

RESEARCH ARTICLE

Open Access

# Genome-wide association and systems genetic analyses of residual feed intake, daily feed consumption, backfat and weight gain in pigs

Duy Ngoc Do<sup>1</sup>, Tage Ostensen<sup>2</sup>, Anders Bjerring Strathe<sup>1,2</sup>, Thomas Mark<sup>1</sup>, Just Jensen<sup>3</sup> and Haja N Kadarmideen<sup>1\*</sup>

## Abstract

**Background:** Feed efficiency is one of the major components determining costs of animal production. Residual feed intake (RFI) is defined as the difference between the observed and the expected feed intake given a certain production. Residual feed intake 1 (RFI1) was calculated based on regression of individual daily feed intake (DFI) on initial test weight and average daily gain. Residual feed intake 2 (RFI2) was as RFI1 except it was also regressed with respect to backfat (BF). It has been shown to be a sensitive and accurate measure for feed efficiency in livestock but knowledge of the genomic regions and mechanisms affecting RFI in pigs is lacking. The study aimed to identify genetic markers and candidate genes for RFI and its component traits as well as pathways associated with RFI in Danish Duroc boars by genome-wide associations and systems genetic analyses.

**Results:** Phenotypic and genotypic records (using the Illumina Porcine SNP60 BeadChip) were available on 1,272 boars. Fifteen and 12 loci were significantly associated ( $p < 1.52 \times 10^{-6}$ ) with RFI1 and RFI2, respectively. Among them, 10 SNPs were significantly associated with both RFI1 and RFI2 implying the existence of common mechanisms controlling the two RFI measures. Significant QTL regions for component traits of RFI (DFI and BF) were detected on pig chromosome (SSC) 1 (for DFI) and 2 for (BF). The SNPs within *MAP3K5* and *PEX7* on SSC 1, *ENSSSCG00000022338* on SSC 9, and *DSCAM* on SSC 13 might be interesting markers for both RFI measures. Functional annotation of genes in 0.5 Mb size flanking significant SNPs indicated regulation of protein and lipid metabolic process, gap junction, inositol phosphate metabolism and insulin signaling pathway are significant biological processes and pathways for RFI, respectively.

**Conclusions:** The study detected novel genetic variants and QTLs on SSC 1, 8, 9, 13 and 18 for RFI and indicated significant biological processes and metabolic pathways involved in RFI. The study also detected novel QTLs for component traits of RFI. These results improve our knowledge of the genetic architecture and potential biological pathways underlying RFI; which would be useful for further investigations of key candidate genes for RFI and for development of biomarkers.

**Keywords:** GWAS, Residual feed intake, Backfat, Pigs, Systems genetics, Pathways

## Background

Residual feed intake (RFI), defined as the difference between the observed feed intake and the predicted feed intake based on average daily gain and backfat, is a sensitive and accurate indicator for feed efficiency in livestock [1]. Selection for reduced RFI can improve the

efficiency of energy utilization without reducing the feed intake capacity that is required for production [2]. Recent studies showed lower RFI selection resulted in better feed conversion efficiency and meat quality in pigs [3] and lower environmental impact in pigs [4], sheep [5] and cattle [6]. Therefore, selection for reduced RFI is important for both economic and environmental aspects of production. Residual feed intake has moderate heritabilities (0.34-0.38) in Danish Duroc pigs and responds to selection [7].

\* Correspondence: hajak@sund.ku.dk

<sup>1</sup>Section of Animal Genetics, Bioinformatics and Breeding, Department of Veterinary Clinical and Animal Sciences, Faculty of Health and Medical Sciences, University of Copenhagen, 1870 Frederiksberg C, Denmark  
Full list of author information is available at the end of the article

