ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics

doi: 10.1111/age.12641

A targeted genotyping approach enhances identification of variants in taste receptor and appetite/reward genes of potential functional importance for obesity-related porcine traits

S. Cirera*[1], A. Clop[†1], M. J. Jacobsen*, M. Guerin[‡], P. Lesnik[‡], C. B. Jørgensen*, M. Fredholm* and P. Karlskov-Mortensen*

*Department of Veterinary and Animal Sciences, Faculty of Health and Medical Sciences, University of Copenhagen, Frederiksberg, Denmark. [†]CSIC-IRTA-UAB-UB, Campus UAB, Centre for Research in Agricultural Genomics (CRAG), 08193 Cerdanyola del Valles, Catalonia, Spain. [‡]INSERM, UMRS U1166, team 4 "Integrative Biology of Atherosclerosis", Sorbonne Universités, UPMC Université Paris 6, Paris, France.

Summary

Taste receptors (TASRs) and appetite and reward (AR) mechanisms influence eating behaviour, which in turn affects food intake and risk of obesity. In a previous study, we used next generation sequencing to identify potentially functional mutations in TASR and AR genes and found indications for genetic associations between identified variants and growth and fat deposition in a subgroup of animals ($n = 38$) from the UNIK resource pig population. This population was created for studying obesity and obesity-related diseases. In the present study we validated results from our previous study by investigating genetic associations between 24 selected single nucleotide variants in TASR and AR gene variants and 35 phenotypes describing obesity and metabolism in the entire UNIK population ($n = 564$). Fifteen variants showed significant association with specific obesity-related phenotypes after Bonferroni correction. Six of the 15 genes, namely *SIM1*, *FOS*, *TAS2R4*, *TAS2R9*, *MCHR2* and *LEPR*, showed good correlation between known biological function and associated phenotype. We verified a genetic association between potentially functional variants in TASR/AR genes and growth/obesity and conclude that the combination of identification of potentially functional variants by next generation sequencing followed by targeted genotyping and association studies is a powerful and cost-effective approach for increasing the power of genetic association studies.

Keywords eating behavior, genetic association, GWAS, metabolism, pig, single nucleotide polymorphisms

Introduction

Growth and obesity are partially regulated by a set of genes modulating the sense of taste and energy homeostasis. The perception of taste in the taste buds of the tongue is carried out mainly by taste receptors (TASRs). The selective activation of these receptors indicates whether a food is nutritious and safe. The set of TASRs includes two groups of G-protein coupled proteins, the TAS1Rs (TAS1R1, TAS1R2 and TAS1R3), which sense the palatable and energy-rich sweet and umami tastes produced by sugars and amino acids, and the TAS2Rs, including a large, diverse and highly polymorphic catalogue of genes that detect the bitter tastes associated with toxins and results in the rejection of potentially poisonous food (Breslin 2013). In the gut, TASRs sense the ingested molecules and regulate the appetite and reward (AR) circuitries and, consequently, regulate eating behaviour, food transit time, digestion and nutrient absorption (Kiuchi *et al.* 2006; Loper *et al.* 2015). The AR mechanisms involve the gastro-intestinal tract, pancreas, liver, brain, muscle and adipose tissues by engaging nutrients, neuropeptides, neurotransmitters, hormones and their related receptors and enzymes. These networks inhibit or excite the dopamine, epinephrine, norepinephrine, serotonin and glutamate receptor pathways and ultimately modulate food intake and energy

Address for correspondence

P. Karlskov-Mortensen, Department of Veterinary and Animal Sciences, Faculty of Health and Medical Sciences, University of Copenhagen, Frederiksberg, Denmark.
E-mail: pkm@sund.ku.dk

[1]These authors have contributed equally.

Accepted for publication 18 December 2017

balance (Fulton 2010; Loper *et al.* 2015). These mechanisms, when unbalanced, can promote obesity, metabolic diseases and diabetes (Berg & Kaunitz 2016; Chao *et al.* 2016). Hence, DNA polymorphisms altering the coding sequence of TASRs or AR genes could alter taste preferences and food intake with a resulting impact on energy balance, obesity and the outcome of diabetes or other metabolic disorders (Schembre *et al.* 2013; Clop *et al.* 2016).

Pigs and humans share many anatomical, physiological, metabolic, genetic and patho-physiological similarities. For this reason, the pig is a highly relevant model in which to study human conditions like obesity (Koopmans & Schuurman 2015). In the last decades the pig has been used extensively in biomedical research to investigate different aspects of obesity and its co-morbidities; from diet interventions studies using specific pig breeds (Göttingen Minipig, Ossabaw, production pigs) to more genetic-focused studies using custom-made pedigrees to gain basic knowledge about genes and pathways involved in obesity traits (Houpt *et al.* 1979; Andersson 1996; Pomp 1997; Spurlock & Gabler 2008). Recently, an F2 resource population designed to elucidate the genetics involved in development of obesity and obesity-related phenotypes has been established: the UNIK population (Pant *et al.* 2015). In this population, a total of 229 quantitative trait loci (QTL) associated with adiposity and metabolic phenotypes have been identified at genome-wide significance levels by genotyping single nucleotide polymorphisms (SNPs) on the Illumina 60K SNPchip and using a genome-wide linkage disequilibrium linkage analysis (LDLA) mapping approach.

