rs334140732	LEPROT	HGNC	29477	J9JIK7	UPI00025E1692	-	4/4	-
6:146104670	C	ENSSSCG00000028506	ENSSSCT00000026980	5'UTR	14	-	-	-	-	rs337560399	YIPF1	HGNC	25231	I3L692	UPI0002105DA9	-	1/10	-
6:152572491	C	ENSSSCG00000003902	ENSSSCT00000004317	3'UTR	2614	-	-	-	-	rs330737439	NSUN4	HGNC	31802	F1S3W1	UPI0001C95FB3	-	6/6	-
6:155538713	C	ENSSSCG00000003955	ENSSSCT00000004374	missense	694	694	232	C/R	Tgc/Cgc	rs326370902	-	-	-	F1SFB6	UPI00025DE800	tol(1)	1/1	-
7:24748234	G	ENSSSCG00000001231	ENSSSCT00000036412	3'UTR	1338	-	-	-	-	-	SLA-1	Clone_based_vega_gene	-	K7GNK3	UPI00028F4960	-	7/7	-
7:24748234	G	ENSSSCG00000001231	ENSSSCT00000001334	3'UTR	1404	-	-	-	-	-	SLA-1	Clone_based_vega_gene	-	F1RZ59	UPI00017EFC69	-	8/8	-
7:24748271	C	ENSSSCG00000001231	ENSSSCT00000036412	3'UTR	1301	-	-	-	-	-	SLA-1	Clone_based_vega_gene	-	K7GNK3	UPI00028F4960	-	7/7	-
7:24748271	C	ENSSSCG00000001231	ENSSSCT00000001334	3'UTR	1367	-	-	-	-	-	SLA-1	Clone_based_vega_gene	-	F1RZ59	UPI00017EFC69	-	8/8	-
7:24748277	A	ENSSSCG00000001231	ENSSSCT00000036412	3'UTR	1295	-	-	-	-	-	SLA-1	Clone_based_vega_gene	-	K7GNK3	UPI00028F4960	-	7/7	-
7:24748277	A	ENSSSCG00000001231	ENSSSCT00000001334	3'UTR	1361	-	-	-	-	-	SLA-1	Clone_based_vega_gene	-	F1RZ59	UPI00017EFC69	-	8/8	-
7:25969489	C	ENSSSCG00000001294	ENSSSCT00000001409	missense	896	896	299	S/C	tCc/tGc	rs333621807	-	-	-	F1RXL5	UPI0001C954EA	deleterious(0)	1/1	-
7:26432546	T	ENSSSCG00000001340	ENSSSCT00000001456	missense	383	383	128	R/H	cGt/cAt	rs343678707	-	-	-	F1RUD1	UPI0001C954D3	tol(1)	2/5	-
7:26432757	A	ENSSSCG00000001340	ENSSSCT00000001456	missense	173	173	58	P/L	cCc/cTc	-	-	-	-	F1RUD1	UPI0001C954D3	deleterious(0)	1/5	-
7:26432788	C	ENSSSCG00000001340	ENSSSCT00000001456	missense	142	142	48	L/V	Tta/Gta	rs344318735	-	-	-	F1RUD1	UPI0001C954D3	tol(0.93)	1/5	-
7:26432797	T	ENSSSCG00000001340	ENSSSCT00000001456	missense	133	133	45	V/I	Gtc/Atc	rs322757273	-	-	-	F1RUD1	UPI0001C954D3	tol(1)	1/5	-
7:33644483	C	ENSSSCG00000001499	ENSSSCT00000001671	missense	13397	13397	4466	M/T	aTg/aCg	rs322822803	DST	HGNC	1090	F1RZU8	UPI00025E19A2	-	51/97	-
7:39242923	T	ENSSSCG00000027778	ENSSSCT00000029888	3'UTR	3344	-	-	-	-	rs322735737	GLO1	HGNC	4323	I3LDM7	UPI00017EFBAB	-	6/6	-
7:39242950	C	ENSSSCG00000027778	ENSSSCT00000029888	3'UTR	3317	-	-	-	-	rs333664915	GLO1	HGNC	4323	I3LDM7	UPI00017EFBAB	-	6/6	-
7:55707198	C	ENSSSCG00000001788	ENSSSCT00000002003	3'UTR	977	-	-	-	-	rs332018296	STARD5	HGNC	18065	F1RID6	UPI00017EFE4D	-	6/6	-
7:64039109	G	ENSSSCG00000001912	ENSSSCT00000002141	missense	1325	1321	441	A/P	Gca/Cca	-	PML	HGNC	9113	F1SID1	UPI0001C95A30	deleterious(0.02)	5/9	-
7:80271028	C	ENSSSCG00000001991	ENSSSCT00000002229	missense	719	473	158	L/P	cTc/cCc	-	DHRS1	HGNC	16445	F1SGQ0	UPI0001C9562C	deleterious(0.01)	5/9	-
7:80434967	G	ENSSSCG00000002005	ENSSSCT00000002243	3'UTR	2837	-	-	-	-	-	EMC9	HGNC	20273	F1SGM2	UPI00017F09D3	-	5/5	-
7:86420677	G	ENSSSCG00000002246	ENSSSCT00000002501	missense	1563	1563	521	E/D	gaG/gaC	rs322350459	CHRM5	HGNC	1954	F1SCL4	UPI0001C958DE	tol(0.25)	1/1	-
7:103573259	A	ENSSSCG00000002366	ENSSSCT00000002630	3'UTR	1333	-	-	-	-	rs81492008	NPC2	Uniprot_gn	-	-	UPI00001303DB	-	4/4	-
7:119945119	C	ENSSSCG00000002443	ENSSSCT00000002713	3'UTR	4591	-	-	-	-	rs323422699	TC2N	HGNC	19859	F1SD88	UPI00017F0336	-	11/11	-
7:120278258	A	ENSSSCG00000002447	ENSSSCT00000002717	5'UTR	75	-	-	-	-	rs327788819	-	-	-	F1SD85	UPI00025DFB8C	-	1/9	-
7:122968448	G	ENSSSCG00000030371	ENSSSCT00000002764	missense	1214	1066	356	K/E	Aaa/Gaa	-	-	-	-	F1SCC6	UPI00017F0640	tol(0.77)	4/5	-
7:124707839	A	ENSSSCG00000027030	ENSSSCT00000022656	3'UTR	2637	-	-	-	-	rs327987366	BDKRB2	Clone_based_vega_gene	-	I3LMT9	UPI00025DF5A1	-	2/2	-

7:124708086	A	ENSSSCG00000027030	ENSSSCT00000022656	3'UTR	2884	-	-	-	-	rs332021578	BDKRB2	Clone_based_vega_gene	-	I3LMT9	UPI00025DF5A1	-	2/2	-
7:124708091	T	ENSSSCG00000027030	ENSSSCT00000022656	3'UTR	2889	-	-	-	-	-	BDKRB2	Clone_based_vega_gene	-	I3LMT9	UPI00025DF5A1	-	2/2	-
7:124708111	C	ENSSSCG00000027030	ENSSSCT00000022656	3'UTR	2909	-	-	-	-	-	BDKRB2	Clone_based_vega_gene	-	I3LMT9	UPI00025DF5A1	-	2/2	-
7:124749090	G	ENSSSCG00000002501	ENSSSCT00000002779	missense	1039	1039	347	F/V	Ttc/Gtc	rs344088349	BDKRB1	HGNC	1029	F1SAR0	UPI0001C95741	tol(0.07)	1/1	-
7:129277651	G	ENSSSCG00000024055	ENSSSCT00000028166	3'UTR	2745	-	-	-	-	rs336268948	-	-	-	I3LRS6	UPI00025E0009	-	10/10	-
7:133511931	T	ENSSSCG00000028324	ENSSSCT00000002883	stop_gained	298	298	100	Q/*	Cag/Tag	rs330852659	-	-	-	F1S7G7	UPI00025DE7B1	-	1/1	-
8:31034153	C	ENSSSCG00000008774	ENSSSCT00000009599	missense	661	661	221	M/L	Atg/Ctg	-	TBC1D1	HGNC	11578	F1S4K5	UPI00025E0DD2	tol(0.15)	4/8	-
8:92675076	A	ENSSSCG00000009060	ENSSSCT00000009924	5'UTR	1037	-	-	-	-	rs336076935	MAML3	HGNC	16272	F1RRE2	UPI00025DFC26	-	1/4	-
8:93405750	T	ENSSSCG00000009063	ENSSSCT00000009928	missense	40	21	7	L/F	ttG/ttT	rs81402316	NDUFC1	Uniprot_gn	-	F1RRC9	UPI0001C96A0A	deleterious(0.03)	1/4	-
8:142136650	T	ENSSSCG00000009228	ENSSSCT00000010107	3'UTR	4405	-	-	-	-	rs345301446	MAPK10	HGNC	6872	F1RW16,K7GP94	UPI0001C96F0A	-	12/12	-
9:481427	A	ENSSSCG00000014575	ENSSSCT00000015925	missense	859	859	287	D/N	Gat/Aat	rs81420553	SCUBE2	HGNC	30425	F1RNJ7	UPI00025DF458	deleterious(0)	11/26	-
9:481475	G	ENSSSCG00000014575	ENSSSCT00000015925	missense	907	907	303	I/V	Atc/Gtc	rs327670752	SCUBE2	HGNC	30425	F1RNJ7	UPI00025DF458	tol(1)	11/26	-
9:519375	T	ENSSSCG00000014575	ENSSSCT00000015925	3'UTR	3172	-	-	-	-	-	SCUBE2	HGNC	30425	F1RNJ7	UPI00025DF458	-	26/26	-
9:4895896	A	ENSSSCG00000014690	ENSSSCT00000016041	missense	737	737	246	S/N	aGt/aAt	rs319089708	UBQLNL	HGNC	28294	F1RJB2	UPI0001C98227	tol(1)	1/1	-
9:5185090	T	ENSSSCG00000024353	ENSSSCT00000032278	missense	931	931	311	A/S	Gct/Tct	rs322302302	-	-	-	I3LUY6	UPI00025DE8A2	tol(0.42)	1/1	-
9:7136936	G	ENSSSCG00000014794	ENSSSCT00000016145	missense	1333	1333	445	K/E	Aag/Gag	-	NUP98	HGNC	8068	F1SUZ2	UPI00025E0E1C	deleterious(0)	10/30	-
9:8563008	G	ENSSSCG00000014822	ENSSSCT00000016174	missense	1502	1502	501	A/G	gCg/gGg	rs327104494	ARHGEF17	HGNC	21726	F1SUT6	UPI00025E17ED	tol(1)	1/21	-
9:44445094	C	ENSSSCG00000015035	ENSSSCT00000026093	3'UTR	659	-	-	-	-	rs334543928	C11orf52	HGNC	30531	I3LI29	UPI00025DF5E3	-	5/5	-
9:44445094	C	ENSSSCG00000015035	ENSSSCT00000016400	3'UTR	662	-	-	-	-	rs334543928	C11orf52	HGNC	30531	F1SMA8	UPI00025DF5E2	-	4/4	-
9:50378829	T	ENSSSCG00000015085	ENSSSCT00000016451	3'UTR	3227	-	-	-	-	rs341177467	IL10RA	Clone_based_vega_gene	-	F1SAM5	UPI00025DFE93	-	6/6	-
9:50378829	T	ENSSSCG00000015085	ENSSSCT00000036681	3'UTR	3311	-	-	-	-	rs341177467	IL10RA	Clone_based_vega_gene	-	K7GSL4	UPI0001E87F5C	-	7/7	-
9:50757145	A	ENSSSCG00000023777	ENSSSCT00000032072	3'UTR	2260	-	-	-	-	rs321269671	-	-	-	Q7YS30,F1SAK6	UPI0001C98169	-	3/3	-
9:50867236	C	ENSSSCG00000029395	ENSSSCT00000025141	missense	877	877	293	R/G	Cgg/Ggg	-	TMEM25	HGNC	25890	I3LNP1	UPI00025E07D1	deleterious(0)	6/8	-
9:51639118	T	ENSSSCG00000015122	ENSSSCT00000035739	3'UTR	1933	-	-	-	-	rs334857517	CD90	Uniprot_gn	-	K7GL43	UPI00028F4A79	-	2/2	-
9:51639118	T	ENSSSCG00000015122	ENSSSCT00000016488	3'UTR	2104	-	-	-	-	rs334857517	CD90	Uniprot_gn	-	B9ZSM8	UPI000195C788	-	4/4	-
9:73855980	G	ENSSSCG00000015654	ENSSSCT00000033310	missense	255	107	36	Y/C	tAt/tGt	rs81419973	IL20	HGNC	6002	F1SEZ1,A0FH92,K7GPP8	UPI00028F4615	tol(0.21)	1/6	-
9:73856015	C	ENSSSCG00000015654	ENSSSCT00000033310	missense	290	142	48	F/L	Ttt/Ctt	rs335456270	IL20	HGNC	6002	F1SEZ1,A0FH92,K7GPP8	UPI00028F4615	tol(0.22)	1/6	-
9:74477548	G	ENSSSCG00000028674	ENSSSCT00000028859	missense	2507	2323	775	T/A	Acg/Gcg	rs342482622	CR2	HGNC	2336	I3LK97	UPI00025DFCBB	tol(0.11)	13/27	-

9:85232767	C	ENSSSCG00000015342	ENSSSCT00000016715	missense	2174	2174	725	Q/R	cAa/cGa	rs345144554	COL28A1	HGNC	22442	F1SF83	UPI00025E1980	tol(0.22)	28/34	-
9:119580777	G	ENSSSCG00000023584	ENSSSCT00000022370	3'UTR	798	-	-	-	-	rs335008877	-	-	-	I3LPP0	UPI00025E0CBC	-	5/5	-
9:125530018	C	ENSSSCG00000015480	ENSSSCT00000016863	3'UTR	9239	-	-	-	-	-	PRRC2C	HGNC	24903	F1S7T1	UPI00025E14C5	-	35/35	-
9:125530036	G	ENSSSCG00000015480	ENSSSCT00000016863	3'UTR	9257	-	-	-	-	-	PRRC2C	HGNC	24903	F1S7T1	UPI00025E14C5	-	35/35	-
9:125530084	C	ENSSSCG00000015480	ENSSSCT00000016863	3'UTR	9305	-	-	-	-	-	PRRC2C	HGNC	24903	F1S7T1	UPI00025E14C5	-	35/35	-
9:125530689	A	ENSSSCG00000015480	ENSSSCT00000016863	3'UTR	9910	-	-	-	-	rs81415979	PRRC2C	HGNC	24903	F1S7T1	UPI00025E14C5	-	35/35	-
9:137190180	C	ENSSSCG00000015561	ENSSSCT00000016951	missense	337	337	113	N/D	Aat/Gat	rs344816883	APOBEC4	HGNC	32152	F1S657	UPI0001C981E4	tol(1)	1/1	-
9:137190185	C	ENSSSCG00000015561	ENSSSCT00000016951	missense	332	332	111	N/S	aAc/aGc	rs328587334	APOBEC4	HGNC	32152	F1S657	UPI0001C981E4	tol(0.57)	1/1	-
9:144598302	T	ENSSSCG00000015604	ENSSSCT00000016997	missense	1239	1129	377	T/S	Acc/Tcc	rs329352790	NEK2	HGNC	7745	Q9GKX1,F1S2U5	UPI0001C98450	tol(0.44)	8/8	-
9:146400404	G	ENSSSCG00000015611	ENSSSCT00000017005	5'UTR	601	-	-	-	-	rs80930679	-	-	-	F1SF51	UPI00025DFD8F	-	1/7	-
9:146400438	T	ENSSSCG00000015611	ENSSSCT00000017005	5'UTR	567	-	-	-	-	rs321786623	-	-	-	F1SF51	UPI00025DFD8F	-	1/7	-
9:146400445	G	ENSSSCG00000015611	ENSSSCT00000017005	5'UTR	560	-	-	-	-	rs332860110	-	-	-	F1SF51	UPI00025DFD8F	-	1/7	-
9:146400464	G	ENSSSCG00000015611	ENSSSCT00000017005	5'UTR	541	-	-	-	-	rs327609517	-	-	-	F1SF51	UPI00025DFD8F	-	1/7	-
9:146400492	G	ENSSSCG00000015611	ENSSSCT00000017005	5'UTR	513	-	-	-	-	rs336488179	-	-	-	F1SF51	UPI00025DFD8F	-	1/7	-
9:146400619	C	ENSSSCG00000015611	ENSSSCT00000017005	5'UTR	386	-	-	-	-	rs337719270	-	-	-	F1SF51	UPI00025DFD8F	-	1/7	-
10:17750738	A	ENSSSCG00000021132	ENSSSCT00000027974	3'UTR	1857	-	-	-	-	rs331466003	-	-	-	I3L6B1	UPI00025DEC5E	-	2/2	-
10:17750914	G	ENSSSCG00000021132	ENSSSCT00000027974	3'UTR	1681	-	-	-	-	rs321969127	-	-	-	I3L6B1	UPI00025DEC5E	-	2/2	-
10:22259333	T	ENSSSCG00000030286	ENSSSCT00000028202	missense	425	176	59	A/E	gCg/gAg	-	-	-	-	I3LHR5	UPI00025DEE7F	tol(1)	2/3	-
10:35065461	C	ENSSSCG00000026291	ENSSSCT00000030955	3'UTR	2698	-	-	-	-	rs81248415	C9orf64	HGNC	28144	I3LAW0	UPI0002106381	-	4/4	-
10:35065492	C	ENSSSCG00000026291	ENSSSCT00000030955	3'UTR	2667	-	-	-	-	rs81248414	C9orf64	HGNC	28144	I3LAW0	UPI0002106381	-	4/4	-
10:47746778	G	ENSSSCG00000011030	ENSSSCT00000012072	3'UTR	6681	-	-	-	-	rs323462543	CUBN	HGNC	2548	F1RWC3	UPI00025DF7CE	-	39/39	-
10:51524885	C	ENSSSCG00000011050	ENSSSCT00000012094	3'UTR	512	-	-	-	-	rs328412791	MEIG1	HGNC	23429	F1RW82	UPI0000F5DBB3	-	2/2	-
10:51524900	T	ENSSSCG00000011050	ENSSSCT00000012094	3'UTR	497	-	-	-	-	rs341138488	MEIG1	HGNC	23429	F1RW82	UPI0000F5DBB3	-	2/2	-
10:51525021	T	ENSSSCG00000011050	ENSSSCT00000012094	3'UTR	376	-	-	-	-	rs326329076	MEIG1	HGNC	23429	F1RW82	UPI0000F5DBB3	-	2/2	-
10:75708339	A	ENSSSCG00000011162	ENSSSCT00000012224	3'UTR	2563	-	-	-	-	-	LARP4B	HGNC	28987	F1RX46	UPI0002106584	-	17/17	-
10:78823222	G	ENSSSCG00000011176	ENSSSCT00000012239	missense	125	125	42	L/P	cTg/cCg	-	-	-	-	F1RX26	UPI00025E08E7	deleterious(0.04)	1/11	-
10:78823225	A	ENSSSCG00000011176	ENSSSCT00000012239	missense	122	122	41	K/M	aAg/aTg	-	-	-	-	F1RX26	UPI00025E08E7	deleterious(0.04)	1/11	-
11:7830830	A	ENSSSCG00000009335	ENSSSCT00000010230	missense	798	753	251	D/E	gaC/gaA	rs336420078	B3GALTL	HGNC	20207	F1RSS9	UPI00025DF92B	tol(0.51)	9/14	-