The understanding of genetic mechanisms underlying traits could potentially be important for setting up priors for (genetic) variances in genomic selection or help finding candidate genes for marker- or gene-assisted selection [8,9]. Many approaches including linkage analyses, genome-wide association studies (GWAS), candidate gene association and transcriptomic profiling for RFI have been performed to unravel the genetic background behind the complex trait in many species. For instance, GWAS and linkage studies were performed by [10-12], candidate gene approaches were carried out in [13-19] or transcriptomic studies were used in cattle [20]. These studies revealed many candidate genes and offer background information for genetic studies of RFI in cattle. Compared to cattle, fewer genomic studies have been conducted for RFI in pigs. Gilbert *et al.* [21] detected a QTL on pig chromosome (SSC) 5 and 9 for RFI in growing pigs in a Piétrain–Large White backcross. Fat *et al.* [22] indicated SNPs in the *FTO* and *TCF7L2* gene were significantly associated with RFI in a candidate gene study. Using transcriptomic approaches, Lkhagvadorj *et al.* [23] found many genes in fat (311) and in liver (147) that were differently expressed in low and high RFI pigs in response to caloric restriction and indicated that lipid metabolic pathways was important for regulation of RFI. A recent GWAS has revealed several major QTLs on SSC 7 and 14 influencing RFI in Yorkshire pigs [24]. Jiao *et al.* [25] detected a significant region for FCR on SSC 4 but did not find any significant QTL for RFI in 1,022 Duroc boars. Sahana *et al.* [26] performed GWAS for FCR and found a significant QTL for the trait on SSC 14 in Duroc pigs. Feed conversion ratio is closely related to RFI and is currently included in the selection index for the Danish pig breeds. However, ratio traits such as FCR are not ideal for statistical and biological reasons [27] and there is still a debate about what exactly is the best definition for feed efficiency in production animals. RFI is not ideal under all circumstances either [2], but are well established and increasingly used as a measure for animal breeding and selection in all livestock species. Also, for the purpose of understanding the biology behind feed efficiency it was preferred to have a measure which is independent of daily gain. Therefore, the GWAS and systems genetics study was performed on RFI and its component traits to identify genetic variants and potential candidate genes for these traits as well as possible biological mechanism controlling feed efficiency in Duroc boars.

## Methods

### Estimation of residual feed intake and deregressed estimated breeding values

All phenotypic data used in this study of 7,358 Duroc pigs were recorded at a central test station (Bøgildgård)

for a period of 4 years (2008 to 2011) and were supplied by the Pig Research Centre, Danish Agriculture and Food Council [7]. Pigs were fed ad libitum from 30 kg to approximately 100 kg BW with the same feed composition. Feed intake was recorded by ACEMA64 (ACEMO, Pontivy, France) automatic dry feeding stations. All data records in test station including the amount of food in each visits, number of visit to feeder per day and time spent in each visit were sent to Pig Research Center for further processes. The boars were weighted by the technical staff on arrival and at regular intervals (~twice a month) during the growth period. The methods of calculation of RFI have been discussed in [7]. In summary, residual feed intake was computed as the difference between the observed average daily feed intake and the predicted daily feed intake. Two models were used to define RFI. In the first model for RFI1, predicted daily feed intake was estimated using linear regression of DFI on initial test weight (BWd) and average daily gain from 30 kg to 100 kg (ADG). In the second model for RFI2, BWd, ADG and backfat were used as regressors. The two measures of RFI are closely related with an overall phenotypic correlation of 0.96 in Duroc pigs [7]. The estimated breeding values (EBVs) for RFI and these component traits were calculated by single-trait animal model with fixed effect of herd-week-section, and random effects of pen, additive genetics and residuals [7]. These EBVs were calculated using Best Linear Unbiased Predictions [28] based on pedigree traced back to 1970, including 34 generations. Deregressed estimated breeding values (dEBVs) were used as dependent variable in GWAS. The deregression procedure of Garrick *et al.* [29] was used. It adjusts for ancestral information, such that the dEBV only contains their own and the descendant's information on each animal. Deregressed EBVs have unequal variances and therefore, should be used in a weighted analysis. The weight for the  $i^{\text{th}}$  animal was estimated as:

$$w_i = (1-h^2) / \left[ \left( c + \frac{(1-r_i^2)}{r_i^2} \right) h^2 \right],$$

in which  $c$  was the part of the genetic variance that was assumed to be not explained by markers ( $c = 0.1$ ),  $h^2$  was the heritability of the trait, and  $r_i^2$  was the reliability of the dEBV of the  $i^{\text{th}}$  animal. Summary of raw phenotypes, dEBV and weight factors of dEBV for RFI and its component traits of genotyped animals is shown in Table 1.