Several studies have employed next generation sequencing methods to identify genetic variation in porcine TASR and AR genes. Two of these studies were aimed at the identification of differences in TASR genes between domestic and wild pig (da Silva *et al.* 2014; Clop *et al.* 2016; Ribani *et al.* 2017). In these studies, more than 300 coding variants were found. The allele frequencies of 13 SNPs in five TASR genes were shown to be significantly different between two pools of Large White pigs grouped according to extreme and divergent estimated breeding values for backfat thickness and, thus, deemed to be associated with backfat thickness (Ribani *et al.* 2017). Similarly, we previously conducted a study to identify variations in 10 TASR and 191 AR genes by targeted exome sequencing in more than 300 pigs from different breeds (Clop *et al.* 2016). A total of 38 F2 UNIK animals selected for extreme values for daily weight gain and retroperitoneal fat content were included in this study, and nine variants in five TASR genes and 57 variants in 26 AR genes displayed significant differences in the predicted allele frequencies (pAFs) between the phenotypically divergent F2 pools indicating genetic association between these variants and growth and fat content (Clop *et al.* 2016).

The aim of the present work was to further demonstrate in a large cohort the associations between some of the most pertinent variants in TASR and AR genes and obesity-related traits that were discovered in our previous study. Thirty SNPs in TASR and AR genes that showed significant difference in pAF in the two phenotypically divergent F2 pools were selected and genotyped in the entire UNIK pedigree ($n = 564$), and the genetic associations between successfully genotyped SNPs and 35 phenotypes related to obesity and metabolism were evaluated.

Materials and methods

Animals used in the study

We previously established an F2 pig resource population specifically designed to investigate the genetic mechanisms involved in obesity by crossing the obesity prone Göttingen Minipig breed with two lean production pig breeds: Duroc and Yorkshire. This population has been described elsewhere (Kogelman *et al.* 2013). Genotyping was performed in a total of 564 animals (23, 87 and 454 animals belonging to the parental, F1 and F2 generations respectively).

Phenotypes included in the association studies

Thirty-five different metabolic and physical obesity-related phenotypes were included in the association study (for a list of the phenotypes and their abbreviations see Table 1). The collection of phenotypes has been described elsewhere (Kogelman *et al.* 2013; Pant *et al.* 2015). Blood samples were collected, and different phenotypes were measured at three time points: at about 2 months of age (63 ± 10 days = Age1), at approximately 7 months of age (218 ± 45 days = Age2) and at slaughter (242 ± 48 days = Age3). Concentrations of total cholesterol (CT), triglycerides (TG) and high density lipoprotein cholesterol (HDL-C) in plasma were determined using standard techniques. Low density lipoprotein cholesterol (LDL-C) levels were calculated using the Friedewald formula. Body composition was determined after weaning using dual-energy x-ray absorptiometry (DEXA) scanning at Age1. Plasma CETP (cholesteryl ester transfer protein) activity (CETP_per) was assayed using the method described by Guerin *et al.* (1994).

Criteria for selection of SNPs for genotyping

Among the SNPs identified by Clop *et al.* (2016), eight SNPs in TASR genes were selected based on their predicted impact on protein function according to SIFT and SNPEFF (Cingolani *et al.* 2012). Additionally, among SNPs in AR genes, 22 SNPs that displayed significant differences in the pAFs between the phenotypically divergent sample pools in the study by Clop *et al.* (2016) were selected. Hence, 30 SNPs were selected for genotyping (see Table 2). It was ensured that all selected SNPs segregated at pAF of the minor allele at 0.08 or greater in the 38 UNIK animals that had been

Table 1 Phenotypes included in the study.