11:16080433	A	ENSSSCG00000009378	ENSSSCT00000010275	missense	374	289	97	G/S	Ggt/Agt	rs81215718	CKAP2	HGNC	1990	F1RMD2	UPI0001C96DAD	tol(0.62)	4/9	-
11:20956321	T	ENSSSCG00000009406	ENSSSCT00000010305	missense	1384	1378	460	S/C	Agc/Tgc	rs339873320	HTR2A	HGNC	5293	F1RK09	UPI0001C96ECB	tol(0.1)	3/3	-
11:86854866	G	ENSSSCG00000023633	ENSSSCT00000025705	3'UTR	1798	-	-	-	-	rs321800457	-	-	-	I3L5A2	UPI0001E87F4A	-	2/2	-
12:1439231	C	ENSSSCG00000028593	ENSSSCT00000022504	3'UTR	2793	-	-	-	-	rs325772802	ENTHD2	HGNC	26458	I3L691	UPI0002106 8D9	-	13/13	-
12:14953133	C	ENSSSCG00000017279	ENSSSCT00000018811	3'UTR	4767	-	-	-	-	rs340502271	ERN1	HGNC	3449	F1RSL5	UPI0001C98E1B	-	20/20	-
12:14976306	T	ENSSSCG00000017280	ENSSSCT00000034499	5'UTR	451	-	-	-	-	rs322711991	ICAM2	HGNC	5345	K7GQI6,K7GMR5	UPI00028F48A3	-	3/6	-
12:14976306	T	ENSSSCG00000017280	ENSSSCT00000033230	5'UTR	488	-	-	-	-	rs322711991	ICAM2	HGNC	5345	K7GMR5	UPI00028F4A82	-	2/4	-
12:14976323	T	ENSSSCG00000017280	ENSSSCT00000034499	5'UTR	468	-	-	-	-	rs334696360	ICAM2	HGNC	5345	K7GQI6,K7GMR5	UPI00028F48A3	-	3/6	-
12:14976323	T	ENSSSCG00000017280	ENSSSCT00000033230	5'UTR	505	-	-	-	-	rs334696360	ICAM2	HGNC	5345	K7GMR5	UPI00028F4A82	-	2/4	-
12:15293741	T	ENSSSCG00000017293	ENSSSCT00000018825	3'UTR	2889	-	-	-	-	-	TACO1	HGNC	24316	F1RRW7	UPI0001C98E57	-	5/5	-
12:15293743	C	ENSSSCG00000017293	ENSSSCT00000018825	3'UTR	2887	-	-	-	-	-	TACO1	HGNC	24316	F1RRW7	UPI0001C98E57	-	5/5	-
12:18762511	C	ENSSSCG00000017347	ENSSSCT00000018883	5'UTR	128	-	-	-	-	rs328842864	HIGD1B	HGNC	24318	F1RQZ8	UPI00025E134B	-	1/4	-
12:19500806	A	ENSSSCG00000017367	ENSSSCT00000018906	3'UTR	2057	-	-	-	-	rs338248908	MPP2	HGNC	7220	F1S1J3	UPI00025DF6C1	-	11/11	-
12:25811003	T	ENSSSCG00000023586	ENSSSCT00000030769	3'UTR	1552	-	-	-	-	rs334137615	SPOP	HGNC	11254	I3LA15	UPI00025DF28D	-	7/7	-
12:26401665	A	ENSSSCG00000028776	ENSSSCT00000030597	5'UTR	161	-	-	-	-	rs322093323	HILS1	Clone_based_vega_gene	-	I3LB72	UPI00028F44CE	-	1/1	-
12:26719194	A	ENSSSCG00000017576	ENSSSCT00000019133	missense	1556	1556	519	A/E	gCg/gAg	rs81497321	-	-	-	F1RT74	UPI00025DF847	tol(0.33)	9/12	-
12:26828555	T	ENSSSCG00000017566	ENSSSCT00000019121	missense	689	664	222	L/F	Ctt/Ttt	rs337171306	ACSF2	HGNC	26101	F1RT96	UPI00025DFCC7	tol(0.71)	6/16	-
12:36043165	G	ENSSSCG00000017643	ENSSSCT00000035025	3'UTR	1623	-	-	-	-	rs332413484	SEPT4	HGNC	9165	K7GM78	UPI0002106 9F7	-	12/12	-
12:36043165	G	ENSSSCG00000017643	ENSSSCT00000019205	3'UTR	1722	-	-	-	-	rs332413484	SEPT4	HGNC	9165	F1RRN6	UPI0001E880D2	-	12/12	-
12:38930552	G	ENSSSCG00000026665	ENSSSCT00000025171	missense	550	522	174	I/M	atA/atG	rs81213318	-	-	-	I3LH52	UPI00025DF624	tol(1)	3/4	-
12:41224886	G	ENSSSCG00000017705	ENSSSCT00000034994	3'UTR	1323	-	-	-	-	rs319705920	CCL5	Uniprot_gn	-	K7GMD2	UPI00028F48DE	-	2/2	-
12:46537688	T	ENSSSCG00000024626	ENSSSCT00000032150	3'UTR	2705	-	-	-	-	rs320388237	TNFAIP1	HGNC	11894	I3LM60	UPI00022CEABB	-	7/7	-
12:48622362	G	ENSSSCG00000017799	ENSSSCT00000019371	3'UTR	7545	-	-	-	-	rs338508857	CPD	HGNC	2301	F1RN68	UPI0001C98C66	-	21/21	-
12:49455051	A	ENSSSCG00000017812	ENSSSCT00000019386	3'UTR	4517	-	-	-	-	rs334984943	VPS53	HGNC	25608	F1RHI3	UPI0001C98C74	-	22/22	-
12:49455138	C	ENSSSCG00000017812	ENSSSCT00000019386	3'UTR	4604	-	-	-	-	rs345402897	VPS53	HGNC	25608	F1RHI3	UPI0001C98C74	-	22/22	-
12:51068404	A	ENSSSCG00000023490	ENSSSCT00000025106	missense	847	847	283	V/L	Gta/Tta	rs330421773	-	-	-	I3LI12	UPI0002106854	tol(1)	1/1	-
12:51068467	G	ENSSSCG00000023490	ENSSSCT00000025106	missense	784	784	262	M/L	Atg/Ctg	rs331208499	-	-	-	I3LI12	UPI0002106854	tol(0.75)	1/1	-
12:51068506	G	ENSSSCG00000023490	ENSSSCT00000025106	missense	745	745	249	V/L	Gtg/Ctg	rs343601164	-	-	-	I3LI12	UPI0002106854	tol(1)	1/1	-

12:51576632	C	ENSSSCG00000017864	ENSSSCT00000019440	missense	1511	881	294	V/G	gTa/gGa	rs339590896	SHPK	HGNC	1492	F1RLI9	UPI00025DF253	tol(1)	6/7	-
12:53917090	T	ENSSSCG00000025235	ENSSSCT00000030882	missense	761	404	135	S/N	aGt/aAt	rs343573166	ZFP3	HGNC	12861	F1RFY7	UPI0001C98D23	tol(0.12)	2/2	-
12:53967824	A	ENSSSCG00000017900	ENSSSCT00000019480	3’UTR	5024	-	-	-	-	rs321617679	KIF1C	HGNC	6317	F1RFY6	UPI0001C98D24	-	23/23	-
12:53967828	C	ENSSSCG00000017900	ENSSSCT00000019480	3’UTR	5020	-	-	-	-	rs328226118	KIF1C	HGNC	6317	F1RFY6	UPI0001C98D24	-	23/23	-
12:53967857	A	ENSSSCG00000017900	ENSSSCT00000019480	3’UTR	4991	-	-	-	-	rs337848011	KIF1C	HGNC	6317	F1RFY6	UPI0001C98D24	-	23/23	-
12:53967892	T	ENSSSCG00000017900	ENSSSCT00000019480	3’UTR	4956	-	-	-	-	rs319065388	KIF1C	HGNC	6317	F1RFY6	UPI0001C98D24	-	23/23	-
12:62027976	C	ENSSSCG00000018028	ENSSSCT00000019623	5’UTR	120	-	-	-	-	rs332436320	TEKT3	HGNC	14293	F1SDE8	UPI0001C98E40	-	2/7	-
13:28861858	A	ENSSSCG00000030065	ENSSSCT00000029929	3’UTR	3908	-	-	-	-	rs324851956	ZBTB47	HGNC	26955	I3LMM3	UPI00025E009F	-	6/6	-
13:28861858	A	ENSSSCG00000030065	ENSSSCT00000027090	3’UTR	4868	-	-	-	-	rs324851956	ZBTB47	HGNC	26955	I3LTP9	UPI00025E009E	-	6/6	-
13:28927966	C	ENSSSCG00000011291	ENSSSCT00000012361	5’UTR	23	-	-	-	-	rs339898672	CCDC13	HGNC	26358	F1SRE6	UPI0002106895	-	1/15	-
13:28927987	A	ENSSSCG00000011291	ENSSSCT00000012361	5’UTR	2	-	-	-	-	rs323786711	CCDC13	HGNC	26358	F1SRE6	UPI0002106895	-	1/15	-
13:33023398	T	ENSSSCG00000011326	ENSSSCT00000012401	missense	1811	1666	556	L/F	Ctc/Ttc	rs330276009	PTH1R	Uniprot_gn	-	-	UPI00001329C1	deleterious(0.03)	15/15	-
13:72985315	T	ENSSSCG00000011543	ENSSSCT00000012633	3’UTR	1373	-	-	-	-	rs80861050	LHFPL4	HGNC	29568	F1SR30	UPI000155FA4D	-	3/3	-
13:73202126	T	ENSSSCG00000011555	ENSSSCT00000012645	3’UTR	1413	-	-	-	-	-	RPUSD3	HGNC	28437	F1SQE3	UPI00017EFB54	-	9/9	-
13:77746279	A	ENSSSCG00000011601	ENSSSCT00000012698	5’UTR	117	-	-	-	-	-	-	-	-	F2Z5S4	UPI0000003E6A	-	1/4	-
13:80086689	C	ENSSSCG00000024134	ENSSSCT00000030864	3’UTR	3220	-	-	-	-	rs319356357	MGLL	Uniprot_gn	-	B8XSJ9	UPI00018958FC	-	8/8	-
13:82443831	A	ENSSSCG00000011640	ENSSSCT00000012740	missense	1031	980	327	R/K	aGg/aAg	rs81211478	TF	Uniprot_gn	-	B3CL06	UPI0001754F86	tol(0.34)	8/17	-
13:87895740	C	ENSSSCG00000011661	ENSSSCT00000012766	3’UTR	2768	-	-	-	-	rs330334469	MRPS22	HGNC	14508	F1SL55	UPI0001C9797E	-	9/9	-
13:89298564	C	ENSSSCG00000011666	ENSSSCT00000012771	3’UTR	4485	-	-	-	-	rs343978707	CLSTN2	HGNC	17448	F1SL50	UPI00025E0E82	-	15/15	-
13:89298586	C	ENSSSCG00000011666	ENSSSCT00000012771	3’UTR	4507	-	-	-	-	rs324809353	CLSTN2	HGNC	17448	F1SL50	UPI00025E0E82	-	15/15	-
13:89298646	T	ENSSSCG00000011666	ENSSSCT00000012771	3’UTR	4567	-	-	-	-	rs345812107	CLSTN2	HGNC	17448	F1SL50	UPI00025E0E82	-	15/15	-
13:89298672	T	ENSSSCG00000011666	ENSSSCT00000012771	3’UTR	4593	-	-	-	-	rs327234780	CLSTN2	HGNC	17448	F1SL50	UPI00025E0E82	-	15/15	-
13:103795396	C	ENSSSCG00000028855	ENSSSCT00000025813	missense	735	358	120	K/Q	Aaa/Caa	-	-	-	-	-	UPI00025DF861	tol(0.06)	4/9	-
13:117162212	C	ENSSSCG00000026829	ENSSSCT00000032001	5’UTR	460	-	-	-	-	-	MYNN	HGNC	14955	I3LGC1	UPI0002106A7B	-	1/8	-
13:117162212	C	ENSSSCG00000026829	ENSSSCT00000029120	5’UTR	460	-	-	-	-	-	MYNN	HGNC	14955	I3L5S0	UPI00025DF3B8	-	1/7	-
13:117162214	G	ENSSSCG00000026829	ENSSSCT00000032001	5’UTR	462	-	-	-	-	-	MYNN	HGNC	14955	I3LGC1	UPI0002106A7B	-	1/8	-
13:117162214	G	ENSSSCG00000026829	ENSSSCT00000029120	5’UTR	462	-	-	-	-	-	MYNN	HGNC	14955	I3L5S0	UPI00025DF3B8	-	1/7	-
13:133973467	T	ENSSSCG00000011803	ENSSSCT00000012915	5’UTR	31	-	-	-	-	rs330468296	EIF4A2	Uniprot_gn	-	A6M930	UPI0000000DD2	-	1/1	-

13:141231577	G	ENSSSCG00000011827	ENSSSCT00000012939	missense	1144	1097	366	I/T	aTc/aCc	rs337739868	LSG1	HGNC	25652	F1SFG2	UPI0001C97659	deleterious (0)	8/14	-
13:146237527	C	ENSSSCG00000020935	ENSSSCT00000030879	3'UTR	3080	-	-	-	-	rs333472939	PTPLB	HGNC	9640	I3LDV2	UPI00025DFEE3	-	3/3	-
13:168193207	A	ENSSSCG00000011963	ENSSSCT00000013090	3'UTR	1857	-	-	-	-	rs323890203	-	-	-	F1SKZ8	UPI00025E1331	-	3/3	-
14:138042	T	ENSSSCG00000009578	ENSSSCT00000010499	3'UTR	1891	-	-	-	-	rs338877815	CDK20	Uniprot_gn	-	D3K5N0	UPI00017EFD48	-	8/8	-
14:138069	G	ENSSSCG00000009578	ENSSSCT00000010499	3'UTR	1864	-	-	-	-	rs324441857	CDK20	Uniprot_gn	-	D3K5N0	UPI00017EFD48	-	8/8	-
14:138213	C	ENSSSCG00000009578	ENSSSCT00000010499	3'UTR	1720	-	-	-	-	rs345373422	CDK20	Uniprot_gn	-	D3K5N0	UPI00017EFD48	-	8/8	-
14:138217	T	ENSSSCG00000009578	ENSSSCT00000010499	3'UTR	1716	-	-	-	-	rs327458652	CDK20	Uniprot_gn	-	D3K5N0	UPI00017EFD48	-	8/8	-
14:138256	A	ENSSSCG00000009578	ENSSSCT00000010499	3'UTR	1677	-	-	-	-	rs80991586	CDK20	Uniprot_gn	-	D3K5N0	UPI00017EFD48	-	8/8	-
14:138380	G	ENSSSCG00000009578	ENSSSCT00000010499	3'UTR	1553	-	-	-	-	rs80993802	CDK20	Uniprot_gn	-	D3K5N0	UPI00017EFD48	-	8/8	-
14:138391	G	ENSSSCG00000009578	ENSSSCT00000010499	3'UTR	1542	-	-	-	-	rs319767165	CDK20	Uniprot_gn	-	D3K5N0	UPI00017EFD48	-	8/8	-
14:138411	G	ENSSSCG00000009578	ENSSSCT00000010499	3'UTR	1522	-	-	-	-	rs329344876	CDK20	Uniprot_gn	-	D3K5N0	UPI00017EFD48	-	8/8	-
14:892057	T	ENSSSCG00000009580	ENSSSCT00000010501	5'UTR	294	-	-	-	-	rs333793143	S1PR3	Uniprot_gn	-	F1RN19,B5MBU5	UPI00017EFB24	-	2/2	-
14:892061	C	ENSSSCG00000009580	ENSSSCT00000010501	5'UTR	298	-	-	-	-	rs342106223	S1PR3	Uniprot_gn	-	F1RN19,B5MBU5	UPI00017EFB24	-	2/2	-
14:2853232	T	ENSSSCG00000021967	ENSSSCT00000031967	3'UTR	550	-	-	-	-	rs327928546	SYK	Clone_based_vega_gene	11491	I3LPY9	UPI00025E08AC	-	4/4	-
14:2853318	A	ENSSSCG00000021967	ENSSSCT00000031967	3'UTR	636	-	-	-	-	rs342232766	SYK	Clone_based_vega_gene	11491	I3LPY9	UPI00025E08AC	-	4/4	-
14:2853539	C	ENSSSCG00000021967	ENSSSCT00000031967	3'UTR	857	-	-	-	-	rs318709876	SYK	Clone_based_vega_gene	11491	I3LPY9	UPI00025E08AC	-	4/4	-
14:11709332	C	ENSSSCG00000030130	ENSSSCT00000031898	3'UTR	2808	-	-	-	-	rs323252665	DPYSL2	HGNC	3014	I3LJE2	UPI00017F0284	-	14/14	-
14:11709372	C	ENSSSCG00000030130	ENSSSCT00000031898	3'UTR	2848	-	-	-	-	rs335186536	DPYSL2	HGNC	3014	I3LJE2	UPI00017F0284	-	14/14	-
14:13710947	G	ENSSSCG00000009680	ENSSSCT00000010614	3'UTR	2842	-	-	-	-	-	EXTL3	HGNC	3518	F1RJQ6	UPI00025DE96F	-	8/8	-
14:15483004	A	ENSSSCG00000009692	ENSSSCT00000010627	5'UTR	56	-	-	-	-	rs80938059	PINX1	HGNC	30046	F1RJ22	UPI00017F0675	-	1/7	-
14:21067922	C	ENSSSCG00000009708	ENSSSCT00000010644	5'UTR	88	-	-	-	-	rs329116642	AADAT	HGNC	17929	F1RIZ5	UPI00025DF528	-	1/14	-
14:23778814	C	ENSSSCG00000009724	ENSSSCT00000010662	missense	563	293	98	V/A	gTc/gCc	-	ANHX	HGNC	40024	F1RIV9	UPI00025DF28A	tol(1)	2/2	-
14:31685058	G	ENSSSCG00000009786	ENSSSCT00000010728	3'UTR	4492	-	-	-	-	rs345698995	HIP1R	HGNC	18415	F1REX8	UPI00025DF4E0	-	32/32	-
14:32817239	G	ENSSSCG00000025464	ENSSSCT00000023675	3'UTR	1352	-	-	-	-	-	ORAI1	Uniprot_gn	-	D3U7X2	UPI0001C97406	-	2/2	-
14:32817240	A	ENSSSCG00000025464	ENSSSCT00000023675	3'UTR	1351	-	-	-	-	-	ORAI1	Uniprot_gn	-	D3U7X2	UPI0001C97406	-	2/2	-
14:41129184	A	ENSSSCG00000009878	ENSSSCT00000010827	3'UTR	1153	-	-	-	-	-	RITA1	HGNC	25925	F1RKB0	UPI0001C97082	-	3/3	-
14:43991022	G	ENSSSCG00000009935	ENSSSCT00000023549	missense	2759	2759	920	V/A	gTg/gCg	rs334916098	MYO1H	HGNC	13879	I3LCV6	UPI00025E0448	tol(0.17)	27/31	-
14:46357034	A	ENSSSCG00000009957	ENSSSCT00000010912	missense	3109	3109	1037	V/M	Gtg/Atg	rs335877758	MYO18B	HGNC	18150	F1RG85	UPI00025E0EF0	tol(0.08)	16/42	-