### Genotyping and SNP data validation

The details of the genotyping method have been described previously [30,31]. In summary, genomic DNA was isolated from all specimens by treatment with proteinase K followed by sodium chloride precipitation and

**Table 1 Statistical description for residual feed intake and its component traits of genotyped animals used in the study<sup>1</sup>**

Traits <sup>2</sup>	Phenotype	dEBV	Reliability of dEBV	Weight factors for dEBV
DFI(kg)	2.29 ± 0.31	-0.20 ± 0.41	0.18 ± 0.08	0.64 ± 0.54
ADG(kg/day)	1.10 ± 0.10	0.10 ± 0.13	0.19 ± 0.08	0.60 ± 0.47
BF(mm)	7.69 ± 1.04	0.53 ± 0.06	0.43 ± 0.08	1.01 ± 0.58
RFI1(kg)	-0.08 ± 0.30	-0.20 ± 0.25	0.24 ± 0.09	1.07 ± 0.96
RFI(g)	-0.08 ± 0.29	-0.20 ± 0.24	0.25 ± 0.09	1.13 ± 0.99

<sup>1</sup>: The mean residual feed intake was calculated from only genotyped animals. The mean of RFI in the whole population was zero. The negative mean values from current data set implied that these animals have better feed efficiency. The values are shown as mean ± SD.

<sup>2</sup>: DFI: Daily feed intake, ADG: Average daily gain, BF: Backfat, RFI1: Residual feed intake 1, RFI2: Residual feed intake 2, dEBV: Deregressed EBV.

SNPs were genotyped on the PorcineSNP60 Illumina iSelect BeadChip. The criteria for screening the genomic data was a call rate per animal of 0.95, call rate per SNP marker of 0.95, Hardy Weinberg equation test with  $p < 0.0001$ , and minor allele frequency  $> 0.05$ .

### Statistical analyses

#### *Linear mixed model used for genome wide association studies*

To control the false positives due to family structure, the following linear mixed model was used to analyze the association between the SNP (modeled individually; one at a time) and the phenotypes:

$$y = \mathbf{1}\mu + \mathbf{Z}a + \mathbf{m}g + e$$

where  $y$  is the vector of dEBVs for RFI (also for other phenotypes including ADG, DFI and BF),  $\mathbf{1}$  is a vector of 1s with length equal to number of observations,  $\mu$  is the general mean,  $Z$  is an incidence matrix relating phenotypes to the corresponding random polygenic effect,  $a$  is a vector of the random polygenic effect  $\sim N(0, A\sigma_u^2)$ , where  $A$  is the additive relationship matrix and  $\sigma_u^2$  is the polygenic variance,  $m$  is a vector with genotypic indicators (-1, 0, or 1) associating records to the marker effect,  $g$  is a scalar of the associated additive effect of the SNP, and  $e$  is a vector of random environmental deviates  $\sim N(0, W^{-1}e^2)$ , where  $\sigma_e^2$  is the general error variance and  $W$  is the diagonal matrix containing weights of the dEBVs. The model was analysed by restricted maximum likelihood (REML) using the DMU software [32] and testing was done using a Wald test against a null hypothesis of  $g = 0$ . The genome-wide significant association following Bonferroni multiple testing correction at 5% significant level was a  $p$  value of  $1.52 \times 10^{-6}$ . The Bonferroni correction is highly conservative and may result in too stringent a threshold and hence many false negative results [33]. Therefore, we also considered a more liberal significant threshold where a SNP was considered to have moderate or suggestive significant association with  $p < 5 \times 10^{-5}$  [34]. Both significant

and suggestive SNPs were used in bioinformatics analysis.

#### Detection of linkage disequilibrium block and haplotypes

Linkage disequilibrium (LD) block analyses were performed for the chromosomal regions with multiple significant SNPs clustered. The blocks were defined using Haploview [35] with the criteria suggested by Gabriel *et al.* [36] to further characterize candidate regions affecting RFI. The criteria by Gabriel *et al.* [36] defined a haplotype block as a region over which 95% of informative SNP pairs show strong LD (strong LD is defined if the one-sided upper 95% confidence bound on  $D'$  is  $> 0.98$  and the lower bound is above 0.7)