Obesity phenotypes	Phenotype description	Model covariates
Birth_wgt	Birth weight (measured in kgs)	Sex
ADG	Average daily gain from birth to age 3	Sex, Age3
BF1	Thickness of subcutaneous adipose tissue in lower trunk (mm) at age 3	Sex, Age3, $(Age3)^2$
BF2	Thickness of subcutaneous adipose tissue in upper trunk (mm) at age 3	Sex, Age3, $(Age3)^2$
BAI_Age2	Body adiposity index, measured at age 2	Sex, Age2, $(Age2)^2$
BAI_Age1	Body adiposity index, measured at age 1	Sex, Age1
BMI_Age2	Body mass index, measured at age 2	Sex, Age2, $(Age2)^2$
BMI_Age1	Body mass index, measured at age 1	Sex, Age1
DG1	Average daily weight gain from birth to age 1 (weight in kgs/day)	Sex, Age1
DG2	Average daily weight gain from age 1 to age 2 (weight in kgs/day)	Sex, Age2, $(Age2)^2$
Retro_fat	Blunt removal of retroperitoneal fat (weight in kgs)	Sex, Age3, $(Age3)^2$, Length at Age3
Ome_fat	Blunt removal of greater omentum (weight in grams)	Sex, Age3, $(Age3)^2$, Length at Age3
Int_fat	Intestinal fat	Sex, Age3
Tr_pfat	Fat percentage trunk region, measured by DEXA scanning	Sex, Age1
WB_lean	Total lean mass in whole body (weight in kgs) measured by DEXA scanning	Sex, Age1, Length at Age1
WB_pfat	Fat percentage in whole body measured by DEXA scanning	Sex, Age1
WB_tf	Total fat in whole body (weight in kgs) measured by DEXA scanning	Sex, Age1, Length at Age 1
PL_Age1[a]	Phospholipids (mmol/l) at age 1	Sex, Age1
TG_Age1[a]	Triglycerides (mmol/l) at age 1	Sex, Age1
TG_Age3[a]	Triglycerides measured at age 3	Sex, Age3
CT_Age1[a]	Total cholesterol (mmol/l) at age 1	Sex, Age1
CT_Age2[a]	Total cholesterol (mmol/l) at age 3	Sex, Age3, $(Age\ 3)^2$
CE_Age1[a]	Esterified cholesterol (mmol/l) at age 1	Sex, Age1
CF_Age1[a]	Free cholesterol (mmol/l) at age 1	Sex, Age1
HDL-C_Age1[a]	High-density-lipoprotein cholesterol (mmol/l) at age 1	Sex, Age1
HDL-C_Age2[a]	High-density-lipoprotein cholesterol (mmol/l) at age 3	Sex, Age3, $(Age\ 3)^2$
LDL-C_Age1[a]	Low-density-lipoprotein cholesterol (mmol/l) at age 1	Sex, Age1
LDL-C_Age2[a]	Low-density-lipoprotein cholesterol (mmol/l) at age 3	Sex, Age3, $(Age\ 3)^2$
hdl_CE_Age1[b]	Esterified cholesterol in high-density-lipoprotein fraction (mmol/l) at age 1	Sex, Age1
hdl_CF_Age1[b]	Free cholesterol in high-density-lipoprotein fraction (mmol/l) at age 1	Sex, Age1
hdl_CT_Age1[b]	Cholesterol in high-density-lipoprotein fraction (mmol/l) at age 1	Sex, Age1
hdl_PL_Age1[b]	Phospholipids in high-density-lipoprotein fraction (mmol/l) at age 1	Sex, Age1
hdl_TG_Age1[b]	Triglycerides in high-density-lipoprotein fraction (mmol/l) at age 1	Sex, Age1
CETP_per[a]	Cholesteryl ester transfer protein activity (CETP activity – percentage) at age 1	Sex, Age1
Fasting_glu[a]	Fasting glucose (mmol/l) at age 3	Sex, Age3, $(Age3)^2$

Age 1, 63 ± 10 days; age 2, 218 ± 45 days; age 3, 242 ± 48 days.
[a]Measured directly in plasma.
[b]Measured in LDL-depleted plasma.

sequenced previously (Clop *et al.* 2016), to maximize the probability of having all the genotypic classes represented in at least three animals in the whole F2 UNIK population ($n = 564$) and, thus, increase the power to detect genetic associations in the experiment.

Genotyping

The Fluidigm D3 ASSAY DESIGN online tool (https://d3.fluidigm.com) was used to design genotyping assays. SNP positions and 100 bp flanking the SNPs were submitted for assay design. SNP assays were run on a total of 564 animals using 96.96 IFC Dynamic Arrays and the SNPtype reagents Kit (Cat# BMK-M10-96.96GT-SNP), according to the manufacturer's recommendations. Samples with known genotypes were included as controls on each array. Of the 30 assays, 24 qualified for final analysis after excluding four assays that failed during the run and two SNPs that showed a minor allele frequency less than 5%.

Association analysis using PLINK

All phenotype data were checked for normality, and log or square root transformations were applied if required. SNP association analyses were performed using a mixed linear model in PLINK (Purcell *et al.* 2007). Cross and sex were fitted as fixed effects in all analyses. Additionally, age, age^2 and body length were included as covariates when appropriate (see Table 1). The PLINK option–adjust was used to calculate false discovery rate and Bonferroni corrected *P*-values.

Breed-specific allele frequencies

Alleles were counted for each SNP marker, and breed-specific allele frequencies were calculated based on allele

Table 2 SNPs selected for genotyping.

Gene name	Chromosome	Position[a]	RS	Ref[b]	Alt[c]	Amino acid change	SIFT class	SIFT score	Effect impact
GRM1	1	19464419	rs342691419	G	A	None			Low
MCHR2	1	66837958	rs325938608	C	T	None			Low
SIM1	1	67278993	rs80966673	G	A	None			Low
HTR1B	1	88598352	rs337239700	A	G	None			Low
ALDH1B1	1	239112489	rs80790173	T	C	None			Low
ALDH3B2	2	4926095		A	C	Asn275His	Deleterious	0.00	Moderate
GPRC5B	3	25882411		A	C	Val282Gly	Deleterious	0.00	Moderate
TAS2R9	5	61254113	rs341774888	G	A	Met62Ile	Tolerated	1.00	Moderate
TAS1R3	6	63616520	rs325926527	G	A	None			Low
TAS1R1	6	67409179	rs330666697	C	G	Leu579Val	Deleterious	0.01	Moderate
LEPR	6	146826063		C	T	None			Low
ALDH6A1	7	97375219		A	C	None			Low
FOS	7	98450413		T	C	Cys50Arg	Deleterious	0.02	Moderate
MTNR1B	9	25556746	rs345547737	A	G	Ile97Val	Tolerated	0.76	Moderate
GPRC5C	12	6701608		C	G	Leu358Val	Tolerated	0.23	Moderate
GPR179	12	23746149		G	T	Ser1779Ile	Deleterious	0.05	Moderate
HTR3C	13	121993580	rs327477069	G	A	None			Low
P2RX2	14	22858286	rs324278003	A	G	None			Low
P2RX7	14	31418726	rs323444034	A	G	None			Low
P2RX7	14	31425255	rs332631772	A	G	None			Low
ALDH2	14	39902743	rs320621677	T	C	Lys515Arg	Tolerated	0.09	Moderate
ADRB1	14	124401286		C	T	None			Low
GABRA6	16	61865942	rs336366772	C	T	None			Low
TAS2R41	18	6781625	rs325525037	G	A	Ala36Val	Tolerated	0.19	Moderate
TAS2R60	18	6808587	rs322889806	C	T	Ala3Thr	Tolerated	0.31	Moderate
TAS2R40	18	7025996	rs319780668	C	T	Gly234Asp	Tolerated	1.00	Moderate
TAS2R39	18	7068985	rs331114472	T	C	Leu37Ser	Tolerated	0.06	Moderate
TAS2R4	18	8135073		A	T	Ile67Phe	Tolerated	0.19	Moderate
GRM8	18	21301813	rs327863473	C	T	None			Low
GABRA3	X	123610295		A	C	[d]			High