14:50286076	C	ENSSSCG00000009997	ENSSSCT00000010954	5'UTR	113	-	-	-	-	rs80784455	OSM	HGNC	8506	F1RFE6	UPI00017F06B8	-	1/3	-
14:51617626	T	ENSSSCG00000010033	ENSSSCT00000010990	3'UTR	10203	-	-	-	-	rs321858502	PRR14L	HGNC	28738	F1RLV3	UPI0001C97223	-	9/9	-
14:52596800	G	ENSSSCG00000010046	ENSSSCT00000011007	3'UTR	2286	-	-	-	-	rs321311416	GNAZ	HGNC	4395	F1RLT6	UPI00017F0439	-	2/2	-
14:53016324	A	ENSSSCG00000010059	ENSSSCT00000011020	3'UTR	1354	-	-	-	-	rs343638247	GUCD1	HGNC	14237	F1RL42	UPI0001E886FE	-	6/6	-
14:79137612	G	ENSSSCG00000010267	ENSSSCT00000011237	missense	691	473	158	V/A	gTc/gCc	-	LRRC20	HGNC	23421	F1SUC9	UPI00025E116A	tol(0.85)	5/5	-
14:97822938	A	ENSSSCG00000010389	ENSSSCT00000011373	missense	2405	2405	802	R/Q	cGa/cAa	rs332655422	C10orf71	HGNC	26973	F1SEJ4	UPI00025DE7E0	tol(1)	1/1	-
14:97824231	A	ENSSSCG00000010389	ENSSSCT00000011373	missense	3698	3698	1233	P/H	cCt/cAt	rs332724862	C10orf71	HGNC	26973	F1SEJ4	UPI00025DE7E0	deleterious(0.02)	1/1	-
14:117676624	G	ENSSSCG00000010509	ENSSSCT00000011499	3'UTR	4623	-	-	-	-	-	PIK3AP1	HGNC	30034	F1SBF8	UPI00025DFB0A	-	17/17	-
14:123083941	C	ENSSSCG00000010577	ENSSSCT00000011570	3'UTR	1311	-	-	-	-	rs336349382	ELOVL3	Uniprot_gn	-	D0G6S7	UPI0001BAEF11	-	4/4	-
14:124870406	C	ENSSSCG00000010606	ENSSSCT00000011602	3'UTR	5617	-	-	-	-	rs336291088	SLK	HGNC	11088	F1S5P2	UPI00017EFFBC	-	18/18	-
14:142977582	C	ENSSSCG00000028063	ENSSSCT00000029273	missense	5188	5188	1730	S/P	Tct/Cct	rs330949980	TACC2	HGNC	11523	I3LM30	UPI00025E183F	-	2/23	-
14:142977582	C	ENSSSCG00000028063	ENSSSCT00000027768	missense	5188	5188	1730	S/P	Tct/Cct	rs330949980	TACC2	HGNC	11523	I3L8P8,K7GQ96	UPI00025E1840	tol(1)	2/20	-
14:143149846	C	ENSSSCG00000010701	ENSSSCT00000030603	splice_donor	-	-	-	-	-	rs324334912	BTBD16	HGNC	26340	I3LJE4	UPI00025E1868	-	-	3/14
14:152345028	A	ENSSSCG00000010762	ENSSSCT00000011774	3'UTR	2614	-	-	-	-	rs81450263	DPYSL4	HGNC	3016	F1SDI9	UPI00025DF2CE	-	18/18	-
15:54683333	T	ENSSSCG00000028592	ENSSSCT00000023110	3'UTR	2374	-	-	-	-	rs327022710	TM2D2	HGNC	24127	I3LTJ8	UPI00017F0BC0	-	4/4	-
15:54683333	T	ENSSSCG00000026608	ENSSSCT00000026862	3'UTR	2885	-	-	-	-	rs327022710	HTRA4	Uniprot_gn	-	I3L7K4,C7C1J1	UPI00025DF1EC	-	10/10	-
15:80374265	G	ENSSSCG00000015905	ENSSSCT00000017318	missense	1969	1969	657	I/V	Atc/Gtc	rs340042800	SCN2A	HGNC	10588	F1RPN2	UPI00025DF824	tol(1)	11/26	-
15:105739067	T	ENSSSCG00000016047	ENSSSCT00000017472	missense	269	157	53	L/I	Cta/Ata	-	MSTN	Uniprot_gn	-	Q9TSY4,Q9TSY2,Q95MF3,E9KYT4	UPI0000037259	tol(0.44)	1/3	-
15:105739095	A	ENSSSCG00000016047	ENSSSCT00000017472	missense	241	129	43	M/I	atG/atT	-	MSTN	Uniprot_gn	-	Q9TSY4,Q9TSY2,Q95MF3,E9KYT4	UPI0000037259	tol(0.19)	1/3	-
15:105739096	G	ENSSSCG00000016047	ENSSSCT00000017472	missense	240	128	43	M/T	aTg/aCg	-	MSTN	Uniprot_gn	-	Q9TSY4,Q9TSY2,Q95MF3,E9KYT4	UPI0000037259	tol(0.58)	1/3	-
15:105739315	T	ENSSSCG00000016047	ENSSSCT00000017472	5'UTR	21	-	-	-	-	-	MSTN	Uniprot_gn	-	Q9TSY4,Q9TSY2,Q95MF3,E9KYT4	UPI0000037259	-	1/3	-
15:105739318	A	ENSSSCG00000016047	ENSSSCT00000017472	5'UTR	18	-	-	-	-	-	MSTN	Uniprot_gn	-	Q9TSY4,Q9TSY2,Q95MF3,E9KYT4	UPI0000037259	-	1/3	-
15:117869790	C	ENSSSCG00000016117	ENSSSCT00000017548	missense	268	268	90	I/L	Att/Ctt	rs319035841	CARF	HGNC	14435	F1SHE7	UPI0001C98ACB	tol(0.47)	2/14	-
15:122804304	A	ENSSSCG00000016146	ENSSSCT00000017579	missense	757	479	160	T/M	aCg/aTg	rs331254965	C2orf80	HGNC	34352	F1SSU0	UPI00025DF474	tol(0.49)	7/7	-
15:155276711	C	ENSSSCG00000016385	ENSSSCT00000017840	missense	527	527	176	Y/S	tAt/tCt	rs346240966	-	-	-	F1SI00	UPI0001C9891C	deleterious(0)	5/10	-
15:155276720	A	ENSSSCG00000016385	ENSSSCT00000017840	missense	536	536	179	R/H	cGc/cAc	rs325260837	-	-	-	F1SI00	UPI0001C9891C	tol(0.07)	5/10	-
15:155279828	G	ENSSSCG00000016385	ENSSSCT00000017840	3'UTR	1193	-	-	-	-	rs333059360	-	-	-	F1SI00	UPI0001C9891C	-	10/10	-
15:155279844	T	ENSSSCG00000016385	ENSSSCT00000017840	3'UTR	1209	-	-	-	-	rs324465137	-	-	-	F1SI00	UPI0001C9891C	-	10/10	-

15:155280005	G	ENSSSCG00000016385	ENSSSCT00000017840	3'UTR	1370	-	-	-	-	-	-	-	-	F1SI00	UPI0001C9891C	-	10/10	-
15:155280012	G	ENSSSCG00000016385	ENSSSCT00000017840	3'UTR	1377	-	-	-	-	rs327758118	-	-	-	F1SI00	UPI0001C9891C	-	10/10	-
15:155280100	C	ENSSSCG00000016385	ENSSSCT00000017840	3'UTR	1465	-	-	-	-	rs343150693	-	-	-	F1SI00	UPI0001C9891C	-	10/10	-
15:155280106	A	ENSSSCG00000016385	ENSSSCT00000017840	3'UTR	1471	-	-	-	-	rs325668425	-	-	-	F1SI00	UPI0001C9891C	-	10/10	-
15:155280122	C	ENSSSCG00000016385	ENSSSCT00000017840	3'UTR	1487	-	-	-	-	-	-	-	-	F1SI00	UPI0001C9891C	-	10/10	-
15:155280123	A	ENSSSCG00000016385	ENSSSCT00000017840	3'UTR	1488	-	-	-	-	-	-	-	-	F1SI00	UPI0001C9891C	-	10/10	-
15:155311053	G	ENSSSCG00000016384	ENSSSCT00000017839	missense	16	16	6	T/A	Aca/Gca	-	-	-	-	F1SI01	UPI00025E07FE	deleterious(0)	1/5	-
15:155329495	A	ENSSSCG00000016383	ENSSSCT00000017838	missense	929	929	310	P/L	cCt/cTt	-	-	-	-	F1SI02	UPI00025E083F	tol(0.11)	4/8	-
15:155329499	A	ENSSSCG00000016383	ENSSSCT00000017838	missense	925	925	309	M/L	Atg/Ttg	-	-	-	-	F1SI02	UPI00025E083F	tol(1)	4/8	-
15:155329501	A	ENSSSCG00000016383	ENSSSCT00000017838	missense	923	923	308	G/V	gGc/gTc	-	-	-	-	F1SI02	UPI00025E083F	tol(0.51)	4/8	-
15:155329505	A	ENSSSCG00000016383	ENSSSCT00000017838	stop_gained	919	919	307	Q/*	Cag/Tag	-	-	-	-	F1SI02	UPI00025E083F	-	4/8	-
15:155329531	G	ENSSSCG00000016383	ENSSSCT00000017838	missense	893	893	298	G/A	gGg/gCg	rs344723792	-	-	-	F1SI02	UPI00025E083F	deleterious(0.02)	4/8	-
15:155456749	G	ENSSSCG00000021922	ENSSSCT00000031322	missense	57	57	19	H/Q	caC/caG	-	-	-	-	I3LAF1	UPI00025E0946	tol(0.3)	1/12	-
15:155458985	G	ENSSSCG00000021922	ENSSSCT00000031322	missense	430	430	144	S/A	Tcc/Gcc	rs346310609	-	-	-	I3LAF1	UPI00025E0946	tol(1)	6/12	-
15:155459094	C	ENSSSCG00000021922	ENSSSCT00000031322	missense	539	539	180	L/S	tTg/tCg	-	-	-	-	I3LAF1	UPI00025E0946	deleterious(0.02)	6/12	-
15:155459355	C	ENSSSCG00000021922	ENSSSCT00000031322	missense	800	800	267	V/A	gTa/gCa	rs329005715	-	-	-	I3LAF1	UPI00025E0946	tol(1)	6/12	-
15:155462239	G	ENSSSCG00000021922	ENSSSCT00000031322	missense	1399	1399	467	T/A	Acc/Gcc	rs339943645	-	-	-	I3LAF1	UPI00025E0946	tol(1)	9/12	-
15:155462375	C	ENSSSCG00000021922	ENSSSCT00000031322	missense	1480	1480	494	T/P	Aca/Cca	rs333849873	-	-	-	I3LAF1	UPI00025E0946	tol(0.41)	10/12	-
15:155462508	G	ENSSSCG00000021922	ENSSSCT00000031322	missense	1613	1613	538	D/G	gAt/gGt	-	-	-	-	I3LAF1	UPI00025E0946	tol(0.1)	10/12	-
15:155463623	C	ENSSSCG00000021922	ENSSSCT00000031322	missense	1879	1879	627	E/Q	Gag/Cag	-	-	-	-	I3LAF1	UPI00025E0946	deleterious(0.03)	12/12	-
15:155463810	G	ENSSSCG00000021922	ENSSSCT00000031322	missense	2066	2066	689	A/G	gCc/gGc	-	-	-	-	I3LAF1	UPI00025E0946	tol(1)	12/12	-
15:155463904	G	ENSSSCG00000021922	ENSSSCT00000031322	missense	2160	2160	720	N/K	aaC/aaG	rs326361934	-	-	-	I3LAF1	UPI00025E0946	tol(0.71)	12/12	-
15:155463942	T	ENSSSCG00000021922	ENSSSCT00000031322	missense	2198	2198	733	G/V	gGt/gTt	-	-	-	-	I3LAF1	UPI00025E0946	tol(0.62)	12/12	-
15:155464145	T	ENSSSCG00000021922	ENSSSCT00000031322	missense	2401	2401	801	V/L	Gtg/Ttg	rs327403430	-	-	-	I3LAF1	UPI00025E0946	deleterious(0.05)	12/12	-
15:155464343	A	ENSSSCG00000021922	ENSSSCT00000031322	missense	2599	2599	867	P/T	Cca/Aca	rs343366566	-	-	-	I3LAF1	UPI00025E0946	tol(1)	12/12	-
16:22312168	C	ENSSSCG00000016832	ENSSSCT00000018327	3'UTR	1486	-	-	-	-	rs81456876	IL7R	HGNC	6024	F1SNC0,K7GLT0,K7GLC0	UPI0001C98A29	-	8/8	-
16:22312168	C	ENSSSCG00000016832	ENSSSCT00000033303	3'UTR	1782	-	-	-	-	rs81456876	IL7R	HGNC	6024	K7GLT0,K7GLC0	UPI00028F4818	-	9/9	-
16:25621220	T	ENSSSCG00000016855	ENSSSCT00000018355	missense	315	152	51	S/N	aGc/aAc	rs341171034	FYB	HGNC	4036	F1SN73	UPI00025E092E	tol(1)	1/18	-

16:25621220	T	ENSSSCG00000016855	ENSSSCT00000035208	missense	163	122	41	S/N	aGc/aAc	rs341171034	FYB	HGNC	4036	K7GT11	UPI00028F45DE	tol(1)	1/11	-
16:36127334	A	ENSSSCG00000016898	ENSSSCT00000018403	missense	6874	6874	2292	L/F	Ctc/Ttc	rs345832663	-	-	-	F1SLP8	UPI00025E07AB	tol(0.31)	11/11	-
16:74633305	A	ENSSSCG00000017072	ENSSSCT00000018588	3'UTR	5265	-	-	-	-	rs338682833	GALNT10	Uniprot_gn	-	F1RQC2,K9IVH1	UPI00025E0A9B	-	12/12	-
16:80298217	T	ENSSSCG00000017100	ENSSSCT00000018618	3'UTR	3170	-	-	-	-	rs341319373	MTRR	HGNC	7473	F1S0T1	UPI0001C98AB1	-	15/15	-
17:9700043	C	ENSSSCG00000007003	ENSSSCT00000007671	missense	515	515	172	V/A	gTc/gCc	rs81241673	HGSNAT	HGNC	26527	F1SE48	UPI00025E0479	tol(0.59)	5/8	-
17:35255359	A	ENSSSCG00000007135	ENSSSCT00000007807	missense	2926	2657	886	R/H	cGt/cAt	rs323161357	NINL	HGNC	29163	F1SAS0	UPI0001C9660E	tol(0.15)	23/36	-
17:35255416	C	ENSSSCG00000007135	ENSSSCT00000007807	missense	2983	2714	905	D/A	gAt/gCt	rs332657072	NINL	HGNC	29163	F1SAS0	UPI0001C9660E	tol(0.1)	23/36	-
17:36229022	T	ENSSSCG00000023811	ENSSSCT00000031284	3'UTR	2745	-	-	-	-	rs341667059	AP5S1	HGNC	15875	I3LLR3	UPI00025E0630	-	3/3	-
17:36229961	A	ENSSSCG00000023811	ENSSSCT00000031284	3'UTR	1806	-	-	-	-	rs337392343	AP5S1	HGNC	15875	I3LLR3	UPI00025E0630	-	3/3	-
17:41538035	A	ENSSSCG00000029295	ENSSSCT00000025159	missense	331	331	111	P/S	Cct/Tct	rs339617998	BPIFB4	HGNC	16179	I3LH43	UPI00025DEE79	tol(1)	3/11	-
18:333729	T	ENSSSCG00000027174	ENSSSCT00000026338	missense	1906	1906	636	G/R	Ggg/Agg	-	-	-	-	I3L8T9	UPI00025DFE4E	tol(0.15)	13/13	-
18:333746	A	ENSSSCG00000027174	ENSSSCT00000026338	missense	1889	1889	630	A/V	gCc/gTc	-	-	-	-	I3L8T9	UPI00025DFE4E	deleterious(0.01)	13/13	-
18:334269	G	ENSSSCG00000027174	ENSSSCT00000026338	missense	1755	1755	585	E/D	gaG/gaC	rs342089965	-	-	-	I3L8T9	UPI00025DFE4E	deleterious(0.01)	12/13	-
18:1556885	T	ENSSSCG00000026652	ENSSSCT00000025384	missense	1214	1144	382	D/Y	Gat/Tat	rs323480955	-	-	-	I3LNT7	UPI00025E1226	deleterious(0.02)	8/8	-
18:18578858	A	ENSSSCG00000016548	ENSSSCT00000023443	3'UTR	2558	-	-	-	-	rs332187018	PODXL	HGNC	9171	I3LBX8	UPI00025DF72E	-	8/8	-
18:18578858	A	ENSSSCG00000016548	ENSSSCT00000018018	3'UTR	3189	-	-	-	-	rs332187018	PODXL	HGNC	9171	F1SNF3	UPI00025DF72C	-	9/9	-
18:18578858	A	ENSSSCG00000016548	ENSSSCT00000023302	3'UTR	2765	-	-	-	-	rs332187018	PODXL	HGNC	9171	I3L8I4	UPI00025DF72D	-	12/12	-
18:20911889	C	ENSSSCG00000016581	ENSSSCT00000018052	3'UTR	2620	-	-	-	-	rs339034014	CALU	HGNC	1458	F1SMN1	UPI0001C98657	-	7/7	-
18:35538731	G	ENSSSCG00000016643	ENSSSCT00000018121	3'UTR	2902	-	-	-	-	rs81211088	TMEM168	HGNC	25826	F1SJB3	UPI00025E1863	-	5/5	-
GL893171.1:19310	T	ENSSSCG00000023914	ENSSSCT00000023083	5'UTR	123	-	-	-	-	-	-	-	-	I3L822	UPI00025E0A01	-	1/4	-
GL894259.1:32546	G	ENSSSCG00000024726	ENSSSCT00000029110	3'UTR	480	-	-	-	-	-	RNASE1	Uniprot_gn	-	I3LGK3,D0PSF7	UPI00025E05E3	-	2/2	-
GL894914.1:23475	G	ENSSSCG00000023457	ENSSSCT00000029284	3'UTR	3660	-	-	-	-	-	-	-	-	I3LRT4	UPI00025DFE59	-	5/5	-
GL895152.1:21453	C	ENSSSCG00000021644	ENSSSCT00000028421	missense	1607	1607	536	V/G	gTt/gGt	-	BRD7	HGNC	14310	I3L640	UPI00025E0237	tol(0.2)	17/20	-
GL895152.1:21454	A	ENSSSCG00000021644	ENSSSCT00000028421	missense	1606	1606	536	V/F	Gtt/Ttt	-	BRD7	HGNC	14310	I3L640	UPI00025E0237	tol(0.65)	17/20	-
GL896090.1:17317	A	ENSSSCG00000027742	ENSSSCT00000025741	3'UTR	2022	-	-	-	-	-	RNF114	Uniprot_gn	-	I3LV30	UPI00025DEB87	-	5/5	-
GL896120.1:24142	G	ENSSSCG00000020871	ENSSSCT00000027390	3'UTR	1926	-	-	-	-	-	ARSD	HGNC	717	I3LM95	UPI00025DFB8F	-	12/12	-
GL896120.1:24170	C	ENSSSCG00000020871	ENSSSCT00000027390	3'UTR	1898	-	-	-	-	-	ARSD	HGNC	717	I3LM95	UPI00025DFB8F	-	12/12	-
GL896248.1:7776	G	ENSSSCG00000029303	ENSSSCT00000022761	missense	262	65	22	E/G	gAg/gGg	-	MRGBP	HGNC	15866	I3L9P3	UPI00025DEBFA	tol(1)	4/5	-

GL896252.1:2816	T	ENSSSCG00000022129	ENSSSCT00000025117	missense	56	36	12	E/D	gaA/gaT	-	ARSE	HGNC	719	I3L814	UPI00025DFDE9	tol(0.34)	1/11	-
GL896494.1:13835	A	ENSSSCG00000020884	ENSSSCT00000024402	missense	6266	6266	2089	R/K	aGg/aAg	-	GPR179	HGNC	31371	I3LP15	UPI00025DEE96	tol(1)	22/22	-
X:43903952	A	ENSSSCG00000025444	ENSSSCT00000032343	missense	611	611	204	R/H	cGc/cAc	rs342998794	EFHC2	HGNC	26233	I3L6K5	UPI00025DF57F	tol(0.65)	8/19	-
X:68708641	A	ENSSSCG00000012425	ENSSSCT00000013586	3’UTR	1656	-	-	-	-	-	UPRT	HGNC	28334	F1RPJ7	UPI00017F018F	-	7/7	-
X:142297482	T	ENSSSCG00000012796	ENSSSCT00000035620	3’UTR	5117	-	-	-	-	rs319221784	MECP2	HGNC	6990	K7GM47	UPI0001E88CBC	-	3/3	-

Table S3. Overlapping positions between sequencing and PorcineSNP60 Beadchip genotyping data.