#### Systems genetics analyses

SNP positions were updated according to the newest release from Ensembl (Sscrofa10.2 genome version). Identification of the closest genes to significant and suggestive SNPs was obtained using Ensembl annotation of Sscrofa10.2 genome version ([http://ensembl.org/Sus\\_scrofa/Info/Index](http://ensembl.org/Sus_scrofa/Info/Index)). The positional candidate genes within 1 Mb bin size of significant and suggestive SNPs were scanned using function *GetNeighGenes()* in the NCBI2R R-package at <http://cran.r-project.org/web/packages/NCBI2R/index.html> [37]. Investigation of functional categories and the relevant KEGG pathways for the genes within 1 Mb bin size of significant SNPs was performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID) available at <http://david.abcc.ncifcrf.gov/> [38]. The selection of 1 Mb bin size or 0.5 Mb flanking regions of significant SNPs was based on previous results of Sahana *et al.* [26] who observed very high LD in this Duroc pig population (average  $r^2 = 0.3$  between two markers in 1 Mb distance). This result suggests a similar distance (1 Mb bin size) can be used to capture the causal genes/SNPs. The positional candidate genes identified by NCBI2R package were first assigned to the KEGG pathways (<http://www.genome.jp/kegg/pathway.html>) and GO terms (<http://www.geneontology.org/>). Then, these related pathways/GO terms were tested if they appear by random chance by using modified Fisher exact test. The pathway/GO terms were

declared to be important for the traits if they do not appear by a random chance with  $p < 0.05$  [38].

## Results

### Genome wide association results for residual feed intake

Following quality control, 23,795 markers were first excluded as having a low (<5%) minor allele frequency, secondly 1,836 markers were excluded because of low (<95%) call rate and finally 3,463 markers were excluded because they were not in HWE ( $p < 0.0001$ ), two animals were removed because of low (<95%) call rate. A final set of 33,241 SNPs and 1,272 pigs was retained for GWAS. Fifteen and twelve SNPs were significantly associated to RFI1 and RFI2 at  $p < 1.52 \times 10^{-6}$  (Bonferroni correction), respectively in which nine SNPs associated with both traits (Tables 2 and 3). The highest contribution of a significant SNP to additive genetic variance was only 0.21% in each trait. Moreover, 138 and 176 SNPs were found to have suggestive (moderately significant) association with RFI1 and RFI2 at  $p < 5 \times 10^{-5}$ , of which 124 SNPs have been found to be associated with both traits. High numbers of significant SNPs for both RFI were found on SSC 1, 8, 9, 11, 12, 13, 14, 15, 16 and 18, while none of them were found on SSC 2, 10 and 17. Several regions: 27-33 Mb on SSC 1, 89-91 Mb on SSC 8, 119-122 Mb on SSC 9 and 26-35 Mb on SSC 18 harbored many associated SNPs for both RFI. Genome wide Manhattan plots displaying the GWA results with the respect to their position are shown in Figure 1a and b. Lists of genes located within 0.5 Mb window from the significant and suggestive SNPs is provided in Additional file 1.

### Genome wide association results for component traits of residual feed intake

The genome wide association analysis showed only one significantly associated SNP ( $p = 6.10 \times 10^{-7}$ ) for DFI at 64 Mb position on SSC 1 (Table 4). Moreover, 25 other suggestive associated SNP were also found on SSC 1, 3, 7, 9, 14, and 16 and two suggestive SNPs have not yet been mapped on any chromosome (Additional file 1). None of significant SNP was found for ADG, however, 15 suggestive SNPs were identified on SSC 6, 15 and 17 (Additional file 1). Thirteen of them were located in 53–54 Mb on SSC17 and marker ASGA0077977 was tightest association with trait at  $p = 1.67 \times 10^{-6}$  (Table 4). All of six significant SNPs associated with BF were located on a region of 82-86 Mb on SSC 2 (Table 4). Moreover, 73 suggestive SNPs for BF were also located in the same region (Additional file 1). Fifteen suggestive SNP for BF were located on region of 60-68 Mb on SSC 1 and 7 SNPs was not mapped onto any chromosome. Genome wide Manhattan plots displaying the GWA results for DFI, ADG and BF with the respect to their positions are shown in Figure 2a, b and c, respectively.