[a]All positions refer to the Sscrofa11.1 assembly.
[b]Ref, reference allele.
[c]Alt, alternative allele.
[d]Splice donor-site mutation.

count in purebred Göttingen Minipig, Duroc and Yorkshire founder animals (see Table S1).

Results and discussion

Association study

Out of 66 variants in TASR and AR genes that were identified previously by studying divergent F2 pools of animals from the UNIK population (Clop *et al.* 2016), 24 SNPs located in eight TASR and 15 AR genes were successfully genotyped. Fifteen of these SNPs were significantly associated with at least one of 35 obesity-related phenotypes at a Bonferroni corrected significance level ($P \leq 0.05$) (Table 3 & Table S1). None of the TAS1R genes and only two TAS2R genes showed significant genetic associations: *TAS2R9* with growth (DG2) and *TAS2R4* with fat deposition (retro_fat) and blood lipid levels (CT_Age2, HDL-C_Age2). These results point to *TAS2R4* as a common genetic determinator for both obesity and blood lipid levels. Thirteen variants in AR genes displayed significant genetic associations with at least one obesity-related phenotype. The variant in *ALDH3B2* was associated with both blood lipids (hdl_CF_Age1) and fat deposition (retro_fat) suggesting that this gene also is a common genetic determinator for these traits. The remaining 12 variants in AR genes had an effect on either growth/fat deposition or blood lipid traits.

Variants in *SIM1*, *ALDH1B1*, *P2RX7* and *LEPR* were associated with different measures of growth (DG1, DG2, ADG). For all these loci, the reference alleles, which also were the most frequent alleles in Duroc, had a positive effect on growth. Conversely, the alternative alleles with growth-decreasing effects in general had a high frequency in Göttingen Minipig and Yorkshire (i.e. except for one of the *P2RX7* SNPs, for which both Duroc and Yorkshire were fixed for the reference allele). Variants in *GRM1* and *SIM1* were associated with fat deposition at Age1 measured by DEXA scanning (Tr_pfat, WB_pfat, WB_tf, WB_lean), whereas variants in *P2RX7* were associated with subcutaneous fat deposition at Age3 (BF1). Other genes associated with blood lipid levels were *FOS*, *MTNR1B* and *GPRC5C*, i.e. the SNP in *MTNR1B* and in *FOS* showed association with the level of total and esterified cholesterol and phospholipids measured at young age (CE_Age1, CT_Age1, PL_Age1),

Table 3 Genome wide significant associations

Chromosome	Position[a]	Gene	Phenotype[b]	Beta[c]	*P*-value	*P*-value (Bonf)[d]
1	19464419	*GRM1*	Tr_pfat	−0.874	6.80E-04	0.01633
			WB_pfat	−0.752	3.08E-04	0.00739
			WB_tf	−139.1	7.84E-05	0.00188
1	66837958	*MCHR2*	Fasting_glu	−0.275	1.70E-03	0.04088
1	67278993	*SIM1*	DG1	0.013	0.000636	0.01527
			WB_lean	283.2	0.001232	0.02956
1	239112489	*ALDH1B1*	BMI_Age2	0.674	2.53E-05	0.00061
			DG2	0.038	5.77E-06	0.00014
			ADG	0.032	3.55E-06	8.53E-05
2	4926095	*ALDH3B2*	hdl_CF_Age1	−0.022	0.001	0.03121
			Leaf_fat	0.318	0.0009	0.02069
5	61254113	*TAS2R9*	DG2	0.027	0.0004	0.00934
6	146826063	*LEPR*	ADG	0.021	0.0015	0.03547
7	98450413	*FOS*	hdl_CE_Age1	−0.049	0.0008	0.01958
			HDL-C_Age1	−0.061	0.0017	0.04041
			PL_Age1	−0.109	0.00048	0.01148
9	25556746	*MTNR1B*	CE_Age1	0.101	0.00017	0.00398
			CT_Age1	0.115	0.0005	0.01190
			PL_Age1	0.096	0.0002	0.00469
12	6701608	*GPRC5C*	CT_Age2	0.159	0.0005	0.01194
12	23746149	*GPR179*	Cetp_per	2.001	0.0002	0.00594
14	31418726	*P2RX7*	ADG	0.027	1.80E-03	0.04426
14	31425255	*P2RX7*	Fasting_glu	0.328	7.36E-05	0.00177
			BF1	2.518	4.32E-04	0.01037
14	39902743	*ALDH2*	Cetp_per	1.804	3.00E-04	0.00724
18	8135073	*TAS2R4*	Cetp_per	2.079	1.81E-07	4.33E-06
			CT_Age2	0.161	6.22E-05	0.00149
			HDL-C_Age2	0.111	1.98E-05	0.00048
			Leaf_fat	0.279	0.0002	0.00536
			LDL-C_Age2	0.085	0.001	0.02512

[a]All positions refer to the Sscrofa11.1 assembly.
[b]Phenotype definitions can be found in Table 1.
[c]Beta for the reference allele.
[d]Bonferroni corrected *P*-value.

whereas the SNP in *GPRC5C* displayed association with levels of low density lipoprotein cholesterol and total cholesterol measured at older age (LDL-C_Age2, CT_Age2).