Chr: Chromosome; Pos: Nucleotide position on the Sscrofa10.2 reference genome; Marker = Marker in the PorcineSNP60 Beadchip; SNP_Chip: Variation reported in the PorcineSNP60 Beadchip. A single allele is reported when all individuals are homozygous for that allele; Minor Allele: Allele identified as minor in the genotyping. The sign '-' indicates the absence of the allele; Major Allele: Allele identified as major in the genotyping. The sign '-' indicates the absence of the allele; MAF: Minor Allele Frequency in the genotyping; Reference: Reference nucleotide in the Sscrofa10.2 genome; N_Ref: Number of reads equal to the reference nucleotide; N_Alt: Number of reads alternative to the reference nucleotide; Q_Ref: Mean quality of the reads covering the reference nucleotide; Q_Alt: Mean quality of the reads covering the alternative nucleotide; SNP_NGS: Variation identified by the sequencing. A single allele is reported when no polymorphism was identified in the sequencing; SNAPE: positions called as SNP by SNAPE are labeled as SNAPE; Freq_MAF: allele frequency detected by sequencing of the allele identified by genotyping as minor; N_reads: Total number of reads covering the position.

Chr	Pos	Marker	SNP_Chip	Minor Allele	Major Allele	MAF	Reference	N_Ref	N_Alt	Q_Ref	Q_Alt	SNP	SNAPE	Freq_MAF	N_Reads
1	4873634	ALGA0000331	GA	G	A	0,432	C	1	2	62	60	TC	SNAPE	0,333	3
1	6036055	MARC0007285	AG	A	G	0,048	C	3	0	63	63	C	-	0,000	3
1	6834332	MARC0069522	AG	A	G	0,233	A	0	3	63	63	G	SNAPE	0,000	3
1	10318726	MARC0078284	AG	A	G	0,459	G	2	1	68	63	GA	-	0,333	3
1	16221205	MARC0023380	GA	G	A	0,096	A	3	0	66	66	A	-	0,000	3
1	17532523	ASGA0001310	GA	G	A	0,014	T	3	0	65	65	T	-	0,000	3
1	21835849	ASGA0001541	A	-	A	0,000	A	2	1	54	64	AC	-	0,333	3
1	24018825	ASGA0001590	AG	A	G	0,027	G	3	0	59	59	G	-	0,000	3
1	49192624	ASGA0002552	GA	G	A	0,171	A	2	3	67	57	GA	SNAPE	0,600	5
1	50857125	MARC0080112	GA	G	A	0,295	C	0	3	63	63	T	SNAPE	0,000	3
1	61043511	ALGA0003548	AG	A	G	0,356	G	4	0	57	57	G	-	0,000	4
1	63174967	ALGA0003646	GA	G	A	0,397	T	2	1	69	64	TC	-	0,333	3
1	65135487	ASGA0003080	AG	A	G	0,363	C	0	3	71	71	T	SNAPE	1,000	3
1	65878511	MARC0026746	AG	A	G	0,178	G	3	0	63	63	G	-	0,000	3
1	68187096	ASGA0003164	GA	G	A	0,384	T	1	3	68	63	CT	SNAPE	0,750	4
1	76056763	ASGA0003306	AC	A	C	0,240	T	1	2	60	67	GT	SNAPE	0,333	3
1	80751430	ASGA0003370	GA	G	A	0,123	T	3	0	65	65	T	-	0,000	3
1	81902314	ASGA0106258	AG	A	G	0,451	G	2	1	59	62	GA	-	0,333	3
1	82159992	MARC0079055	AG	A	G	0,445	C	2	1	64	68	CT	-	0,333	3
1	84928449	MARC0039482	AG	A	G	0,014	C	4	0	71	71	C	-	0,000	4
1	87072923	SIRI0000273	AG	A	G	0,438	T	3	0	67	67	T	-	1,000	3
1	87132357	MARC0093965	GA	G	A	0,110	A	4	0	62	62	A	-	0,000	4
1	93456091	ASGA0003694	GA	G	A	0,404	A	1	3	63	67	GA	SNAPE	0,750	4
1	97039899	H3GA0002210	GA	G	A	0,459	A	2	1	65	61	AG	-	0,333	3
1	100021690	ASGA0003793	AC	A	C	0,486	C	3	0	65	65	C	-	0,000	3
1	100578918	H3GA0053259	G	-	G	0,000	C	39	0	64	64	C	-	0,000	39
1	100579558	ASGA0084414	C	-	C	0,000	G	116	1	67	63	GT	-	0,009	117
1	100580257	ALGA0106810	GA	G	A	0,213	T	10	0	62	62	T	-	0,000	10
1	107343474	ALGA0005119	AG	A	G	0,438	G	4	0	67	67	G	-	0,000	4
1	109731334	M1GA0001077	AG	A	G	0,349	C	1	4	52	59	TC	SNAPE	0,800	5
1	112223763	ALGA0005274	C	-	C	0,000	C	3	0	64	64	C	-	0,000	3
1	120894943	MARC0007969	AG	A	G	0,411	C	0	3	65	65	T	SNAPE	1,000	3

1	125648304	ALGA0005613	AG	A	G	0,432	T	1	2	66	70	CT	SNAPE	0,333	3
1	132595455	ALGA0005754	GA	G	A	0,253	G	1	2	59	61	AG	SNAPE	0,333	3
1	140673154	ASGA0004475	GA	G	A	0,425	A	2	1	63	56	AG	-	0,333	3
1	144340414	MARC0029067	GA	G	A	0,055	T	2	1	67	69	TC	-	0,333	3
1	145464586	MARC0029823	GA	G	A	0,370	G	0	3	58	58	A	SNAPE	0,000	3
1	156055338	MARC0036647	GA	G	A	0,308	A	2	1	58	62	AG	-	0,333	3
1	157387413	SIRI0001167	AG	A	G	0,082	G	4	0	65	65	G	-	0,000	4
1	162225692	MARC0037074	GA	G	A	0,178	C	0	3	63	63	T	SNAPE	0,000	3
1	163106611	M1GA0001154	AC	A	C	0,336	T	1	2	69	67	GT	SNAPE	0,333	3
1	165658636	MARC0026211	GA	G	A	0,247	T	5	0	55	55	T	-	0,000	5
1	180335861	MARC0053567	AG	A	G	0,007	G	3	0	67	67	G	-	0,000	3
1	181495505	MARC0081465	AG	A	G	0,007	C	3	0	65	65	C	-	0,000	3
1	182485401	ASGA0005122	GA	G	A	0,185	A	3	0	57	57	A	-	0,000	3
1	193774677	ALGA0007038	GA	G	A	0,151	A	6	1	61	64	AG	-	0,143	7
1	198039938	DRGA0001730	G	-	G	0,000	C	3	0	68	68	C	-	0,000	3
1	204197370	ALGA0007224	GA	G	A	0,014	T	2	1	72	64	TC	-	0,333	3
1	204963217	MARC0013490	A	-	A	0,000	A	3	0	49	49	A	-	0,000	3
1	214209984	ALGA0007421	C	-	C	0,000	C	3	0	55	55	C	-	0,000	3
1	227282063	MARC0015301	AG	A	G	0,062	T	1	2	46	51	CT	SNAPE	0,333	3
1	232441844	INRA0005879	GA	G	A	0,466	G	3	0	59	59	G	-	1,000	3
1	246529471	MARC0004023	AG	A	G	0,055	C	3	0	46	46	C	-	0,000	3
1	251917501	H3GA0003784	GA	G	A	0,281	C	0	3	64	64	T	SNAPE	0,000	3
1	253518562	MARC0060067	G	-	G	0,000	C	3	0	64	64	C	-	0,000	3
1	268547138	MARC0027678	GA	G	A	0,144	T	3	0	67	67	T	-	0,000	3
1	268795123	ASGA0006375	AG	A	G	0,438	G	1	2	53	64	AG	SNAPE	0,667	3
1	282033624	MARC0032462	GA	G	A	0,068	T	3	0	67	67	T	-	0,000	3
1	284478293	ALGA0009414	AG	-	-	-	A	3	2	64	73	AG	SNAPE	0,600	5
1	284763000	ASGA0104727	AG	A	G	0,500	T	3	1	62	66	TC	-	0,750	4
1	286071507	ASGA0006959	GA	G	A	0,116	A	4	0	62	62	A	-	0,000	4
1	286400906	MARC0083500	A	-	A	0,000	T	5	0	64	64	T	-	0,000	5
1	291237694	ALGA0009831	GA	G	A	0,315	T	3	0	63	63	T	-	0,000	3
1	304515559	MARC0007390	G	-	G	0,000	G	5	0	62	62	G	-	0,000	5
1	309068867	MARC0015377	GA	G	A	0,130	A	2	1	67	64	AG	-	0,333	3
2	609952	ALGA0104042	AC	A	C	0,137	A	0	5	57	57	C	SNAPE	0,000	5
2	963130	MARC0022036	CA	C	A	0,240	G	0	4	64	64	T	SNAPE	0,000	4

2	2631862	ALGA0115217	AG	A	G	0,144	C	1	3	63	65	TC	SNAPE	0,750	4
2	4258564	ALGA0116722	AG	A	G	0,363	A	0	3	65	65	G	SNAPE	0,000	3
2	6998911	ASGA0008809	GA	G	A	0,075	G	0	3	62	62	A	SNAPE	0,000	3
2	7809332	ASGA0008859	AG	A	G	0,014	G	4	0	61	61	G	-	0,000	4
2	8681914	MARC0058937	AG	-	-	-	C	2	1	66	55	CT	-	0,667	3
2	9481805	ASGA0082164	GA	G	A	0,397	T	2	1	68	67	TC	-	0,333	3
2	10913056	ASGA0009021	AG	A	G	0,267	G	2	1	55	40	GA	-	0,333	3
2	16382771	ASGA0009299	GA	G	A	0,130	A	3	1	64	67	AG	-	0,250	4
2	19206291	H3GA0006190	GA	G	A	0,151	A	3	1	59	43	AG	-	0,250	4
2	19991449	MARC0013393	AG	A	G	0,014	C	3	0	58	58	C	-	0,000	3
2	23647170	MARC0050788	AG	A	G	0,486	A	2	1	70	68	AG	-	0,667	3
2	27857915	MARC0007620	AG	A	G	0,116	C	4	1	64	65	CT	-	0,200	5
2	28699945	MARC0028771	AG	A	G	0,267	C	2	1	61	60	CT	-	0,333	3
2	31566031	MARC0040962	AG	A	G	0,021	T	0	3	60	60	C	SNAPE	0,000	3
2	33623071	ASGA0084451	AG	A	G	0,158	A	1	2	54	65	GA	SNAPE	0,333	3
2	39628085	MARC0095742	CA	C	A	0,466	G	3	1	64	67	GT	-	0,750	4
2	41046132	MARC0090409	AG	A	G	0,295	C	3	0	61	61	C	-	0,000	3
2	41417366	ALGA0013110	AC	A	C	0,308	C	2	1	63	64	CA	-	0,333	3
2	41700706	ALGA0013140	AC	A	C	0,342	C	3	0	57	57	C	-	0,000	3
2	44299936	ASGA0010155	AG	A	G	0,438	T	3	0	63	63	T	-	1,000	3
2	44654907	ASGA0010162	AC	A	C	0,075	G	4	0	66	66	G	-	0,000	4
2	49437814	ASGA0097854	AG	A	G	0,370	C	2	1	67	68	CT	-	0,333	3
2	52593141	ALGA0013712	GA	G	A	0,185	G	0	3	55	55	A	SNAPE	0,000	3
2	79772013	MARC0075686	G	-	G	0,000	C	3	0	63	63	C	-	0,000	3
2	81675738	ALGA0014021	GA	G	A	0,315	A	3	1	59	65	AG	-	0,250	4
2	94133554	MARC0021401	AG	A	G	0,041	C	5	0	62	62	C	-	0,000	5
2	96053472	MARC0055959	AG	A	G	0,418	A	1	2	44	63	GA	SNAPE	0,333	3
2	100809723	MARC0016745	AG	A	G	0,377	T	0	3	67	67	C	SNAPE	0,000	3
2	101270993	MARC0013462	GA	G	A	0,041	C	0	3	66	66	T	SNAPE	0,000	3
2	109765071	MARC0087102	GA	G	A	0,423	C	2	1	59	66	CT	-	0,667	3
2	112531132	ALGA0014905	AC	A	C	0,313	G	2	1	67	63	GT	-	0,333	3
2	118204016	H3GA0007374	AG	A	G	0,473	G	4	0	64	64	G	-	0,000	4
2	118635459	MARC0021477	GA	G	A	0,459	T	3	1	55	68	TC	-	0,250	4
2	120009260	MARC0052388	GA	G	A	0,432	C	1	2	63	64	TC	SNAPE	0,333	3
2	127532245	ASGA0011632	GA	G	A	0,144	C	0	4	67	67	T	SNAPE	0,000	4

2	127651611	MARC0026920	AG	A	G	0,226	G	3	0	62	62	G	-	0,000	3
2	127886829	MARC0031266	AC	A	C	0,158	T	1	2	64	64	GT	SNAPE	0,333	3
2	139700871	ALGA0016182	AG	A	G	0,233	C	4	1	61	50	CT	-	0,200	5
2	141602306	MARC0019943	AG	-	-	-	G	5	0	63	63	G	-	1,000	5
2	142547802	ALGA0016393	GA	G	A	0,370	A	3	0	63	63	A	-	0,000	3
2	143061116	MARC0041805	AC	-	-	-	G	3	0	63	63	G	-	1,000	3
2	145637499	DIAS0003825	AC	A	C	0,389	C	3	1	56	66	CA	-	0,250	4
2	148345545	DIAS0004579	AG	A	G	0,192	T	0	3	68	68	C	SNAPE	0,000	3
2	154116865	H3GA0055347	AG	A	G	0,247	T	0	4	67	67	C	SNAPE	0,000	4
2	154582274	ASGA0094169	GA	G	A	0,329	A	3	0	61	61	A	-	0,000	3
2	156498625	ALGA0118122	AG	A	G	0,473	C	1	2	66	70	TC	SNAPE	0,667	3
2	157060689	ASGA0085387	AG	A	G	0,418	C	3	0	65	65	C	-	0,000	3
2	157845289	M1GA0003551	GA	G	A	0,493	T	0	3	59	59	C	SNAPE	1,000	3
3	5869767	ASGA0013114	GA	G	A	0,322	A	1	2	62	61	GA	SNAPE	0,667	3
3	13776008	MARC0052206	AG	A	G	0,438	A	0	4	49	49	G	SNAPE	0,000	4
3	13803309	MARC0073628	AC	A	C	0,451	A	3	1	63	66	AC	-	0,750	4
3	16201343	H3GA0055073	CA	C	A	0,384	C	3	2	56	61	CA	SNAPE	0,600	5
3	23876494	SIRI0001150	AG	A	G	0,212	T	0	5	61	61	C	SNAPE	0,000	5
3	23922431	DRGA0017485	G	-	G	0,000	G	3	0	57	57	G	-	0,000	3
3	26166212	MARC0078314	GA	G	A	0,500	G	3	0	70	70	G	-	1,000	3
3	30799997	ALGA0018337	TC	-	-	-	G	4	0	63	63	G	-	1,000	4
3	32604309	MARC0081077	AG	A	G	0,421	G	1	2	63	62	AG	SNAPE	0,667	3
3	36152439	MARC0009965	GA	G	A	0,089	T	3	0	52	52	T	-	0,000	3
3	38949614	ALGA0108469	AG	A	G	0,027	T	0	4	61	61	C	SNAPE	0,000	4
3	48578805	ALGA0018786	G	-	G	0,000	C	3	0	57	57	C	-	0,000	3
3	48924957	ALGA0106691	CA	C	A	0,082	A	3	0	66	66	A	-	0,000	3
3	49528314	ASGA0090160	AG	A	G	0,082	G	3	0	64	64	G	-	0,000	3
3	52245426	MARC0028829	CA	C	A	0,158	T	3	0	58	58	T	-	0,000	3
3	58059120	ASGA0014713	GA	G	A	0,048	T	3	0	57	57	T	-	0,000	3
3	59525623	MARC0058680	GA	G	A	0,363	C	1	4	63	64	TC	SNAPE	0,200	5
3	61588892	MARC0054644	AG	A	G	0,171	G	2	1	71	52	GA	-	0,333	3
3	63217196	MARC0009806	AG	A	G	0,205	G	2	1	54	64	GA	-	0,333	3
3	66168616	MARC0040824	AG	A	G	0,151	G	4	0	58	58	G	-	0,000	4
3	72016063	MARC0050071	GA	G	A	0,240	A	3	0	63	63	A	-	0,000	3
3	79508577	MARC0095438	AC	A	C	0,048	G	3	0	57	57	G	-	0,000	3