### LD block, haplotype analysis and functional categories of positional candidate genes for residual feed intake

Four and three LD blocks were identified in regions 30.5-31.5 Mb on SSC 1 and 120.5-121.5 Mb on SSC 9, respectively. The Manhattan plot, LD blocks and Ensembl genome for candidate regions on SSC 1 and SSC 9 were shown in Figures 3 and 4, respectively. Frequency of each haplotype for different LD blocks on SSC 1 and SSC 9 was shown in Additional files 2 and 3,

**Table 2 The significant SNPs associated with residual feed intake 1 (RFI1)**

SNP	SSC <sup>1</sup>	Position	p-value	MAF <sup>2</sup>	Var.Exp <sup>3</sup>	Nearest gene	Gene name
MARC0013869	0	0	$5.13 \times 10^{-7}$	0.47	0.21		
MARC0112693	1	30734120	$1.52 \times 10^{-6}$	0.44	0.19	<i>PEX7</i>	PEX7 peroxisomal biogenesis factor 7
H3GA0001223	1	30,769,583	$1.41 \times 10^{-6}$	0.44	0.19	<i>PEX7</i>	PEX7 peroxisomal biogenesis factor 7
ASGA0082396	1	30,941,797	$3.73 \times 10^{-7}$	0.45	0.20	<i>MAP3K5</i>	mitogen-activated protein kinase kinase kinase 5
ALGA0106992	1	30,962,276	$3.13 \times 10^{-7}$	0.45	0.20	<i>MAP3K5</i>	mitogen-activated protein kinase kinase kinase 5
ASGA0094502	1	30,978,281	$3.73 \times 10^{-7}$	0.45	0.20	<i>MAP3K5</i>	mitogen-activated protein kinase kinase kinase 5
ALGA0107451	1	31,008,523	$3.07 \times 10^{-7}$	0.45	0.20	<i>MAP3K5</i>	mitogen-activated protein kinase kinase kinase 5
H3GA0001228	1	31,202,546	$1.89 \times 10^{-7}$	0.44	0.21	<i>MAP3K5</i>	mitogen-activated protein kinase kinase kinase 5
ALGA0003540	1	60,869,380	$5.44 \times 10^{-7}$	0.15	0.15	<i>NT5E</i>	5'-nucleotidase, ecto
ALGA0003690	1	64,094,344	$4.04 \times 10^{-7}$	0.26	0.18	<i>GABRR2</i>	gamma-aminobutyric acid (GABA) A receptor, rho 2
ALGA0108119	9	120,773,379	$5.10 \times 10^{-7}$	0.47	0.21	<i>ENSSSCG00000022338</i>	Uncharacterized protein
ALGA0054579	9	120,972,491	$3.86 \times 10^{-7}$	0.47	0.21	<i>ENSSSCG00000022338</i>	Uncharacterized protein
DRGA0009690	9	121,359,360	$9.48 \times 10^{-7}$	0.47	0.20	<i>ENSSSCG00000022338</i>	Uncharacterized protein
H3GA0028049	9	121,407,081	$7.61 \times 10^{-7}$	0.48	0.20	<i>ENSSSCG00000022338</i>	Uncharacterized protein
H3GA0038097	13	21,3691,291	$1.29 \times 10^{-6}$	0.12	0.13	<i>DSCAM</i>	Down syndrome cell adhesion molecule

<sup>1</sup>: Pig Chromosome ; <sup>2</sup>: Minor Allele Frequency ; <sup>3</sup>: Additive variance explained by SNP.