None of the significant genetic associations identified in the present study were recognised in a previous LDLA mapping study (Pant *et al.* 2015). This demonstrates the power of association studies with genetic variants of potential phenotypic importance. A traditional genome-wide association study, such as the LDLA previously performed in the UNIK pedigree (Pant *et al.* 2015), can detect genetic associations only when a causal variant is in linkage disequilibrium (LD) with one or more of the genetic markers under analysis. However, if LD is weak, the QTL will remain undetected. The present work shows that this was the case in our previous LDLA. It should be noted that the same animals and phenotypes were used in both experiments. Even though different statistical models were applied, it is clear that none of the genetic associations found in the present study were reported using the 60K SNPchip (Table S1). As illustrated in Figs S1 & S2, this was not due to a lack of genetic markers in the previous association study. Instead, it was due to lack of robust LD between the coding variants tested now and the markers of the SNP chip.

It might be argued that the SNPs tested in the present study would not have been deemed genetically associated with a trait at a genome-wide significance level had they been included in a GWAS with a set of 60K markers. Although this is true, this mainly illustrates the limited power of the animal material ($n = 564$) and emphasizes that, especially for material with low power, it is highly beneficial to make an informed selection of markers to be included in the association study. The limitation of our approach is, of course, that only QTL in selected genes can be detected, leaving most of the genome unexplored.

Fifteen out of the 24 successfully genotyped variants were genetically associated with obesity-related physical or metabolic phenotypes. Additionally, one locus (*HTR1B*) was found to be associated with borderline significance with high density lipoprotein cholesterol at Age3 (data not shown). Hence, 15 (16 if including the borderline significant locus) of 20 associations from Clop *et al.*

(2016) were confirmed to be associated with growth, obesity or obesity-related traits. Additionally, we identified an association between a variant in *TAS2R9* with daily gain (DG2) that was not detected in the subset of pigs used by Clop *et al.* (2016). In conclusion, according to the present and more powerful method, the analysis performed by Clop *et al.* (2016), based on a subset of extreme animals selected from the UNIK population, reports three false positives (out of 20) and only one false negative. This clearly demonstrates the strength of an approach selecting individuals with extreme phenotypes. However, the present study also clearly demonstrates how the inclusion of more individuals and phenotypes has the power to dissect, define and reveal endophenotypes underlying complex multifactorial phenotypes like obesity and growth. For example, the variant studied in *TAS2R4* is not associated with only average daily gain and retroperitoneal fat content, as described by Clop *et al.* (2016). More precisely, this locus is associated with deposition of retroperitoneal fat and subcutaneous fat but not with weight of omental fat or intestinal fat (data not shown). Additionally, the same locus is associated with blood levels of HDL-cholesterol, triglycerides and CETP, though only in older animals, not in young pigs.

The biology behind the associations

Taste receptor and AR genes influence eating behaviour, affecting food intake and subsequently obesity status. In the present study, 13 SNPs in 12 AR genes and two SNPs in TASR genes showed statistically significant association with one or more obesity-relevant phenotypes. Of these, four AR genes (*SIM1*, *FOS*, *MCHR2* and *LEPR*) and two TASR genes (*TAS2R4* and *TAS2R9*) corroborate previous knowledge on biological function. Out of these six genes, the SNPs in *FOS*, *TAS2R4* and *TAS2R9* give rise to amino acid substitutions, which suggests a direct impact on the phenotype. Variants tested in the remaining genes (*SIM1*, *MCHR2* and *LEPR*) are synonymous, which suggests a more subtle or regulatory effect or linkage disequilibrium with a nearby causative mutation.

SIM1, single-minded homolog 1, is a transcription factor involved in regulation of energy homeostasis through the melanocortin 4 receptor (MC4R) signalling pathway (Xi *et al.* 2013). $Sim1^{(-/-)}$ mice develop hyperphagic early-onset obesity (Michaud *et al.* 2001). In humans, haploinsufficiency of *SIM1* has been found to cause severe early-onset obesity (Holder *et al.* 2000). The association found between the SNP in *SIM1* and daily gain (DG1) and total lean mass (WB_lean), both at young age, indicates that common variants in this gene may have more subtle effects relevant in common obesity.

FOS is a member of the AP-1 family of inducible transcription factors involved in cell proliferation and differentiation, and it also activates the synthesis of phospholipids (Pecchio *et al.* 2011). In the present study, the variation in *FOS* was associated with phospholipid levels (PL_Age1) as well as with high density cholesterol level (HDL-C_Age1) and esterified cholesterol level in LDL-depleted plasma (hdl_CE_Age1), all in the young pig. The Göttingen Minipig variant causes an amino acid change from cysteine (nonpolar) to arginine (basic), which may have a significant impact on protein function.