3	94868661	H3GA0010040	AG	-	-	-	A	3	0	62	62	A	-	1,000	3
3	100173669	DIAS0000405	GA	G	A	0,377	T	3	1	61	68	TC	-	0,250	4
3	102111329	MARC0037203	CA	C	A	0,137	G	0	3	66	66	T	SNAPE	0,000	3
3	107784484	ALGA0108384	CA	C	A	0,445	C	2	2	64	58	CA	-	0,500	4
3	109719833	MARC0076382	AG	A	G	0,247	G	3	0	60	60	G	-	0,000	3
3	114514650	ASGA0015824	GA	G	A	0,288	C	1	2	64	52	TC	-	0,333	3
3	117308221	MARC0038850	AG	A	G	0,379	G	2	1	56	71	GA	-	0,333	3
3	118182476	ASGA0015984	GA	G	A	0,233	A	3	0	61	61	A	-	0,000	3
3	121806793	MARC0013132	AG	A	G	0,322	G	3	0	66	66	G	-	0,000	3
3	122653666	MARC0013481	G	-	G	0,000	G	4	0	65	65	G	-	0,000	4
3	124204416	H3GA0010680	AG	A	G	0,075	T	1	2	66	66	CT	SNAPE	0,333	3
4	4693557	DRGA0004374	AC	A	C	0,110	G	3	1	65	53	GT	-	0,250	4
4	7002305	ASGA0017748	AG	A	G	0,336	G	3	0	57	57	G	-	0,000	3
4	10228266	MARC0017269	AC	A	C	0,377	G	2	1	42	56	GT	-	0,333	3
4	14893206	ASGA0018510	G	-	G	0,000	C	3	0	65	65	C	-	0,000	3
4	30118313	ALGA0024229	A	-	A	0,000	C	0	4	61	61	T	SNAPE	0,000	4
4	33016884	MARC0016020	AG	A	G	0,274	T	0	6	63	63	C	SNAPE	0,000	6
4	37508781	ASGA0019325	GA	G	A	0,192	A	4	0	61	61	A	-	0,000	4
4	38707173	ASGA0019349	GA	G	A	0,299	A	3	0	65	65	A	-	0,000	3
4	42557531	ALGA0024743	GA	G	A	0,130	T	4	0	52	52	T	-	0,000	4
4	45853311	DRGA0004781	G	-	G	0,000	G	3	0	60	60	G	-	0,000	3
4	47249204	ALGA0024905	AG	A	G	0,247	T	1	2	70	65	CT	SNAPE	0,333	3
4	55692680	ALGA0025041	G	-	G	0,000	C	4	0	54	54	C	-	0,000	4
4	60357934	ALGA0025151	A	-	A	0,000	A	3	0	58	58	A	-	0,000	3
4	61322889	ALGA0025197	G	-	G	0,000	G	3	0	63	63	G	-	0,000	3
4	62131797	MARC0004720	GA	G	A	0,336	T	2	4	65	60	CT	SNAPE	0,667	6
4	63988841	MARC0049861	AG	A	G	0,404	A	3	0	50	50	A	-	1,000	3
4	67653229	INRA0014411	AG	A	G	0,048	G	4	0	58	58	G	-	0,000	4
4	78073077	ALGA0025786	AG	A	G	0,397	G	2	1	56	54	GA	-	0,333	3
4	78608608	ALGA0025816	GA	G	A	0,089	C	2	2	71	69	CT	SNAPE	0,500	4
4	81026612	MARC0019918	AC	A	C	0,096	C	3	0	63	63	C	-	0,000	3
4	84206952	MARC0072324	GA	G	A	0,349	T	0	3	67	67	C	SNAPE	1,000	3
4	87370244	MARC0035059	AC	A	C	0,322	T	1	2	69	62	GT	SNAPE	0,333	3
4	90314757	MARC0040853	CA	C	A	0,068	A	3	0	60	60	A	-	0,000	3
4	92520727	MARC0011851	A	-	A	0,000	A	3	0	59	59	A	-	0,000	3

4	93365845	ASGA0020567	AG	A	G	0,438	C	2	2	67	59	CT	SNAPE	0,500	4
4	95276234	ALGA0026501	AG	A	G	0,240	T	2	2	64	68	TC	SNAPE	0,500	4
4	95499399	ASGA0020687	CA	C	A	0,240	A	3	0	55	55	A	-	0,000	3
4	98763083	ASGA0103970	CA	C	A	0,308	C	0	4	68	68	A	SNAPE	0,000	4
4	101953905	ALGA0026861	GA	G	A	0,178	A	2	2	65	61	AG	SNAPE	0,500	4
4	102706847	ALGA0026905	AC	A	C	0,104	G	3	0	66	66	G	-	0,000	3
4	104426777	ASGA0021116	AG	A	G	0,158	T	0	3	64	64	C	SNAPE	0,000	3
4	109928627	MARC0011217	CA	C	A	0,007	T	4	0	62	62	T	-	0,000	4
4	116260443	ALGA0027784	A	-	A	0,000	G	0	3	58	58	A	SNAPE	0,000	3
4	117247420	ASGA0021815	GA	G	A	0,253	C	0	4	64	64	T	SNAPE	0,000	4
4	119763345	ALGA0028034	GA	G	A	0,247	G	0	4	62	62	A	SNAPE	0,000	4
4	120344002	ASGA0022061	GA	G	A	0,027	A	2	1	66	66	AG	-	0,333	3
4	120489522	ASGA0022047	GA	G	A	0,034	T	3	0	68	68	T	-	0,000	3
4	124706183	ALGA0028398	AG	A	G	0,349	C	0	3	53	53	T	SNAPE	1,000	3
4	129079594	MARC0101639	GA	G	A	0,257	T	2	1	68	64	TC	-	0,333	3
4	130044947	ALGA0028716	AG	-	-	-	A	3	0	69	69	A	-	1,000	3
4	138629077	MARC0088218	CA	C	A	0,466	T	1	2	59	62	GT	SNAPE	0,667	3
4	141429369	MARC0077962	GA	G	A	0,432	T	1	2	63	64	CT	SNAPE	0,667	3
4	141823043	ASGA0023488	GA	G	A	0,500	G	2	1	60	70	GA	-	0,667	3
4	143310293	MARC0108103	GA	G	A	0,368	A	3	0	62	62	A	-	0,000	3
5	4410900	MARC0005539	AG	A	G	0,384	A	0	3	62	62	G	SNAPE	0,000	3
5	5932868	ALGA0030127	GA	G	A	0,130	T	3	0	64	64	T	-	0,000	3
5	8534753	H3GA0015531	GA	G	A	0,096	A	3	0	65	65	A	-	0,000	3
5	12659477	MARC0040397	GA	G	A	0,479	C	2	1	50	71	CT	-	0,667	3
5	14161325	H3GA0015764	GA	G	A	0,459	A	3	0	57	57	A	-	0,000	3
5	16776043	ALGA0030853	GA	G	A	0,260	C	0	3	63	63	T	SNAPE	0,000	3
5	18447205	DIAS0000646	AG	A	G	0,274	T	1	2	68	55	CT	SNAPE	0,333	3
5	18808373	ASGA0024930	AG	A	G	0,171	C	6	0	61	61	C	-	0,000	6
5	23163802	H3GA0016066	G	-	G	0,000	C	3	0	62	62	C	-	0,000	3
5	24769365	ASGA0096408	GA	G	A	0,384	C	2	1	63	69	CT	-	0,667	3
5	30237479	ALGA0031434	GA	G	A	0,333	G	0	3	63	63	A	SNAPE	0,000	3
5	38337110	ALGA0031722	AG	A	G	0,205	C	2	2	69	61	CT	SNAPE	0,500	4
5	38477450	MARC0021861	AG	A	G	0,274	T	0	5	62	62	C	SNAPE	0,000	5
5	38765537	ASGA0025444	AG	A	G	0,027	A	0	4	63	63	G	SNAPE	0,000	4
5	44994712	ALGA0031807	C	-	C	0,000	G	3	0	61	61	G	-	0,000	3

5	61894232	ALGA0032070	GA	G	A	0,123	T	3	0	59	59	T	-	0,000	3
5	62654938	ALGA0032158	GA	G	A	0,247	A	2	1	67	55	AG	-	0,333	3
5	67233849	ASGA0086192	GA	G	A	0,329	A	3	0	66	66	A	-	0,000	3
5	69369016	ALGA0032548	AG	A	G	0,247	G	3	0	60	60	G	-	0,000	3
5	69597659	ALGA0032593	AG	A	G	0,370	A	1	2	62	59	GA	SNAPE	0,333	3
5	71817704	MARC0005359	AG	A	G	0,034	C	3	0	65	65	C	-	0,000	3
5	73686403	H3GA0054252	G	-	G	0,000	G	3	0	57	57	G	-	0,000	3
5	83293808	ASGA0026433	GA	G	A	0,096	T	3	0	61	61	T	-	0,000	3
5	89819362	MARC0052288	AG	A	G	0,411	G	1	3	68	58	AG	SNAPE	0,750	4
5	107564798	DRGA0006452	GA	G	A	0,240	C	2	1	61	64	CT	-	0,667	3
6	469860	ALGA0103547	CA	C	A	0,205	A	2	1	62	65	AC	-	0,333	3
6	2085723	MARC0036664	GA	G	A	0,130	A	4	0	63	63	A	-	0,000	4
6	3976131	MARC0033020	AG	A	G	0,007	G	3	0	63	63	G	-	0,000	3
6	7643426	ASGA0095713	GA	G	A	0,438	G	0	3	67	67	A	SNAPE	0,000	3
6	8679743	ALGA0110070	AC	A	C	0,103	C	2	2	65	62	CA	SNAPE	0,500	4
6	8710349	ALGA0111931	AG	A	G	0,137	C	2	1	61	68	CT	-	0,333	3
6	10599132	ALGA0111656	TG	-	-	-	G	0	3	64	64	T	SNAPE	0,000	3
6	15302198	MARC0034453	GA	G	A	0,459	T	1	3	62	58	CT	SNAPE	0,750	4
6	16435748	MARC0046484	CA	C	A	0,219	T	4	0	61	61	T	-	0,000	4
6	17917275	DIAS0000453	AG	A	G	0,295	T	0	3	63	63	C	SNAPE	0,000	3
6	21052994	ALGA0034835	CA	C	A	0,062	A	2	1	71	69	AC	-	0,333	3
6	24347591	ALGA0109690	GA	G	A	0,123	A	3	0	60	60	A	-	0,000	3
6	24354180	MARC0085497	G	-	G	0,000	C	3	0	61	61	C	-	0,000	3
6	24928813	MARC0032851	AG	A	G	0,007	C	4	0	57	57	C	-	0,000	4
6	27654031	ASGA0102960	AG	A	G	0,458	G	1	3	75	62	AG	SNAPE	0,750	4
6	27995521	ASGA0082257	GA	G	A	0,390	A	3	3	67	71	AG	SNAPE	0,500	6
6	30261804	ASGA0028031	AC	A	C	0,418	T	4	1	65	52	TG	-	0,800	5
6	31475831	ALGA0104027	AC	A	C	0,041	G	4	0	65	65	G	-	0,000	4
6	40248689	ASGA0095216	A	-	A	0,000	T	4	0	63	63	T	-	0,000	4
6	42375050	SIRI0000740	GA	G	A	0,315	A	2	1	55	53	AG	-	0,333	3
6	43090860	MARC0017436	GA	G	A	0,425	T	2	2	68	66	TC	SNAPE	0,500	4
6	43254597	ALGA0108257	AG	A	G	0,007	T	0	4	66	66	C	SNAPE	0,000	4
6	44231753	MARC0042606	GA	G	A	0,055	T	3	0	66	66	T	-	0,000	3
6	46463710	MARC0070998	AG	A	G	0,459	A	2	1	55	63	AG	-	0,667	3
6	47286643	ASGA0028202	AG	A	G	0,356	A	0	4	55	55	G	SNAPE	0,000	4

6	55763196	MARC0005355	A	-	A	0,000	A	3	0	66	66	A	-	0,000	3
6	65811584	ALGA0035551	GA	G	A	0,425	C	2	2	62	65	CT	SNAPE	0,500	4
6	66570874	MARC0012843	AG	A	G	0,390	T	1	2	63	67	CT	SNAPE	0,333	3
6	68263560	MARC0035195	AG	A	G	0,356	G	3	0	60	60	G	-	0,000	3
6	69930291	H3GA0056439	GA	G	A	0,322	T	4	0	64	64	T	-	0,000	4
6	70317828	ALGA0115349	AG	A	G	0,219	C	4	0	60	60	C	-	0,000	4
6	70386367	MARC0098064	AG	A	G	0,315	C	*							
6	73745517	MARC0011512	GA	G	A	0,370	G	1	3	54	58	AG	SNAPE	0,250	4
6	75365266	ASGA0086694	AG	A	G	0,123	G	3	0	62	62	G	-	0,000	3
6	75417134	ALGA0110359	GA	G	A	0,292	T	1	2	55	59	CT	SNAPE	0,667	3
6	84047659	ASGA0101193	GA	G	A	0,322	A	3	0	66	66	A	-	0,000	3
6	84333986	ALGA0116144	GA	G	A	0,295	A	3	0	61	61	A	-	0,000	3
6	88655325	MARC0033580	AG	A	G	0,096	C	4	0	56	56	C	-	0,000	4
6	91640556	ALGA0121255	GA	G	A	0,295	T	2	1	68	66	TC	-	0,333	3
6	92447955	MARC0072646	GA	G	A	0,219	C	0	3	55	55	T	SNAPE	0,000	3
6	101205770	M1GA0008862	CA	C	A	0,007	T	4	0	66	66	T	-	0,000	4
6	107414338	MARC0011750	AC	A	C	0,397	A	1	2	70	70	CA	SNAPE	0,333	3
6	119148364	MARC0031744	GA	G	A	0,089	A	3	0	65	65	A	-	0,000	3
6	124297805	ALGA0121932	GA	G	A	0,425	T	1	2	63	66	CT	SNAPE	0,667	3
6	138314076	ASGA0029738	TC	-	-	-	T	0	3	69	69	C	SNAPE	0,000	3
6	150135326	MARC0009266	GA	G	A	0,452	C	2	2	66	65	CT	SNAPE	0,500	4
6	150772134	ALGA0121862	GA	G	A	0,226	G	1	2	60	61	AG	SNAPE	0,333	3
6	153108429	ALGA0037706	GA	G	A	0,179	C	2	2	66	67	CT	SNAPE	0,500	4
6	155441353	ASGA0030357	GA	G	A	0,384	A	3	0	57	57	A	-	0,000	3
7	1150480	MARC0027365	CA	C	A	0,034	G	0	3	62	62	T	SNAPE	0,000	3
7	1989754	SIRI0000004	GA	G	A	0,377	T	2	1	65	64	TC	-	0,333	3
7	5822223	H3GA0019706	AG	A	G	0,281	T	1	2	62	56	CT	SNAPE	0,333	3
7	7206627	ALGA0038327	AC	A	C	0,219	C	3	1	67	65	CA	-	0,250	4
7	16801051	ASGA0031489	AC	A	C	0,260	C	4	0	66	66	C	-	0,000	4
7	18009952	ALGA0039018	GA	G	A	0,308	T	4	0	66	66	T	-	0,000	4
7	18545282	MARC0044748	CA	C	A	0,432	C	1	2	66	67	AC	SNAPE	0,333	3
7	27826201	M1GA0009777	AG	A	G	0,329	G	2	1	66	68	GA	-	0,333	3
7	33618223	MARC0068518	AG	A	G	0,123	T	0	4	62	62	C	SNAPE	0,000	4
7	33790291	H3GA0020692	GA	G	A	0,048	A	5	0	67	67	A	-	0,000	5
7	34507110	ASGA0032442	CG	C	G	0,116	C	5	0	63	63	C	-	0,000	5

7	36453640	ASGA0032592	GA	G	A	0,390	A	4	1	57	67	AG	-	0,200	5
7	37084501	ASGA0032630	AG	A	G	0,459	T	1	2	40	69	CT	SNAPE	0,333	3
7	39985401	ASGA0032851	GA	G	A	0,459	T	1	2	54	49	CT	SNAPE	0,667	3
7	54377612	INRA0025801	GA	G	A	0,212	C	0	3	66	66	T	SNAPE	0,000	3
7	57463816	ALGA0042082	AC	A	C	0,048	G	4	1	55	67	GT	-	0,200	5
7	58515043	ASGA0034178	AG	A	G	0,055	C	4	0	67	67	C	-	0,000	4
7	59245765	H3GA0021868	A	-	A	0,000	A	3	0	57	57	A	-	0,000	3
7	60636683	ALGA0042217	GA	G	A	0,267	C	0	3	55	55	T	SNAPE	0,000	3
7	64503112	ASGA0034291	AG	A	G	0,493	A	2	1	62	71	AG	-	0,667	3
7	69481673	MARC0025572	AC	A	C	0,315	C	2	1	59	64	CA	-	0,333	3
7	70176985	MARC0036603	AC	A	C	0,021	G	3	0	60	60	G	-	0,000	3
7	93658566	DRGA0007954	GA	G	A	0,087	T	2	1	65	64	TC	-	0,333	3
7	98396938	MARC0111338	AG	A	G	0,075	C	3	0	67	67	C	-	0,000	3
7	102792552	ASGA0035486	CA	C	A	0,082	T	3	0	57	57	T	-	0,000	3
7	103105644	MARC0041419	AG	A	G	0,267	C	3	0	65	65	C	-	0,000	3
7	104807152	DBNP0000926	GA	G	A	0,096	C	0	3	59	59	T	SNAPE	0,000	3
7	115625934	MARC0001424	CA	C	A	0,226	A	4	0	63	63	A	-	0,000	4
7	117195566	ASGA0036089	AG	A	G	0,267	T	1	2	64	57	CT	SNAPE	0,333	3
7	119267329	ALGA0044931	AG	A	G	0,233	C	6	0	59	59	C	-	0,000	6
7	122910008	M1GA0010866	AG	A	G	0,123	C	3	0	66	66	C	-	0,000	3
7	128541986	M1GA0011142	AC	A	C	0,322	A	1	2	71	63	CA	SNAPE	0,333	3
7	134005327	ASGA0037326	AG	A	G	0,164	G	2	1	64	71	GA	-	0,333	3
8	924576	ASGA0037412	CA	C	A	0,458	C	1	2	64	63	AC	SNAPE	0,333	3
8	3779973	ASGA0037525	AG	A	G	0,363	C	1	2	67	67	TC	SNAPE	0,667	3
8	4915510	MARC0027651	AG	A	G	0,473	G	1	2	43	63	AG	SNAPE	0,667	3
8	5094390	ASGA0037551	GA	G	A	0,014	C	0	3	67	67	T	SNAPE	0,000	3
8	6758516	MARC0002041	AC	-	-	-	A	0	3	61	61	C	SNAPE	0,000	3
8	14836149	MARC0054361	GA	G	A	0,390	T	4	0	63	63	T	-	0,000	4
8	27238519	DRGA0008451	G	-	G	0,000	C	3	0	60	60	C	-	0,000	3
8	43691362	H3GA0024834	AG	-	-	-	T	1	2	65	67	CT	SNAPE	0,333	3
8	59533854	MARC0062246	AG	A	G	0,041	G	4	0	61	61	G	-	0,000	4
8	64409837	ASGA0038865	AG	A	G	0,493	G	0	3	67	67	A	SNAPE	1,000	3
8	71573944	ASGA0038900	AG	A	G	0,212	A	1	3	69	66	GA	SNAPE	0,250	4
8	75574086	DIAS0001639	AG	A	G	0,226	T	1	2	58	62	CT	SNAPE	0,333	3
8	75716950	MARC0065725	GA	G	A	0,055	C	0	3	62	62	T	SNAPE	0,000	3