**Table 3 Significant SNPs associated with residual feed intake 2 (RFI2)**

SNP	SSC <sup>1</sup>	Position	p-value	MAF <sup>2</sup>	Var.exp <sup>3</sup>	Nearest gene	Gene name
MARC0013869	0	0	$3.15 \times 10^{-7}$	0.474	0.20		
ASGA0082396	1	30,941,797	$5.64 \times 10^{-7}$	0.448	0.19	MAP3K5	Mitogen-activated protein kinase kinase kinase 5
ALGA0106992	1	30,962,276	$4.68 \times 10^{-7}$	0.448	0.19	MAP3K5	Mitogen-activated protein kinase kinase kinase 5
ASGA0094502	1	30,978,281	$5.09 \times 10^{-7}$	0.449	0.19	MAP3K5	Mitogen-activated protein kinase kinase kinase 5
ALGA0107451	1	31,008,523	$4.54 \times 10^{-7}$	0.450	0.19	MAP3K5	Mitogen-activated protein kinase kinase kinase 5
H3GA0001228	1	31,202,546	$3.67 \times 10^{-7}$	0.443	0.20	MAP3K5	Mitogen-activated protein kinase kinase kinase 5
ALGA0108119	9	120,773,379	$3.64 \times 10^{-7}$	0.473	0.20	ENSSSCG00000022338	Uncharacterized protein
ALGA0054579	9	120,972,491	$2.49 \times 10^{-7}$	0.473	0.20	ENSSSCG00000022338	Uncharacterized protein
DRGA0009690	9	121,359,360	$6.02 \times 10^{-7}$	0.475	0.20	ENSSSCG00000022338	Uncharacterized protein
H3GA0028049	9	121,407,081	$5.02 \times 10^{-7}$	0.475	0.20	ENSSSCG00000022338	Uncharacterized protein
ASGA0089950	13	210,531,047	$2.87 \times 10^{-7}$	0.113	0.13	HLSC	Holocarboxylase synthetase
ASGA0097399	13	210,534,054	$3.07 \times 10^{-7}$	0.114	0.13	HLSC	Holocarboxylase synthetase

<sup>1</sup>: Pig Chromosome ; <sup>2</sup>: Minor Allele Frequency ; <sup>3</sup>: Additive variance explained by SNP.

respectively. On chromosome 1, each LD block has similar frequency for major haplotypes with frequency ranging from 0.44 to 0.55. On chromosome 9, haplotype 2112212 for LD block 1 and haplotype 2121112 for LD block 2 appeared more often than other haplotypes (1 is minor allele and 2 is major allele).

Due to high number of common SNPs for both RFI1 and RFI2, we decided to use positional candidate genes found for significant/suggestive SNPs for RFI2 for functional annotation. There were, 619 positional candidate genes for RFI2 to these significant/suggestive SNPs and were used as input in DAVID (Additional file 1). The functional annotation of positional candidate genes based on biological process and KEGG pathways involving in RFI2 is shown in Tables 5 and 6, respectively. The GO termed regulation of protein metabolic process, cellular lipid metabolic process and lipid metabolic process showed significant overrepresentation of genes statistically associated with RFI ( $p < 0.05$ ). The gap junction, phosphatidylinositol signaling system inositol phosphate metabolism and insulin signaling pathways were statistically associated with RFI ( $p < 0.05$ ).

## Discussion

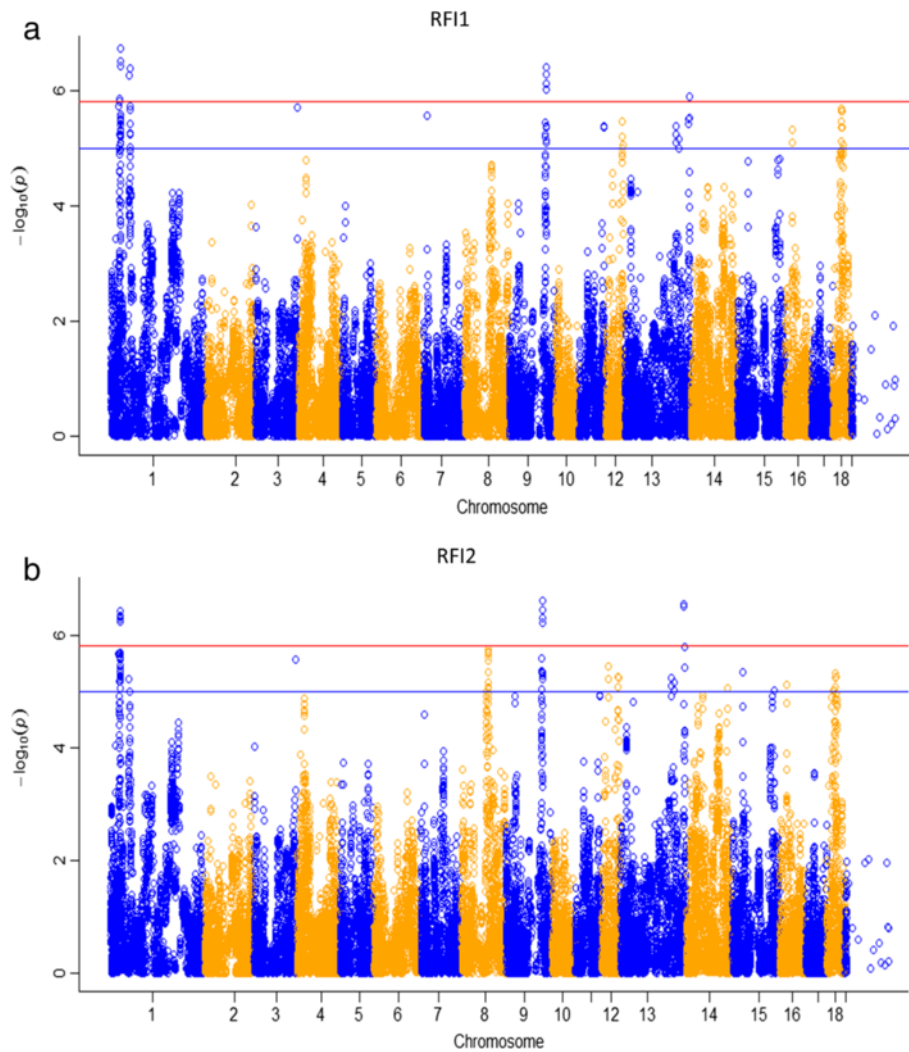
### QTLs, LD blocks and candidate genes for residual feed intake