The *melanin-concentrating hormone receptor 2* (*MCHR2*) gene encodes for a G-protein coupled receptor expressed in the arcuate nucleus and the ventral medial nucleus regions of the brain. Both areas are involved in regulation of food intake (Sailer *et al.* 2001). The gene is present in humans, some other primates and carnivores as well as in pigs but is not found in rodents (Hill *et al.* 2001; Sailer *et al.* 2001). Even though *MCHR2* is not expressed in rodents, a recent study showed that expression of human *MCHR2* in transgenic mice protected against diet-induced obesity (Chee *et al.* 2014). There are several studies in humans linking several polymorphisms in this gene to obesity (Gibson *et al.* 2004; Ghoussaini *et al.* 2007). In the present study, the SNP in this gene was found to be significantly associated with fasting glucose levels. The fact that this gene is found in pig and not in rodents underpins that the pig provides a better model than do rodents for metabolic disturbances associated with glucose metabolism.

Leptin signalling is the main mechanism controling food intake and energy homeostasis in mammals in order to modulate body mass and body composition. Activation of leptin receptors induces modifications in the expression of several hypothalamic neuropeptide genes (i.e. *POMC*, *NPY*, *AGRP*) resulting in hypophagia and increased energy expenditure (Balthasar 2006; Robertson *et al.* 2008). In mice and humans, polymorphisms in both leptin and leptin receptor genes have been widely studied in relation to obesity. In pigs, several leptin variants have been associated with growth, fatness, feed intake and body composition (Munoz *et al.* 2009; Óvilo *et al.* 2010), linking obesity with this appetite-regulating mechanism. These results are in agreement with the association found in the present study that links the SNP in *LEPR* to average daily gain.

TAS2Rs, including TAS2R4 and TAS2R9, mediate nutrient assimilation and endocrine responses (Bachmanov & Beauchamp 2007). Chemosensory signalling pathways involving these TAS2Rs have been implicated in gut hormone release affecting obesity and adipocyte metabolism in mice (Avau *et al.* 2015). Additionally, TAS2R genes are differentially expressed in the gastrointestinal tract in response to a high fat diet in mice (Vegezzi *et al.* 2014). The *TAS2R4* SNP included in this study results in an amino acid substitution (isoleucine to phenylalanine) in a transmembrane domain. This variant is associated with CETP activity (CETP_per) and cholesterol levels in blood (HDL-C_Age2, CT_Age2). This is in agreement with a potential relationship with cholesterol metabolism suggested by Pydi

et al. (2016), who reported that TAS2R4 displays cholesterol sensitivity in its signalling function. The association of the SNP in *TAS2R9* and average daily gain (DG2) is also interesting as it may link taste perception with glucose metabolism and henceforth the feeling of satiety. Our results are in agreement with the study by Dotson *et al.* (2008), who showed that certain variants of TAS2R genes, including variants in *TAS2R9*, are associated with type 2 diabetes and influence glucose homeostasis. The SNP genotyped in *TAS2R9* results in a missense mutation changing a methionine to isoleucine. A functionally compromised TAS2R9 receptor could impact glucose homeostasis, resulting in increase in daily gain, providing an important link between alimentary chemosensation and body growth disturbances that can result in obesity.

Five out of the remaing nine SNPs in AR genes (*MTNR1,B ALDH2, GPR179, GPRC5C* and *ALDH3B2*) that showed statistically significant association with obesity-relevant phenotypes give rise to amino acid substitutions that suggest a direct impact on the phenotype. Variants tested in the remaining genes (*GRM1, ALDH1B1* and *P2RX7*) are synonymous which, as mentioned above, might suggest a regulatory effect or LD with a nearby causative mutation. To our knowledge, the biological function in the context of obesity-related traits for these eight genes has not been clearly demonstrated previously. Three of the genes belong to the aldehyde dehydrogenase family of proteins (*ALDH1B1, ALDH2, ALDH3B2*), whereas the remaining are receptors expressed mainly in the central nervous system. As shown in Table 3, some of the genes seem to be involved in regulation of metabolism, whereas others are involved in fat deposition and daily gain, and in most cases the functional importance is clearly age specific.

Conclusions and perspectives

We have analysed associations of a total of 15 variants located in two TASR and 12 AR genes and verified a genetic association with specific obesity-related phenotypes that we discovered in our previous study. For six variants, the effects on the associated phenotype(s) immediately agree with the known function of the genes. These six variants are located in four AR genes (*SIM1, FOS, MCHR2* and *LEPR*) and two TASR genes (*TAS2R4* and *TAS2R9*). With regards to the remaining eight variants in AR genes (*MTNR1B, ALDH2, GPR179, GPRC5C, ALDH3B2, GRM1, ALDH1B1* and *P2RX7*) this study provides novel insights of potential functional importance due to the unique associations with specific obesity-related phenotypes. Interestingly, most of the 15 statistically significant associations are age specific, indicating that the regulation of fat deposition, weight gain and metabolism differs between adolescents and older individuals. Additional studies need to be performed to provide more specific information on the molecular pathways affected by the gene variants investigated here.