8	87705075	M1GA0011981	AC	A	C	0,068	C	4	0	51	51	C	-	0,000	4
8	95129988	MARC0073866	CA	C	A	0,459	G	0	3	66	66	T	SNAPE	0,000	3
8	117959809	H3GA0025318	AG	A	G	0,486	G	1	3	63	61	AG	SNAPE	0,750	4
8	128703259	ASGA0039757	AG	A	G	0,425	G	1	2	63	68	AG	SNAPE	0,667	3
8	131014130	MARC0053144	AG	A	G	0,137	A	0	4	67	67	G	SNAPE	0,000	4
8	131295589	MARC0026632	GA	G	A	0,021	T	3	0	56	56	T	-	0,000	3
8	136467907	ALGA0049696	AG	-	-	-	T	1	2	65	63	CT	SNAPE	0,333	3
9	2976048	MARC0098905	AG	-	-	-	G	1	2	64	67	AG	SNAPE	0,333	3
9	7178792	ALGA0118900	TA	T	A	0,404	A	4	1	65	68	AT	-	0,200	5
9	9368238	MARC0005906	GA	G	A	0,438	G	4	2	63	62	GA	SNAPE	0,667	6
9	11023938	M1GA0012669	GA	G	A	0,479	A	3	0	62	62	A	-	0,000	3
9	23888625	MARC0113715	A	-	A	0,000	A	4	0	64	64	A	-	0,000	4
9	27528813	ASGA0042170	GA	G	A	0,356	G	1	2	60	69	AG	SNAPE	0,333	3
9	28357982	ALGA0109015	AG	A	G	0,226	G	2	1	53	70	GA	-	0,333	3
9	39183000	MARC0051875	GA	G	A	0,062	T	4	0	64	64	T	-	0,000	4
9	40848510	ASGA0085910	A	-	A	0,000	A	9	0	63	63	A	-	0,000	9
9	48972341	ALGA0052809	AG	A	G	0,438	C	2	4	55	63	TC	SNAPE	0,667	6
9	49981877	MARC0042421	AG	A	G	0,205	C	5	0	64	64	C	-	0,000	5
9	51593990	ALGA0052916	AG	A	G	0,240	C	3	0	66	66	C	-	0,000	3
9	63483264	ASGA0098417	AC	A	C	0,431	T	2	2	56	71	TG	SNAPE	0,500	4
9	67199879	ASGA0084646	GA	G	A	0,406	T	2	1	66	56	TC	-	0,333	3
9	73855980	ALGA0056145	GA	G	A	0,370	A	1	2	53	67	GA	SNAPE	0,667	3
9	75391264	MARC0091838	AG	A	G	0,158	C	3	0	59	59	C	-	0,000	3
9	76936577	ALGA0053741	C	-	C	0,000	G	3	0	64	64	G	-	0,000	3
9	98616809	MARC0111607	G	-	G	0,000	C	3	0	62	62	C	-	0,000	3
9	100612906	MARC0037787	GA	G	A	0,500	G	1	2	75	65	AG	SNAPE	0,333	3
9	104996366	MARC0039925	AG	A	G	0,103	G	3	0	44	44	G	-	0,000	3
9	123847767	MARC0087389	GA	G	A	0,116	G	0	3	65	65	A	SNAPE	0,000	3
9	123898599	MARC0032537	AG	A	G	0,212	C	3	0	67	67	C	-	0,000	3
9	126542393	ASGA0044391	AG	A	G	0,377	G	3	3	58	66	GA	SNAPE	0,500	6
9	131957271	ALGA0054972	GA	G	A	0,390	A	1	2	65	63	GA	SNAPE	0,667	3
9	132659971	H3GA0052528	GA	G	A	0,377	T	4	0	62	62	T	-	0,000	4
9	135773590	MARC0031935	AG	A	G	0,301	G	4	0	48	48	G	-	0,000	4
9	136170919	ALGA0055124	AG	A	G	0,096	C	3	0	64	64	C	-	0,000	3
9	136848852	MARC0005452	GA	G	A	0,384	G	3	1	60	66	GA	-	0,750	4

9	139299087	ALGA0055311	GA	G	A	0,479	C	1	2	58	65	TC	SNAPE	0,333	3
9	144689456	MARC0103295	GA	G	A	0,397	A	0	3	56	56	G	SNAPE	1,000	3
9	145998807	ASGA0045135	AG	A	G	0,164	C	3	0	67	67	C	-	0,000	3
9	148985546	ALGA0122913	CA	C	A	0,486	A	1	2	66	65	CA	SNAPE	0,667	3
10	16834168	ALGA0057334	GA	G	A	0,377	T	2	3	66	62	CT	SNAPE	0,600	5
10	21159125	MARC0102096	AG	A	G	0,041	G	3	0	66	66	G	-	0,000	3
10	29045919	ASGA0047195	AG	A	G	0,151	G	3	1	66	64	GA	-	0,250	4
10	32953632	ALGA0058126	AG	A	G	0,082	T	1	4	67	58	CT	SNAPE	0,200	5
10	35068814	ALGA0113565	AG	A	G	0,329	T	0	3	65	65	C	SNAPE	0,000	3
10	41577189	INRA0033777	G	-	G	0,000	G	3	0	62	62	G	-	0,000	3
10	46203553	ALGA0058642	AG	A	G	0,137	G	3	0	62	62	G	-	0,000	3
10	48844125	MARC0075511	GA	G	A	0,493	A	3	0	65	65	A	-	0,000	3
10	49173528	MARC0057581	TG	-	-	-	T	4	1	68	68	TG	-	0,800	5
10	51947116	ASGA0091625	AG	A	G	0,452	G	1	3	68	65	AG	SNAPE	0,750	4
10	57954470	ASGA0048292	GA	G	A	0,075	T	3	0	61	61	T	-	0,000	3
10	59340783	MARC0005804	A	-	A	0,000	T	4	0	60	60	T	-	0,000	4
10	60897238	ALGA0059417	GA	G	A	0,473	G	2	1	68	64	GA	-	0,667	3
10	62591687	ASGA0048487	CA	C	A	0,288	C	1	3	72	68	AC	SNAPE	0,250	4
10	63439731	MARC0094056	A	-	A	0,000	A	3	0	66	66	A	-	0,000	3
10	68845976	ASGA0104631	AG	A	G	0,295	C	3	0	59	59	C	-	0,000	3
10	69624734	H3GA0030667	AG	A	G	0,329	A	1	3	66	62	GA	SNAPE	0,250	4
10	73852395	ASGA0049073	AG	A	G	0,212	A	0	3	51	51	G	SNAPE	0,000	3
10	76190839	MARC0007196	A	-	A	0,000	T	4	0	64	64	T	-	0,000	4
11	4689240	ASGA0049406	GA	G	A	0,452	T	3	2	61	60	TC	-	0,400	5
11	11739965	MARC0069399	AG	A	G	0,021	A	0	3	66	66	G	SNAPE	0,000	3
11	12518680	H3GA0031362	AG	A	G	0,212	C	3	1	47	62	CT	-	0,250	4
11	20953383	ALGA0061277	AG	A	G	0,438	G	0	3	64	64	A	SNAPE	1,000	3
11	24236768	ALGA0061460	GA	G	A	0,363	A	0	3	56	56	G	SNAPE	1,000	3
11	28140288	MARC0002392	AG	A	G	0,048	C	3	0	62	62	C	-	0,000	3
11	50012027	ASGA0050765	AG	A	G	0,349	G	1	2	64	53	AG	SNAPE	0,667	3
11	52861137	MARC0025052	AG	A	G	0,301	C	3	0	60	60	C	-	0,000	3
11	53388276	ALGA0062287	AG	A	G	0,137	C	3	0	56	56	C	-	0,000	3
11	54758069	INRA0036477	GA	G	A	0,144	T	6	0	62	62	T	-	0,000	6
11	62217369	ALGA0062526	AG	A	G	0,068	C	2	1	61	56	CT	-	0,333	3
11	76244052	ASGA0051662	GA	G	A	0,048	T	3	0	44	44	T	-	0,000	3

11	82867524	ALGA0064019	CA	C	A	0,144	C	2	1	60	62	CA	-	0,667	3
11	84726075	ASGA0052333	GA	G	A	0,212	A	1	2	60	59	GA	SNAPE	0,667	3
12	1449077	ASGA0099135	GA	G	A	0,226	T	0	3	68	68	C	SNAPE	1,000	3
12	4909182	MARC0107758	G	-	G	0,000	C	4	0	64	64	C	-	0,000	4
12	7075951	ASGA0052681	AG	A	G	0,432	A	2	3	66	62	GA	SNAPE	0,400	5
12	15141877	ASGA0053428	CA	C	A	0,225	C	0	3	57	57	A	SNAPE	0,000	3
12	17941131	MARC0053390	CA	C	A	0,386	G	2	1	68	65	GT	-	0,667	3
12	23161884	MARC0113309	AG	A	G	0,219	C	2	1	70	66	CT	-	0,333	3
12	25324124	DIAS0001467	AG	-	-	-	A	0	3	67	67	G	SNAPE	0,000	3
12	27939866	MARC0054687	AG	A	G	0,247	T	1	2	63	65	CT	SNAPE	0,333	3
12	38358087	MARC0039716	AG	A	G	0,500	G	2	1	57	64	GA	-	0,333	3
12	42580910	ALGA0066378	AG	A	G	0,188	C	2	1	66	69	CT	-	0,333	3
12	43987384	MARC0072172	GA	G	A	0,192	G	1	4	56	54	AG	SNAPE	0,200	5
12	48749400	MARC0034121	GA	G	A	0,096	A	3	0	54	54	A	-	0,000	3
12	49919795	ASGA0054868	GA	G	A	0,253	A	3	0	68	68	A	-	0,000	3
12	49934017	M1GA0016848	G	-	G	0,000	C	3	0	66	66	C	-	0,000	3
12	50373701	ALGA0066729	AG	A	G	0,014	G	3	0	69	69	G	-	0,000	3
12	51604624	H3GA0055908	AC	A	C	0,082	A	0	4	63	63	C	SNAPE	0,000	4
12	58370227	MARC0012228	C	-	C	0,000	C	3	0	65	65	C	-	0,000	3
12	60228089	ALGA0103202	AG	-	-	-	A	3	0	57	57	A	-	1,000	3
12	63095062	MARC0114025	AG	A	G	0,257	C	3	0	65	65	C	-	0,000	3
13	1505974	MARC0048559	AG	A	G	0,007	G	3	0	64	64	G	-	0,000	3
13	2676033	ALGA0067450	AG	A	G	0,418	T	3	0	70	70	T	-	1,000	3
13	3697801	MARC0007354	GA	G	A	0,110	T	3	0	63	63	T	-	0,000	3
13	7650574	MARC0037054	AG	A	G	0,466	T	2	1	48	57	TC	-	0,667	3
13	10650337	ASGA0105603	CA	C	A	0,253	T	4	0	55	55	T	-	0,000	4
13	12132345	MARC0071364	AG	A	G	0,107	G	3	1	59	67	GA	-	0,250	4
13	12757494	MARC0008185	AG	A	G	0,500	G	2	2	64	59	GA	SNAPE	0,500	4
13	13594790	ALGA0068039	GA	G	A	0,226	G	3	0	64	64	G	-	1,000	3
13	21806276	ALGA0068736	TC	-	-	-	T	2	1	58	67	TC	-	0,667	3
13	24421663	ALGA0116366	AG	A	G	0,048	C	3	0	67	67	C	-	0,000	3
13	25158584	ALGA0068831	AG	A	G	0,062	T	0	3	62	62	C	SNAPE	0,000	3
13	29929800	ASGA0056949	GA	G	A	0,274	G	1	2	57	63	AG	SNAPE	0,333	3
13	30842222	MARC0053125	TC	-	-	-	T	2	1	61	68	TC	-	0,667	3
13	32830936	ASGA0096353	AG	A	G	0,288	G	3	0	66	66	G	-	0,000	3

13	38925584	MARC0076574	GA	G	A	0,155	A	3	0	61	61	A	-	0,000	3
13	43751281	INRA0040093	G	-	G	0,000	G	3	0	59	59	G	-	0,000	3
13	46574837	MARC0040165	AC	A	C	0,308	C	2	1	62	63	CA	-	0,333	3
13	47488117	ALGA0069872	AG	A	G	0,253	T	0	4	62	62	C	SNAPE	0,000	4
13	50478692	MARC0052375	GA	G	A	0,055	T	3	0	65	65	T	-	0,000	3
13	55201128	DRGA0012367	GA	G	A	0,144	A	2	1	67	57	AG	-	0,333	3
13	60124449	MARC0061219	AG	A	G	0,111	A	1	2	64	63	GA	SNAPE	0,333	3
13	60486947	ASGA0057747	GA	G	A	0,466	G	1	2	64	55	AG	SNAPE	0,333	3
13	65746490	SIRI0001436	CA	C	A	0,219	G	3	2	65	61	GT	SNAPE	0,600	5
13	67262717	ALGA0070563	AG	A	G	0,308	C	1	2	60	65	TC	SNAPE	0,667	3
13	77123152	MARC0052143	GA	G	A	0,390	A	1	2	59	53	GA	SNAPE	0,667	3
13	89335311	ASGA0058393	AG	A	G	0,116	G	1	2	68	69	AG	SNAPE	0,667	3
13	91972877	MARC0075708	AG	A	G	0,137	A	0	3	66	66	G	SNAPE	0,000	3
13	95375353	ASGA0058509	C	-	C	0,000	C	4	0	63	63	C	-	0,000	4
13	97309723	ASGA0058518	GA	G	A	0,015	T	4	0	63	63	T	-	0,000	4
13	99741449	ALGA0071495	AG	A	G	0,041	G	3	0	56	56	G	-	0,000	3
13	101512473	MARC0049284	GA	G	A	0,021	T	4	0	63	63	T	-	0,000	4
13	103171133	DRGA0012768	GA	G	A	0,185	A	2	1	61	67	AG	-	0,333	3
13	103484502	MARC0019359	CA	C	A	0,192	T	3	0	67	67	T	-	0,000	3
13	119280288	ALGA0071703	GA	G	A	0,410	C	0	3	64	64	T	SNAPE	0,000	3
13	124746808	INRA0040886	GA	G	A	0,144	C	0	3	57	57	T	SNAPE	0,000	3
13	126861637	MARC0021264	CA	C	A	0,322	C	1	2	68	67	AC	SNAPE	0,333	3
13	130982311	MARC0019443	G	-	G	0,000	C	4	0	67	67	C	-	0,000	4
13	133941534	MARC0076402	G	-	G	0,000	G	3	0	63	63	G	-	0,000	3
13	134736263	MARC0045469	AC	A	C	0,349	C	2	1	66	70	CA	-	0,333	3
13	136647355	ASGA0058797	GA	G	A	0,014	T	5	0	65	65	T	-	0,000	5
13	138091331	MARC0076957	GA	G	A	0,472	A	1	2	55	71	GA	SNAPE	0,667	3
13	145300325	ALGA0072095	AC	A	C	0,167	C	2	1	62	70	CA	-	0,333	3
13	147536783	MARC0105487	GA	G	A	0,459	T	1	3	66	63	CT	SNAPE	0,750	4
13	150082135	DIAS0000815	AG	A	G	0,212	C	1	2	52	59	TC	SNAPE	0,667	3
13	168220231	MARC0024620	AC	A	C	0,233	C	2	1	66	60	CA	-	0,333	3
13	175687313	ALGA0072751	C	-	C	0,000	G	4	0	64	64	G	-	0,000	4
13	177981732	ALGA0072779	AG	A	G	0,418	G	2	1	60	56	GA	-	0,333	3
13	179432777	MARC0033381	C	-	C	0,000	C	3	0	53	53	C	-	0,000	3
13	185614224	ALGA0072903	AG	A	G	0,370	T	0	3	64	64	C	SNAPE	0,000	3

13	186878981	MARC0037806	G	-	G	0,000	C	3	0	62	62	C	-	0,000	3
13	196739971	MARC0046795	GA	G	A	0,164	T	1	2	64	65	CT	SNAPE	0,667	3
13	199273808	ASGA0059654	AG	A	G	0,130	G	3	0	65	65	G	-	0,000	3
13	208316160	ALGA0073752	GA	G	A	0,123	A	3	1	68	67	AG	-	0,250	4
13	211361839	ASGA0060000	AG	A	G	0,445	G	2	1	66	65	GA	-	0,333	3
14	900865	ALGA0074201	AG	A	G	0,432	C	2	1	66	63	CT	-	0,333	3
14	9819387	MARC0032050	AG	A	G	0,253	A	1	2	70	63	GA	SNAPE	0,333	3
14	13391538	MARC0029560	AG	A	G	0,370	A	0	3	67	67	G	SNAPE	0,000	3
14	13778165	ASGA0061516	GA	G	A	0,130	A	2	1	65	56	AG	-	0,333	3
14	20500464	MARC0114116	GA	G	A	0,258	C	0	3	64	64	T	SNAPE	0,000	3
14	21157940	MARC0010043	GA	G	A	0,222	A	2	2	68	61	AG	SNAPE	0,500	4
14	27518647	ALGA0076388	AG	A	G	0,363	T	1	2	50	60	CT	SNAPE	0,333	3
14	27942025	MARC0065414	G	-	G	0,000	C	4	0	66	66	C	-	0,000	4
14	32650321	ASGA0062526	AG	A	G	0,288	C	1	3	64	66	TC	SNAPE	0,750	4
14	36866634	MARC0075655	AG	A	G	0,041	T	1	2	62	59	CT	SNAPE	0,333	3
14	42567970	H3GA0039888	GA	G	A	0,137	C	3	0	65	65	C	-	1,000	3
14	42997445	INRA0043717	AG	A	G	0,110	T	0	3	59	59	C	SNAPE	0,000	3
14	44167271	DIAS0000018	AG	A	G	0,417	T	0	3	62	62	C	SNAPE	0,000	3
14	45192379	MARC0094177	A	-	A	0,000	T	4	0	50	50	T	-	0,000	4
14	45642051	ASGA0063071	A	-	A	0,000	C	0	3	66	66	T	SNAPE	0,000	3
14	48105182	ASGA0063188	CA	C	A	0,267	A	4	2	67	66	AC	SNAPE	0,333	6
14	49674562	MARC0008762	AG	A	G	0,164	T	0	4	66	66	C	SNAPE	0,000	4
14	59663011	MARC0109427	GA	G	A	0,425	A	1	2	53	65	GA	SNAPE	0,667	3
14	59986607	MARC0009335	AC	A	C	0,034	G	5	0	66	66	G	-	0,000	5
14	62975744	H3GA0040465	AG	A	G	0,397	C	1	2	60	63	TC	SNAPE	0,667	3
14	64116421	MARC0004193	AG	A	G	0,473	C	3	0	63	63	C	-	0,000	3
14	68934841	ALGA0078379	GA	G	A	0,219	G	1	3	66	57	AG	SNAPE	0,250	4
14	70086888	ALGA0078434	AG	A	G	0,107	C	1	2	62	68	TC	SNAPE	0,667	3
14	71288826	MARC0002437	AT	A	T	0,137	A	3	0	60	60	A	-	0,000	3
14	78600780	ASGA0106183	GA	G	A	0,164	C	1	2	45	68	TC	SNAPE	0,333	3
14	93130879	MARC0110415	A	-	A	0,000	A	3	0	57	57	A	-	0,000	3
14	97285338	MARC0027760	AG	A	G	0,205	G	5	1	64	62	GA	-	0,167	6
14	100397581	INRA0045780	GA	G	A	0,116	C	0	3	52	52	T	SNAPE	0,000	3
14	108604082	ALGA0080347	CG	C	G	0,472	G	4	1	61	64	GC	-	0,200	5
14	110441918	ALGA0080466	AG	A	G	0,336	C	1	2	61	70	TC	SNAPE	0,667	3