Despite differences in the estimation models, RFI had very high genetic correlation with each other ( $r_g = 0.96$  [7]). Hence it is not surprising that the GWAS results for RFI1 and RFI2 show highly similar genetic architecture (numbers of top SNPs and their genomic positions). Many significant SNPs for both RFI were located in the same genomic regions on SSC 1, 9 and 13, and approximately 80% of suggestive SNPs (124 SNPs) were also found to be associated with both traits. Likewise, Nkrumah *et al.* [39] also reported many concordant

QTLs between RFI based phenotype (RFI<sub>p</sub>) and RFI based genotype (RFI<sub>g</sub>) in cattle. These authors detected 14 common and 3 distinct QTLs for the two RFI measures. Their high genetic similarity makes it difficult to find unique QTL and candidate genes for each trait.

Two most interesting chromosomal regions for RFI were 30.5-31.5 Mb on SSC 1 and 120.5-121.5 Mb on SSC 9. Seven and five highly significant SNPs for RFI1 and RFI2 were found in chromosomal regions 30.5-31.5 Mb on SSC 1. *MAP3K5* (mitogen-activated protein kinase 5) gene, located from 30,747 to 31,011 kb on SSC 1, might be an interesting candidate gene. *MAP3K5*, also known as apoptosis signal-regulating kinase 1 (*ASK1*), acts as an essential component of the mitogen-activated protein kinase signal transduction pathway in humans [40], it mediates signaling for determination of cell fate such as differentiation and survival in mice [41]. The effect of *MAP3K5* (or in generally, MAPK) on controlling feed intake or RFI may be mediated by variety of pathways such as hormones and growth factors that act through receptor tyrosine kinases (e.g. insulin, epidermal growth factor (EGF) [42]), cytokine receptors (e.g. growth hormone) to vasoactive peptides acting through G protein-coupled, seven-transmembrane receptors (e.g. endothelin) and so on [43]. In cattle, the majority of up-regulated genes in low RFI beef was stimulated by MAPKs [20]. Functional approaches to validate *MAP3K5* as a candidate gene for RFI in pigs is necessary.