We have employed a targeted approach using potentially functional SNPs, previously identified by next generation sequencing, in the present genetic association study. Using these variants we identified a significant number of genetic associations that were missed in a traditional association analysis based on SNP arrays (Pant *et al.* 2015). Thus, this strategy proved to be a powerful and cost-effective approach for increasing the power of genetic association studies for complex traits.

From an animal breeding perspective, our results prompt a discussion on how results from this and similar studies may be used. The biology underlying selected traits has until now remained an unopened black box in animal breeding. Here, we identify functional variants in genes related to eating behaviour, showing a direct association with growth and leanness. Hence, this study discloses genetic variants that have been unknowingly targeted by selection in traditional breeding programs. We suggest that this knowledge may be exploited in precision animal breeding in the future.

Acknowledgements

This study was supported by the Danish Independent Research Council Grant Number DFF 1335-00127. Alex Clop acknowledges the Rasmon y Cajal Fellowship program from the Spanish Ministry of Economy and Competitiveness (RYC-2011-07763).

Conflict of interest

The authors declare not to have any conflict of interest.

References

Andersson L.B. (1996) Genes and obesity. *Annals of Medicine* **28**, 5–7.

Avau B., Bauters D., Steensels S., Vancleef L., Laermans J., Lesuisse J., Buyse J., Lijnen H.R., Tack J. & Depoortere I. (2015) The gustatory signaling pathway and bitter taste receptors affect the development of obesity and adipocyte metabolism in mice. *PLoS One* **10**, e0145538.

Bachmanov A.A. & Beauchamp G.K. (2007) Taste receptor genes. *Annual Review of Nutrition* **27**, 389–414.

Balthasar N. (2006) Genetic dissection of neuronal pathways controlling energy homeostasis. *Obesity* **14**, 222S–7S.

Berg C.J. & Kaunitz J.D. (2016) Gut chemosensing: implications for disease pathogenesis. *F1000Research* **5**, 2424.

Breslin P.A. (2013) An evolutionary perspective on food and human taste. *Current Biology* **23**, R409–18.

Chao D.H.M., Argmann C., Van Eijk M. *et al.* (2016) Impact of obesity on taste receptor expression in extra-oral tissues: emphasis on hypothalamus and brainstem. *Scientific Reports* **6**, 29094.

Chee M.J.S., Pissios P., Prasad D. & Maratos-Flier E. (2014) Expression of melanin-concentrating hormone receptor 2 protects against diet-induced obesity in male mice. *Endocrinology* **155**, 81–8.

Cingolani P., Platts A., Wang L.L., Coon M., Nguyen T., Wang L., Land S.J., Lu X. & Ruden D.M. (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SNPEFF: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly*, **6**, 80–92.

Clop A., Sharaf A., Castelló A. *et al.* (2016) Identification of protein-damaging mutations in 10 swine taste receptors and 191 appetite–reward genes. *BMC Genomics* **17**, 1–17.

Dotson C.D., Zhang L., Xu H., Shin Y.-K., Vigues S., Ott S.H., Elson A.E., Choi H.J., Shaw H. & Egan J.M. (2008) Bitter taste receptors influence glucose homeostasis. *PLoS One* **3**, e3974.

Fulton S. (2010) Appetite and reward. *Frontiers in Neuroendocrinology* **31**, 85–103.

Ghoussaini M., Vatin V., Lecoeur C. *et al.* (2007) Genetic study of the *melanin-concentrating hormone receptor 2* in childhood and adulthood severe obesity. *Journal of Clinical Endocrinology and Metabolism* **92**, 4403–9.

Gibson W.T., Pissios P., Trombly D.J., Luan J., Keogh J., Wareham N.J., Maratos-Flier E., O'Rahilly S. & Farooqi I.S. (2004) Melanin-concentrating hormone receptor mutations and human obesity: functional analysis. *Obesity Research* **12**, 743–9.

Guerin M., Dolphin P.J. & Chapman M.J. (1994) A new *in vitro* method for the simultaneous evaluation of cholesteryl ester exchange and mass transfer between HDL and apoB-containing lipoprotein subspecies. Identification of preferential cholesteryl ester acceptors in human plasma. *Arteriosclerosis, Thrombosis, and Vascular Biology* **14**, 199–206.

Hill J., Duckworth M., Murdock P., Rennie G., Sabido-David C., Ames R.S., Szekeres P., Wilson S., Bergsma D.J. & Gloger I.S. (2001) Molecular cloning and functional characterization of MCH2, a novel human MCH receptor. *Journal of Biological Chemistry* **276**, 20125–9.

Holder J.L. Jr, Butte N.F. & Zinn A.R. (2000) Profound obesity associated with a balanced translocation that disrupts the *SIM1* gene. *Human Molecular Genetics* **9**, 101–8.

Houpt K.A., Houpt T.R. & Pond W.G. (1979) The pig as a model for the study of obesity and of control of food intake: a review. *Yale Journal of Biology and Medicine* **52**, 307–29.

Kiuchi S., Yamada T., Kiyokawa N., Saito T., Fujimoto J. & Yasue H. (2006) Genomic structure of swine *taste receptor family 1 member 3*, *TAS1R3*, and its expression in tissues. *Cytogenetic and Genome Research* **115**, 51–61.

Kogelman L.J., Kadarmideen H.N., Mark T., Karlskov-Mortensen P., Bruun C.S., Cirera S., Jacobsen M.J., Jorgensen C.B. & Fredholm M. (2013) An f2 pig resource population as a model for genetic studies of obesity and obesity-related diseases in humans: design and genetic parameters. *Frontiers in Genetics* **4**, 29.