14	110514320	ALGA0080476	AG	A	G	0,035	G	3	0	65	65	G	-	0,000	3
14	111814949	ALGA0080560	GA	G	A	0,493	A	2	1	67	66	AG	-	0,333	3
14	115580357	H3GA0041844	CA	C	A	0,199	A	1	2	67	68	CA	SNAPE	0,667	3
14	123214577	ASGA0066177	GA	G	A	0,192	C	0	3	67	67	T	SNAPE	0,000	3
14	124879995	ASGA0066251	AG	A	G	0,315	G	3	0	58	58	G	-	0,000	3
14	134601907	MARC0011591	GA	G	A	0,137	G	0	3	54	54	A	SNAPE	0,000	3
14	136565684	MARC0016693	CA	C	A	0,301	A	0	3	67	67	C	SNAPE	1,000	3
14	138428984	DRGA0014684	AG	A	G	0,123	T	1	3	52	58	CT	SNAPE	0,250	4
14	140674302	ALGA0082286	AG	A	G	0,233	C	2	1	68	68	CT	-	0,333	3
14	143936615	ASGA0067340	GA	G	A	0,336	T	3	0	70	70	T	-	0,000	3
14	147687705	MARC0112974	A	-	A	0,000	A	4	0	61	61	A	-	0,000	4
15	22879717	CASI0010169	GA	G	A	0,370	A	2	2	64	69	AG	SNAPE	0,500	4
15	23522584	MARC0114005	AG	A	G	0,493	C	2	1	52	63	CT	-	0,333	3
15	30264695	ASGA0098906	AG	A	G	0,288	T	0	3	57	57	C	SNAPE	0,000	3
15	32083260	MARC0009494	AC	A	C	0,264	C	2	2	63	64	CA	SNAPE	0,500	4
15	35228926	ALGA0110960	GA	G	A	0,282	T	3	0	66	66	T	-	0,000	3
15	36902576	ASGA0100430	GA	G	A	0,055	C	0	3	69	69	T	SNAPE	0,000	3
15	39450304	ASGA0069333	AG	A	G	0,178	C	4	1	65	66	CT	-	0,200	5
15	40821463	ALGA0084968	AG	A	G	0,130	G	3	0	69	69	G	-	0,000	3
15	52619034	MARC0002102	AG	A	G	0,404	G	1	2	59	61	AG	SNAPE	0,667	3
15	65558414	ALGA0085654	GA	G	A	0,273	C	1	2	51	58	TC	SNAPE	0,333	3
15	65787941	ASGA0069769	GA	G	A	0,391	C	2	3	66	63	TC	SNAPE	0,400	5
15	84287331	MARC0089234	GA	G	A	0,007	G	0	3	52	52	A	SNAPE	0,000	3
15	87163603	MARC0074414	GA	G	A	0,041	A	3	0	65	65	A	-	0,000	3
15	91900610	MARC0011192	AG	A	G	0,459	G	0	3	62	62	A	SNAPE	1,000	3
15	107335694	INRA0049898	GA	G	A	0,116	T	3	0	64	64	T	-	0,000	3
15	108470002	INRA0049906	AG	A	G	0,079	T	0	3	65	65	C	SNAPE	0,000	3
15	108549730	INRA0049907	AG	A	G	0,062	G	4	1	66	63	GA	-	0,200	5
15	110011380	ASGA0070210	GA	G	A	0,007	A	4	0	62	62	A	-	0,000	4
15	114055631	MARC0089468	CG	C	G	0,370	C	2	1	66	70	CG	-	0,333	3
15	119982356	H3GA0044814	AG	A	G	0,356	A	2	3	67	61	GA	SNAPE	0,400	5
15	125941914	MARC0044293	AC	-	-	-	A	3	0	51	51	A	-	1,000	3
15	134290727	INRA0050279	AG	A	G	0,027	A	0	4	62	62	G	SNAPE	0,000	4
15	136689942	MARC0028230	AG	A	G	0,110	C	4	0	64	64	C	-	0,000	4
15	142047399	ASGA0091187	GA	G	A	0,425	G	0	3	60	60	A	SNAPE	0,000	3

15	146927575	H3GA0045481	GA	G	A	0,185	T	3	0	58	58	T	-	0,000	3
15	147460703	ALGA0109945	AG	A	G	0,274	T	0	3	46	46	C	SNAPE	0,000	3
15	155782508	MARC0009013	A	-	A	0,000	A	3	0	57	57	A	-	0,000	3
16	12567717	MARC0004453	AC	A	C	0,389	G	0	3	56	56	T	SNAPE	1,000	3
16	17297502	MARC0031868	CA	C	A	0,341	A	2	1	61	68	AC	-	0,333	3
16	21235965	M1GA0020918	GA	G	A	0,075	A	2	1	68	69	AG	-	0,333	3
16	21506197	MARC0081954	GA	G	A	0,315	T	2	2	56	58	TC	SNAPE	0,500	4
16	33866754	MARC0006026	AC	A	C	0,240	G	4	0	63	63	G	-	0,000	4
16	35614082	ASGA0072998	AG	A	G	0,295	C	3	0	61	61	C	-	0,000	3
16	35915624	ASGA0073016	AG	A	G	0,041	C	3	0	66	66	C	-	0,000	3
16	39433001	MARC0068395	GA	G	A	0,034	T	3	0	57	57	T	-	0,000	3
16	44481215	MARC0073104	GA	G	A	0,479	G	0	3	67	67	A	SNAPE	0,000	3
16	48319008	MARC0005524	G	-	G	0,000	G	4	0	58	58	G	-	0,000	4
16	56950872	H3GA0046663	AC	A	C	0,500	A	2	1	60	59	AC	-	0,667	3
16	58118912	M1GA0021048	AG	A	G	0,438	C	2	1	46	52	CT	-	0,333	3
16	71638433	ALGA0091340	AC	A	C	0,219	C	4	0	69	69	C	-	0,000	4
16	71676371	MARC0014855	GA	G	A	0,493	C	0	3	60	60	T	SNAPE	0,000	3
16	72486510	MARC0033850	GA	G	A	0,151	A	0	3	60	60	G	SNAPE	1,000	3
16	74783148	H3GA0047032	GA	G	A	0,205	T	2	2	64	65	TC	SNAPE	0,500	4
16	75044752	DIAS0003143	GA	G	A	0,418	G	1	3	66	68	AG	SNAPE	0,250	4
16	75656200	ASGA0074153	AG	A	G	0,007	G	4	0	63	63	G	-	0,000	4
16	79672812	ALGA0091949	GA	G	A	0,329	C	3	0	63	63	C	-	1,000	3
16	81072970	MARC0051680	GA	G	A	0,130	A	2	1	65	68	AG	-	0,333	3
16	82232964	ALGA0092202	AG	A	G	0,068	G	4	0	60	60	G	-	0,000	4
17	5735844	MARC0101036	GA	G	A	0,313	T	3	0	65	65	T	-	0,000	3
17	9137676	MARC0018046	AG	A	G	0,404	A	2	2	63	56	AG	-	0,500	4
17	9700086	MARC0050729	TC	-	-	-	G	3	4	66	65	AG	SNAPE	0,429	7
17	11263527	ALGA0093187	CA	C	A	0,123	A	3	0	59	59	A	-	0,000	3
17	11637994	ALGA0118709	GA	G	A	0,151	C	0	3	67	67	T	SNAPE	0,000	3
17	12380965	H3GA0047800	TC	-	-	-	T	0	3	65	65	C	SNAPE	0,000	3
17	12877757	ALGA0093121	GA	G	A	0,315	T	3	0	59	59	T	-	0,000	3
17	16899381	INRA0052777	AG	-	-	-	C	0	3	68	68	T	SNAPE	0,000	3
17	17479387	MARC0070553	TG	-	-	-	T	3	1	65	66	TG	-	0,750	4
17	19005250	ALGA0093481	AG	A	G	0,205	C	3	0	66	66	C	-	0,000	3
17	23405226	ALGA0093699	AG	A	G	0,486	T	1	2	51	60	CT	SNAPE	0,333	3

17	26310337	INRA0053127	AC	A	C	0,062	T	0	3	62	62	G	SNAPE	0,000	3
17	28194155	DRGA0016669	AG	A	G	0,486	T	1	4	58	66	CT	SNAPE	0,200	5
17	30048261	ALGA0094022	CA	C	A	0,459	T	3	0	57	57	T	-	0,000	3
17	31436076	ALGA0094114	AG	A	G	0,336	G	2	1	62	51	GA	-	0,333	3
17	33911546	MARC0099174	CA	C	A	0,171	C	1	2	49	61	AC	SNAPE	0,333	3
17	33925357	MARC0115013	AC	A	C	0,418	A	1	3	63	68	CA	SNAPE	0,250	4
17	34917396	ASGA0076317	CA	C	A	0,274	A	4	1	62	68	AC	-	0,200	5
17	35644668	ALGA0094407	AC	A	C	0,486	C	6	2	68	65	CA	SNAPE	0,250	8
17	44622084	ASGA0076872	AG	A	G	0,007	C	5	0	61	61	C	-	0,000	5
17	47595840	ALGA0115746	GA	G	A	0,007	T	4	0	63	63	T	-	0,000	4
17	53166870	MARC0068151	GA	G	A	0,199	C	1	4	56	63	TC	SNAPE	0,200	5
17	53489947	MARC0039024	A	-	A	0,000	T	3	0	62	62	T	-	0,000	3
17	57271405	MARC0047625	TC	-	-	-	C	1	2	52	58	GC	SNAPE	0,333	3
17	64350574	M1GA0022391	AG	A	G	0,336	G	3	1	60	66	GA	-	0,250	4
17	64792101	MARC0040093	G	-	G	0,000	C	3	0	69	69	C	-	0,000	3
17	69302804	M1GA0022894	AG	A	G	0,240	G	1	2	65	65	AG	SNAPE	0,667	3
18	528383	ASGA0100292	G	-	G	0,000	C	5	0	61	61	C	-	0,000	5
18	3080751	ASGA0096637	CA	C	A	0,096	G	1	2	62	64	TG	SNAPE	0,333	3
18	7169700	MARC0045826	AG	A	G	0,390	C	1	2	67	65	TC	SNAPE	0,667	3
18	8792587	ASGA0087469	GA	G	A	0,247	T	3	0	58	58	T	-	0,000	3
18	9044109	MARC0007516	AG	A	G	0,308	C	3	0	69	69	C	-	0,000	3
18	13420674	ALGA0097060	GA	G	A	0,199	T	5	0	65	65	T	-	0,000	5
18	14790355	MARC0086791	AG	A	G	0,384	G	2	4	66	57	AG	SNAPE	0,667	6
18	17477861	H3GA0050438	AG	A	G	0,034	C	3	0	66	66	C	-	0,000	3
18	18507823	ALGA0097258	AG	A	G	0,329	G	4	0	65	65	G	-	0,000	4
18	21375539	H3GA0050524	A	-	A	0,000	T	3	0	66	66	T	-	0,000	3
18	31024686	MARC0005136	AC	A	C	0,110	C	3	0	60	60	C	-	0,000	3
18	34055442	MARC0065378	AC	A	C	0,267	G	3	2	63	42	GT	-	0,400	5
18	35538731	DIAS0002798	GA	G	A	0,459	A	2	2	65	60	AG	SNAPE	0,500	4
18	43807494	MARC0001169	TC	-	-	-	C	4	0	67	67	C	-	1,000	4
18	44492656	ALGA0106178	CA	C	A	0,295	T	5	0	61	61	T	-	0,000	5
18	49618941	ASGA0098397	CA	C	A	0,336	A	2	1	55	65	AC	-	0,333	3
18	51542502	ALGA0098540	GA	G	A	0,144	T	5	0	58	58	T	-	0,000	5
18	52790382	MARC0057774	AG	A	G	0,397	A	3	3	66	62	AG	SNAPE	0,500	6
18	57050998	ASGA0080380	AG	A	G	0,212	G	1	2	62	71	AG	SNAPE	0,667	3

18	58861041	ALGA0098906	GA	G	A	0,171	A	2	2	71	67	AG	SNAPE	0,500	4
GL892106.1	193	MARC0073396	TC	-	-	-	G	4	1	68	64	GA	-	0,800	5
GL892163.1	333	ASGA0025061	GA	G	A	0,158	A	5	0	68	68	A	-	0,000	5
GL892362.2	13276	MARC0052564	AC	A	C	0,116	T	4	0	53	53	T	-	1,000	4
GL892482.1	112782	UMB10000102	AG	A	G	0,158	G	3	0	65	65	G	-	0,000	3
GL892567.1	163	MARC0033979	GA	G	A	0,452	C	3	0	52	52	C	-	1,000	3
GL892686.1	388	H3GA0005146	AG	A	G	0,113	T	0	3	64	64	C	SNAPE	0,000	3
GL892697.1	118	ALGA0117067	AG	A	G	0,336	A	3	0	63	63	A	-	1,000	3
GL893071.1	24967	ALGA0105887	AG	A	G	0,240	C	4	0	63	63	C	-	0,000	4
GL893166.1	13209	ASGA0036883	GA	G	A	0,329	A	2	1	66	56	AG	-	0,333	3
GL893171.1	4267	MARC0054451	AG	A	G	0,411	A	3	0	64	64	A	-	1,000	3
GL893222.2	223546	MARC0028053	CA	C	A	0,151	G	5	0	63	63	G	-	1,000	5
GL893581.2	56372	ALGA0122025	AG	A	G	0,178	T	1	2	62	51	CT	-	0,333	3
GL893618.1	61456	ASGA0047433	AG	A	G	0,274	A	1	3	70	69	GA	SNAPE	0,250	4
GL893704.1	8389	ASGA0095354	GA	G	A	0,034	A	3	0	62	62	A	-	0,000	3
GL893724.1	11172	MARC0082449	AC	A	C	0,299	G	3	0	60	60	G	-	0,000	3
GL893762.1	34201	MARC0014151	AG	A	G	0,267	A	1	2	56	67	GA	SNAPE	0,333	3
GL893809.1	12772	ASGA0083391	AG	A	G	0,247	G	3	0	57	57	G	-	0,000	3
GL893871.1	40980	ASGA0102258	GA	G	A	0,205	C	3	0	62	62	C	-	1,000	3
GL894055.2	45867	MARC0042009	AG	A	G	0,349	T	1	2	65	51	CT	-	0,333	3
GL894259.1	38956	ALGA0117568	GA	G	A	0,349	C	1	3	65	65	TC	SNAPE	0,250	4
GL894737.1	17342	MARC0053028	GA	G	A	0,158	A	2	2	62	62	AG	SNAPE	0,500	4
GL894759.2	7765	MARC0089581	AG	A	G	0,014	G	4	0	65	65	G	-	0,000	4
GL894979.1	1970	H3GA0056755	AG	A	G	0,130	A	0	3	67	67	G	SNAPE	0,000	3
GL894987.2	32906	ALGA0019723	AG	A	G	0,103	C	2	1	66	70	CT	-	0,333	3

GL895151.1	24144	MARC0089685	GA	G	A	0,222	T	3	0	60	60	T	-	0,000	3
GL895235.1	35444	MARC0033875	CA	C	A	0,048	A	3	0	55	55	A	-	0,000	3
GL895342.1	12923	ALGA0068334	AG	A	G	0,445	C	2	1	64	62	CT	-	0,333	3
GL895357.2	25987	ASGA0097613	GA	G	A	0,425	C	0	4	64	64	T	SNAPE	0,000	4
GL895413.1	49579	MARC0005414	A	-	A	0,000	A	3	0	61	61	A	-	0,000	3
GL895475.1	7652	MARC0073362	TG	-	-	-	A	1	2	55	58	CA	SNAPE	0,333	3
GL895827.2	30893	ASGA0100325	G	-	G	0,000	T	3	0	70	70	T	-	1,000	3
GL896623.1	116	ALGA0123993	AG	A	G	0,472	G	1	2	66	65	AG	SNAPE	0,667	3
GL896671.1	133	MARC0044302	AG	-	-	-	A	3	0	63	63	A	-	1,000	3
JH118613.1	125909	ASGA0102421	CA	C	A	0,240	T	4	0	54	54	T	-	0,000	4
JH118641.1	36965	ASGA0064618	GA	G	A	0,089	C	0	4	65	65	T	SNAPE	0,000	4
JH118655.1	49612	MARC0085359	GA	G	A	0,295	A	3	0	64	64	A	-	0,000	3
JH118658.1	22739	MARC0068457	A	-	A	0,000	T	3	0	51	51	T	-	0,000	3
JH118669.1	224278	MARC0003328	GA	G	A	0,336	C	1	3	51	64	TC	SNAPE	0,250	4
JH118695.1	7745	ALGA0010577	CA	C	A	0,014	A	3	0	66	66	A	-	0,000	3
JH118753.1	22252	MARC0050114	GA	G	A	0,158	T	2	1	61	56	TC	-	0,333	3
X	3029185	ALGA0121921	AG	A	G	0,137	T	0	4	59	59	C	SNAPE	0,000	4
X	7269183	MARC0111499	AG	A	G	0,418	A	2	1	62	65	AG	-	0,667	3
X	8044093	ALGA0099273	AG	A	G	0,130	C	3	0	66	66	C	-	0,000	3
X	12508028	MARC0060456	AG	A	G	0,103	G	2	3	65	65	AG	SNAPE	0,600	5
X	12941862	MARC0097036	AG	A	G	0,212	C	2	1	55	60	CT	-	0,333	3
X	12973924	ASGA0090623	AC	A	C	0,411	C	1	2	65	67	AC	SNAPE	0,667	3
X	13152222	ASGA0080850	AG	A	G	0,301	T	1	2	64	66	CT	SNAPE	0,333	3
X	16424567	MARC0032291	AG	A	G	0,219	G	2	1	66	61	GA	-	0,333	3
X	22662992	ALGA0099483	GC	G	C	0,458	C	1	2	66	61	GC	SNAPE	0,667	3
X	34603527	ASGA0081086	AG	A	G	0,493	A	1	3	60	62	GA	SNAPE	0,250	4
X	44092699	H3GA0051764	GA	G	A	0,185	C	0	3	66	66	T	SNAPE	0,000	3
X	48628584	ALGA0099736	AG	A	G	0,445	C	2	2	59	60	CT	SNAPE	0,500	4

X	107023141	MARC0026850	GA	G	A	0,144	T	3	0	58	58	T	-	0,000	3
X	109357428	ALGA0099906	AG	A	G	0,500	G	0	3	53	53	A	SNAPE	1,000	3
X	110208365	MARC0048855	C	-	C	0,000	T	3	0	50	50	T	-	1,000	3
X	113468186	ASGA0081284	AG	A	G	0,048	C	3	0	68	68	C	-	0,000	3
X	113821548	ALGA0099984	CA	C	A	0,226	T	3	0	59	59	T	-	0,000	3
X	114564150	H3GA0051924	AG	A	G	0,178	C	3	1	64	70	CT	-	0,250	4
X	123752552	ASGA0081391	GA	G	A	0,007	A	3	0	61	61	A	-	0,000	3
X	142189065	ASGA0081616	GA	G	A	0,500	T	3	2	54	63	TC	SNAPE	0,400	5

Table S4. List of the 54 SNPs showing significant differences (P<0.05) between the two libraries

Chr = Chromosome; Pos = Nucleotide position on the Sscrofa10.2 reference genome; Reference = Reference nucleotide in the Sscrofa10.2 reference genome; SNP_NGS = Variation identified by the sequencing; N_Ref = Number of reads equal to the reference nucleotide; N_Alt = Number of reads alternative to the reference nucleotide; Q_Ref = Mean quality of the reads covering the reference nucleotide; Q_Alt = Mean quality of the reads covering the alternative nucleotide; $P_{Fischer}$ =P-value of the Fisher's test.

				POSITIVE LIBRARY				NEGATIVE LIBRARY				
Chr	**Pos**	**Reference**	**SNP_NGS**	**N_Ref**	**N_Alt**	**Q_Ref**	**Q_Alt**	**N_Ref**	**N_Alt**	**Q_Ref**	**Q_Alt**	$P_{Fischer}$
1	68514	A	AG	0	3	61	61	4	0	62	62	2,86E-02
1	80363	G	GC	0	3	65	65	5	1	61	70	4,76E-02
1	130684	A	AT	11	0	60	60	1	2	68	64	3,30E-02
1	100578902	A	AG	15	6	68	66	24	1	67	64	3,65E-02
1	100578908	G	GT	19	2	63	62	15	10	64	58	4,05E-02
1	312947644	C	CA	18	2	63	63	5	5	68	66	2,56E-02
3	142508639	G	GT	8	1	63	53	0	3	53	53	1,82E-02
3	142508640	A	AT	8	1	63	68	0	4	62	62	6,99E-03
3	142527674	A	AG	10	5	49	60	13	0	56	56	4,37E-02
3	142732127	C	CT	19	15	60	66	34	8	62	68	2,43E-02
6	171530	C	CG	44	4	58	65	55	16	59	65	4,82E-02
6	859837	T	TC	11	1	66	69	3	4	65	65	3,79E-02
6	873039	A	AG	7	0	63	63	4	5	64	66	3,37E-02
6	873061	A	AG	8	0	64	64	4	7	61	59	1,28E-02
6	875775	C	CG	13	1	50	62	2	3	38	58	3,74E-02
6	956438	G	GA	8	12	61	65	10	2	63	62	2,76E-02
6	957389	C	TC	5	13	58	65	7	2	57	51	3,69E-02
6	957407	G	AG	1	12	69	64	5	3	66	65	1,39E-02
6	67619962	G	AG	0	5	68	68	3	0	63	63	1,79E-02
6	67619975	C	TC	0	5	66	66	3	0	66	66	1,79E-02
6	67619980	G	TG	0	5	64	64	3	0	62	62	1,79E-02
8	77334615	A	GA	25	14	65	67	18	33	66	66	1,03E-02
9	16885924	A	AC	7	1	63	55	4	8	61	63	2,81E-02
9	40852106	T	TG	20	9	66	68	12	20	67	65	2,09E-02
9	40853107	G	GT	13	0	67	67	13	7	67	67	2,66E-02
10	21604345	G	TG	9	8	65	64	2	13	58	65	2,78E-02
10	76923741	G	GA	1	4	64	63	4	0	65	65	4,76E-02
12	48937212	G	GT	30	2	60	67	10	5	62	66	2,63E-02
12	48937258	A	AT	28	0	64	64	12	3	68	65	3,69E-02

12	48937265	G	GT	28	0	65	65	11	3	68	70	3,17E-02
12	48937304	G	GT	15	0	64	64	6	3	68	71	4,15E-02
15	155198719	C	CA	0	3	67	67	4	0	58	58	2,86E-02
15	155253410	C	CA	11	0	63	63	2	3	57	66	1,79E-02
15	155292938	C	CA	1	4	69	64	5	0	65	65	4,76E-02
15	155292948	C	CT	1	4	63	59	4	0	68	68	4,76E-02
15	155318428	A	GA	2	8	65	63	6	1	62	50	1,52E-02
15	155409629	A	AG	5	0	61	61	1	4	67	59	4,76E-02
15	155465431	G	GT	13	4	65	61	6	10	59	57	3,66E-02
15	155466144	C	CT	8	0	60	60	4	7	64	59	1,28E-02
15	155466188	T	TG	18	0	64	64	9	4	64	58	2,27E-02
15	155473144	C	CA	5	5	64	65	18	2	65	59	2,56E-02
16	36781804	G	GT	72	0	52	52	65	5	51	60	2,70E-02
16	36782040	C	CA	5	7	66	69	12	1	66	75	1,12E-02
16	36782170	T	GT	70	111	62	63	94	91	62	63	2,11E-02
16	36782225	G	GA	191	15	66	65	210	3	68	66	3,22E-03
16	36782626	T	TA	81	20	62	58	60	5	59	60	4,44E-02
16	36782695	T	GT	7	41	66	60	12	22	62	61	3,59E-02
16	36782771	T	TG	23	1	61	66	16	6	59	68	4,31E-02
16	36783493	A	AC	138	35	63	68	149	20	64	67	3,95E-02
16	36786592	T	CT	8	18	64	65	2	26	55	65	3,66E-02
16	36786614	A	AC	39	4	66	62	50	0	66	66	4,23E-02
16	36786814	T	TA	64	1	62	57	69	8	64	61	3,90E-02
16	36787213	C	CT	61	3	62	65	44	11	62	68	1,96E-02
16	54726105	A	AG	50	3	63	64	17	6	65	64	1,92E-02
16	54883839	C	CT	28	11	62	61	16	23	56	64	1,15E-02
16	54883877	T	TC	45	29	67	65	29	44	66	65	1,34E-02
17	69407874	C	CG	19	0	65	65	9	3	66	64	4,89E-02
17	69494538	A	AC	10	3	68	62	4	9	70	60	4,72E-02
18	225564	C	CT	24	1	57	43	10	4	63	59	4,69E-02
18	335523	G	GA	2	4	57	64	9	0	53	53	1,10E-02
18	359629	G	GA	6	1	65	68	0	3	67	67	3,33E-02
18	385298	C	CT	5	4	66	61	16	1	67	52	3,45E-02
18	391790	A	AG	8	5	61	63	11	0	62	62	4,11E-02

Table S5. List of SNPs that showed differences in number of alternative reads compared with the genome wide association results.

Chr = Chromosome; Marker = Marker in the PorcineSNP60 Beadchip; Pos_M = Nucleotide position of the marker on the Sscrofa10.2 reference genome; P_{GWAS} = P-value of association in the GWAS; Pos_{SNP} = Nucleotide position on the Sscrofa10.2 reference genome of the SNP having $P_{Fischer} < 0.05$; $P_{Fischer}$ = P-value of the Fisher's test.

Chr	**Marker**	**Pos_M**	**P_{GWAS}**	**Pos_{SNP}**	***P_{fisher}***
1	ALGA0000009	52297	2,75E-03	68514	2,86E-02
1	ALGA0000009	52297	2,75E-03	130684	3,30E-02
1	ALGA0000009	52297	2,75E-03	80363	4,76E-02
1	ALGA0000014	79763	1,74E-05	68514	2,86E-02

1	ALGA0000014	79763	1,74E-05	130684	3,30E-02
1	ALGA0000014	79763	1,74E-05	80363	4,76E-02
6	M1GA0008302	787265	1,65E-06	873061	1,28E-02
6	M1GA0008302	787265	1,65E-06	957407	1,39E-02
6	M1GA0008302	787265	1,65E-06	956438	2,76E-02
6	M1GA0008302	787265	1,65E-06	873039	3,37E-02
6	M1GA0008302	787265	1,65E-06	957389	3,69E-02
6	M1GA0008302	787265	1,65E-06	875775	3,74E-02
6	M1GA0008302	787265	1,65E-06	859837	3,79E-02
6	M1GA0008318	945991	4,41E-04	873061	1,28E-02
6	M1GA0008318	945991	4,41E-04	957407	1,39E-02
6	M1GA0008318	945991	4,41E-04	956438	2,76E-02
6	M1GA0008318	945991	4,41E-04	873039	3,37E-02
6	M1GA0008318	945991	4,41E-04	957389	3,69E-02
6	M1GA0008318	945991	4,41E-04	875775	3,74E-02
6	M1GA0008318	945991	4,41E-04	859837	3,79E-02
6	M1GA0008329	996248	9,35E-05	873061	1,28E-02
6	M1GA0008329	996248	9,35E-05	957407	1,39E-02
6	M1GA0008329	996248	9,35E-05	956438	2,76E-02
6	M1GA0008329	996248	9,35E-05	873039	3,37E-02
6	M1GA0008329	996248	9,35E-05	957389	3,69E-02
6	M1GA0008329	996248	9,35E-05	875775	3,74E-02
6	M1GA0008329	996248	9,35E-05	859837	3,79E-02
9	DRGA0009307	17138159	8,66E-04	16885924	2,81E-02
12	DIAS0000309	48865200	9,96E-04	48937212	2,63E-02
12	DIAS0000309	48865200	9,96E-04	48937265	3,17E-02
12	DIAS0000309	48865200	9,96E-04	48937258	3,69E-02
12	DIAS0000309	48865200	9,96E-04	48937304	4,15E-02

General conclusions

In this work we applied different approaches to identify DNA markers associated to production traits. In particular, we focused the attention on Italian Large White pig breed using GWAS and applying a selective genotyping approach to reduce the number of animals to be genotyped and to increase the power of the analyses. Two traits were targeted, namely Back Fat Thickness and Average Daily Gain, and several regions affecting these traits were identified. For Back Fat Thickness, we identified some Copy Number Variant Regions that may affect fat deposition.

In addition to already developed genotyping platforms, we searched the pig genome to identify new markers potentially affecting production traits exploring Next Generation Sequencing. We used the Ion Torrent Technology, based on a semiconductor detection of pH variation during the sequencing process, and sequenced our samples with a Ion 318 sequencing chip, that can provide up to 1Gb of sequences of 200bp length. In our knowledge this is the first time that the selective genotyping approach and deep sequencing have been combined for SNP discovery.

Other two studies were carried on with a completely different approach. Allele frequency changes for SNPs affecting candidate genes and at Genome Wide level were analysed to identify selection signatures driven by selection program during the last 20 years.

This approach confirmed that a great number of markers may affect production traits and that they are captured by the classical selection programs.

Genome Wide Association Studies works, both for the association with BFT and ADG, showed the emergency of new patterns of genes involved in the differences between fat and lean pigs, and between rapidly growing or slower growing animals. GWAS revealed 123 significant or suggestively significant SNP associated with Backfat Thickenss and 229 associated with Average Daily Gain.

The detection of 16 Copy Number Variant Regions that resulted more frequent in lean or fat pigs showed that different copies of those region could have a limited impact on fat deposition but suggests that a further analysis on low frequency events on a larger set of animal could be useful to increase the power. These regions often appear to be involved in food intake and behaviour, beside affecting genes involved in metabolic pathways and their expression.

By combining RRLs semi-conductor sequencing with selective genotyping approach, new short variants where discovered and at least 54 are worth to be analysed in association studies.

The study of groups of pigs born in different years in a population undergone to stringent selection showed that, in a very short time, the allele frequency of some loci can drastically change if they are close to traits that are interesting for selection schemes. Four polymorphic sites in genes known to affect lean cuts, ham weight, feed:gain ratio and backfat thickness (IGF2, MC4R, VRTN and FTO) showed significant ($P<0.01$) changes in allele frequencies over time due to a progressive and continuous increase of one allele. The study of genotypes frequency shifts at a Genome Wide level in the above described groups revealed genes and pathways that could not been expected from previous knowledge. This idea could also be consistent with the presence of still unrevealed genes of functional pathways in those regions.

All these experiments need to be better understood, studied and validated. It will be also worth to study in detail, at high resolution, the sequences sorrounding the above significant SNPs and one possible way will be to analyse Next Generation Sequencing data at high depth and read quality, since in this way it will be possible to observe all the set of involved SNPs, haplotypes and, in the best case, the exact location of causative alleles.

In future we will perform the application of NGS techniques to all animals and the information will be integrated in genomic selection plans.

General references

Boman IA, Klemetsdal G, Nafstad O, Blichfeldt T, Våge DI. Selection based on progeny testing induces rapid changes in myostatin allele frequencies - a case study in sheep. J Anim Breed Genet. 2011 Feb;128(1):52-5. doi: 10.1111/j.1439-0388.2010.00879.x. Epub 2010 Sep 22. PubMed PMID: 21214644.

Corominas J, Ramayo-Caldas Y, Puig-Oliveras A, Pérez-Montarelo D, Noguera JL, Folch JM, Ballester M. Polymorphism in the ELOVL6 gene is associated with a major QTL effect on fatty acid composition in pigs. PLoS One. 2013;8(1):e53687. doi:10.1371/journal.pone.0053687. Epub 2013 Jan 14. PubMed PMID: 23341976; PubMed Central PMCID: PMC3544903.

Corominas J, Ramayo-Caldas Y, Castelló A, Muñoz M, Ibáñez-Escriche N, Folch JM, Ballester M. Evaluation of the porcine ACSL4 gene as a candidate gene for meat quality traits in pigs. Anim Genet. 2012 Dec;43(6):714-20. doi:10.1111/j.1365-2052.2012.02335.x. Epub 2012 Mar 11. PubMed PMID: 22497636.

Fiona Cunningham, M. Ridwan Amode, Daniel Barrell, Kathryn Beal, Konstantinos Billis, Simon Brent, Denise Carvalho-Silva, Peter Clapham, Guy Coates, Stephen Fitzgerald, Laurent Gil, Carlos Garcín Girón, Leo Gordon, Thibaut Hourlier, Sarah E. Hunt, Sophie H. Janacek, Nathan Johnson, Thomas Juettemann, Andreas K. Kähäri, Stephen Keenan, Fergal J. Martin, Thomas Maurel, William McLaren, Daniel N. Murphy, Rishi Nag, Bert Overduin, Anne Parker, Mateus Patricio, Emily Perry, Miguel Pignatelli, Harpreet Singh Riat, Daniel Sheppard, Kieron Taylor, Anja Thormann, Alessandro Vullo, Steven P. Wilder, Amonida Zadissa, Bronwen L. Aken, Ewan Birney, Jennifer Harrow, Rhoda Kinsella, Matthieu Muffato, Magali Ruffier, Stephen M.J. Searle,

Giulietta Spudich, Stephen J. Trevanion, Andy Yates, Daniel R. Zerbino and Paul Flicek. Ensembl (2015), Nucleic Acids Research. doi: 10.1093/nar/gku1010

Groenen MA, Archibald AL, Uenishi H, Tuggle CK, Takeuchi Y, Rothschild MF, Rogel-Gaillard C, Park C, Milan D, Megens HJ, Li S, Larkin DM, Kim H, Frantz LA, Caccamo M, Ahn H, Aken BL, Anselmo A, Anthon C, Auvil L, Badaoui B, Beattie CW, Bendixen C, Berman D, Blecha F, Blomberg J, Bolund L, Bosse M, Botti S, Bujie Z, Bystrom M, Capitanu B, Carvalho-Silva D, Chardon P, Chen C, Cheng R, Choi SH, Chow W, Clark RC, Clee C, Crooijmans RP, Dawson HD, Dehais P, De Sapio F, Dibbits B, Drou N, Du ZQ, Eversole K, Fadista J, Fairley S, Faraut T, Faulkner GJ, Fowler KE, Fredholm M, Fritz E, Gilbert JG, Giuffra E, Gorodkin J, Griffin DK, Harrow JL, Hayward A, Howe K, Hu ZL, Humphray SJ, Hunt T, Hornshøj H, Jeon JT, Jern P, Jones M, Jurka J, Kanamori H, Kapetanovic R, Kim J, Kim JH, Kim KW, Kim TH, Larson G, Lee K, Lee KT, Leggett R, Lewin HA, Li Y, Liu W, Loveland JE, Lu Y, Lunney JK, Ma J, Madsen O, Mann K, Matthews L, McLaren S, Morozumi T, Murtaugh MP, Narayan J, Nguyen DT, Ni P, Oh SJ, Onteru S, Panitz F, Park EW, Park HS, Pascal G, Paudel Y, Perez-Enciso M, Ramirez-Gonzalez R, Reecy JM, Rodriguez-Zas S, Rohrer GA, Rund L, Sang Y, Schachtschneider K, Schraiber JG, Schwartz J, Scobie L, Scott C, Searle S, Servin B, Southey BR, Sperber G, Stadler P, Sweedler JV, Tafer H, Thomsen B, Wali R, Wang J, Wang J, White S, Xu X, Yerle M, Zhang G, Zhang J, Zhang J, Zhao S, Rogers J, Churcher C, Schook LB. Analyses of pig genomes provide insight into porcine demography and evolution. Nature. 2012 Nov 15;491(7424):393-8. doi: 10.1038/nature11622. PubMed PMID: 23151582; PubMed Central PMCID: PMC3566564.

Marklund S, Kijas J, Rodriguez-Martinez H, Rönnstrand L, Funa K, Moller M, Lange D, Edfors-Lilja I, Andersson L. Molecular basis for the dominant white phenotype in the domestic pig. Genome Res. 1998 Aug;8(8):826-33. PubMed PMID: 9724328; PubMed Central PMCID: PMC310759.

Ramayo-Caldas Y, Mercadé A, Castelló A, Yang B, Rodríguez C, Alves E, Díaz I, Ibáñez-Escriche N, Noguera JL, Pérez-Enciso M, Fernández AI, Folch JM. Genome-wide association study for intramuscular fatty acid composition in an Iberian × Landrace cross. J Anim Sci. 2012 Sep;90(9):2883-93. doi: 10.2527/jas.2011-4900. Epub 2012 Jul 10. PubMed PMID: 22785162.

Ramayo-Caldas Y, Castelló A, Pena RN, Alves E, Mercadé A, Souza CA, Fernández AI, Perez-Enciso M, Folch JM. Copy number variation in the porcine genome inferred from a 60 k SNP BeadChip. BMC Genomics. 2010 Oct 22;11:593. doi: 10.1186/1471-2164-11-593. PubMed PMID: 20969757; PubMed Central PMCID: PMC3091738.

Rubin CJ, Megens HJ, Martinez Barrio A, Maqbool K, Sayyab S, Schwochow D, Wang C, Carlborg Ö, Jern P, Jørgensen CB, Archibald AL, Fredholm M, Groenen MA, Andersson L. Strong signatures of selection in the domestic pig genome. Proc Natl Acad Sci U S A. 2012 Nov 27;109(48):19529-36. doi: 10.1073/pnas.1217149109. Epub 2012 Nov 14. PubMed PMID: 23151514; PubMed Central PMCID: PMC3511700.

Wang J, Wang H, Jiang J, Kang H, Feng X, Zhang Q, Liu JF. Identification of genome-wide copy number variations among diverse pig breeds using SNP genotyping arrays. PLoS One. 2013 Jul 23;8(7):e68683. doi: 10.1371/journal.pone.0068683. Print 2013. PubMed PMID: 23935880; PubMed Central PMCID: PMC3720780.