Koopmans S.J. & Schuurman T. (2015) Considerations on pig models for appetite, metabolic syndrome and obese type 2 diabetes: from food intake to metabolic disease. *European Journal of Pharmacology* **759**, 231–9.

Loper H.B., La Sala M., Dotson C. & Steinle N. (2015) Taste perception, associated hormonal modulation, and nutrient intake. *Nutrition Reviews* **73**, 83–91.

Michaud J.L., Boucher F., Melnyk A., Gauthier F., Goshu E., Levy E., Mitchell G.A., Himms-Hagen J. & Fan C.M. (2001) Sim1 haploinsufficiency causes hyperphagia, obesity and reduction of the paraventricular nucleus of the hypothalamus. *Human Molecular Genetics* **10**, 1465–73.

Munoz G., Ovilo C., Silio L., Tomas A., Noguera J.L. & Rodriguez M.C. (2009) Single- and joint-population analyses of two experimental pig crosses to confirm quantitative trait loci on *Sus scrofa* chromosome 6 and leptin receptor effects on fatness and growth traits. *Journal of Animal Science* **87**, 459–68.

Óvilo C., Fernández A., Fernández A.I., Folch J.M., Varona L., Benítez R., Nuñez Y., Rodríguez C. & Silió L. (2010) Hypothalamic expression of porcine *leptin receptor* (*LEPR*), *neuropeptide Y* (*NPY*), and *cocaine- and amphetamine-regulated transcript* (*CART*) genes is influenced by *LEPR* genotype. *Mammalian Genome* **21**, 583–91.

Pant S.D., Karlskov-Mortensen P., Jacobsen M.J. *et al.* (2015) Comparative analyses of QTLs influencing obesity and metabolic phenotypes in pigs and humans. *PLoS One* **10**, e0137356.

Pecchio A.R.A., Gizzi A.M.C., Renner M.L., Molina-Calavita M. & Caputto B.L. (2011) c-Fos activates and physically interacts with specific enzymes of the pathway of synthesis of polyphosphoinositides. *Molecular Biology of the Cell* **22**, 4716–25.

Pomp D. (1997) Genetic dissection of obesity in polygenic animal models. *Behavior Genetics* **27**, 285–306.

Purcell S., Neale B., Todd-Brown K. *et al.* (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics* **81**, 559–75.

Pydi S.P., Jafurulla M., Wai L., Bhullar R.P., Chelikani P. & Chattopadhyay A. (2016) Cholesterol modulates bitter taste receptor function. *Biochimica et Biophysica Acta (BBA)-Biomembranes* **1858**, 2081–7.

Ribani A., Bertolini F., Schiavo G., Scotti E., Utzeri V.J., Dall'Olio S., Trevisi P., Bosi P. & Fontanesi L. (2017) Next generation semiconductor based sequencing of bitter taste receptor genes in different pig populations and association analysis using a selective DNA pool-seq approach. *Animal Genetics* **48**, 97–102.

Robertson S.A., Leinninger G.M. & Myers M.G. Jr (2008) Molecular and neural mediators of leptin action. *Physiology and Behavior* **94**, 637–42.

Sailer A.W., Sano H., Zeng Z. *et al.* (2001) Identification and characterization of a second melanin-concentrating hormone receptor, MCH-2R. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 7564–9.

Schembre S.M., Cheng I., Wilkens L.R., Albright C.L. & Marchand L.L. (2013) Variations in bitter-taste receptor genes, dietary intake, and colorectal adenoma risk. *Nutrition and Cancer* **65**, 982–90.

da Silva E.C., de Jager N., Burgos-Paz W., Reverter A., Perez-Enciso M. & Roura E. (2014) Characterization of the porcine nutrient and taste receptor gene repertoire in domestic and wild populations across the globe. *BMC Genomics* **15**, 1057.

Spurlock M.E. & Gabler N.K. (2008) The development of porcine models of obesity and the metabolic syndrome. *Journal of Nutrition* **138**, 397–402.

Vegezzi G., Anselmi L., Huynh J., Barocelli E., Rozengurt E., Raybould H. & Sternini C. (2014) Diet-induced regulation of bitter taste receptor subtypes in the mouse gastrointestinal tract. *PLoS One* **9**, e107732.

Xi D., Roizen J., Lai M., Gandhi N. & Kublaoui B. (2013) Paraventricular nucleus Sim1 neuron ablation mediated obesity is resistant to high fat diet. *PLoS One* **8**, e81087.

Supporting information

Additional supporting information may be found online in the supporting information tab for this article:

Figure S1 Comparison of results from a previous LDLA study and the present association analysis. Open symbols represent results for SNP markers used in the LDLA study. Black symbols represent results for the variant in *ALDH1B1* tested in the present study. Position on chromosome 1 (Sscrofa11.1 assembly) is on the *X*-axis; the negative logarithm (base 10) of the raw *P*-value is on the *Y*-axis.

Figure S2 Comparison of results from a previous LDLA study and the present association analysis. Open symbols represent results for SNP markers used in the LDLA study. The black symbol represents the result for the variant in *P2RX7*, position 33289293 tested in the present study. Position on chromosome 14 (Sscrofa11.1 assembly) is on the *X*-axis; the negative logarithm (base 10) of the raw *P*-value is on the *Y*-axis.

Table S1 Results overview.