The LD block helps to get more detail in QTL regions because several significant/suggestive SNPs were found in the same LD block. Therefore, it could imply that the causative mutation might be in these blocks. These approaches have been extensively applied in many species [31,44,45]. It is also worthy to note that two candidate SNPs (ALGA0106992 and ALGA0094502) are tightly linked ( $D' = 0.98$ ) in the LD block 3 (Figure 3). Functional



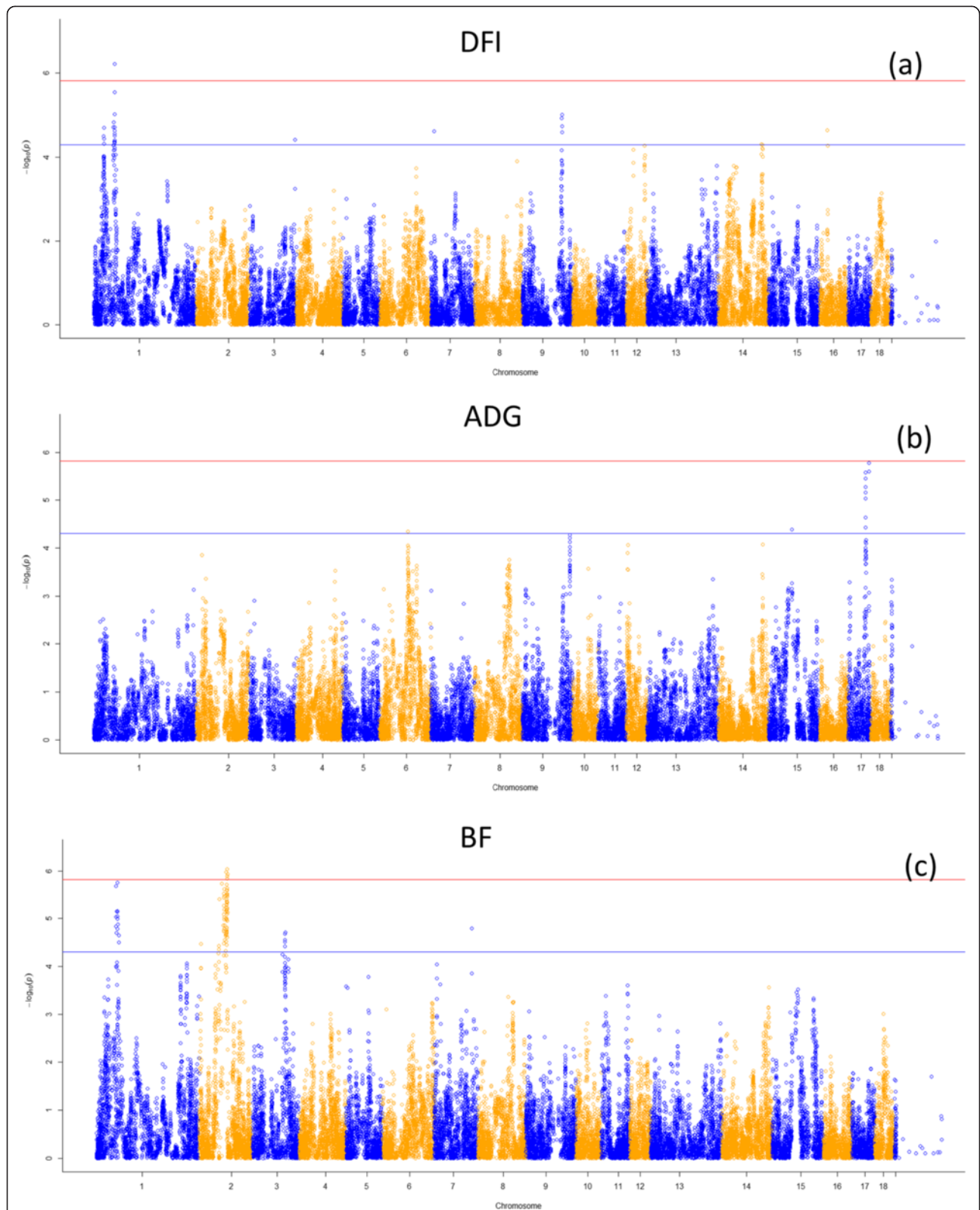
**Figure 1** Manhattan plot of genome-wide  $p$ -values of association for residual feed intake. Legends: **(a)**: Manhattan plot for Residual feed intake 1 (RFI1), **(b)**: Manhattan plot for Residual feed intake 2 (RFI2). The horizontal red and blue lines represent the genome-wide significance threshold at  $p < 1.52 \times 10^{-6}$  and  $p < 5 \times 10^{-5}$ , respectively.

**Table 4** The significant SNPs associated with component traits of residual feed intake

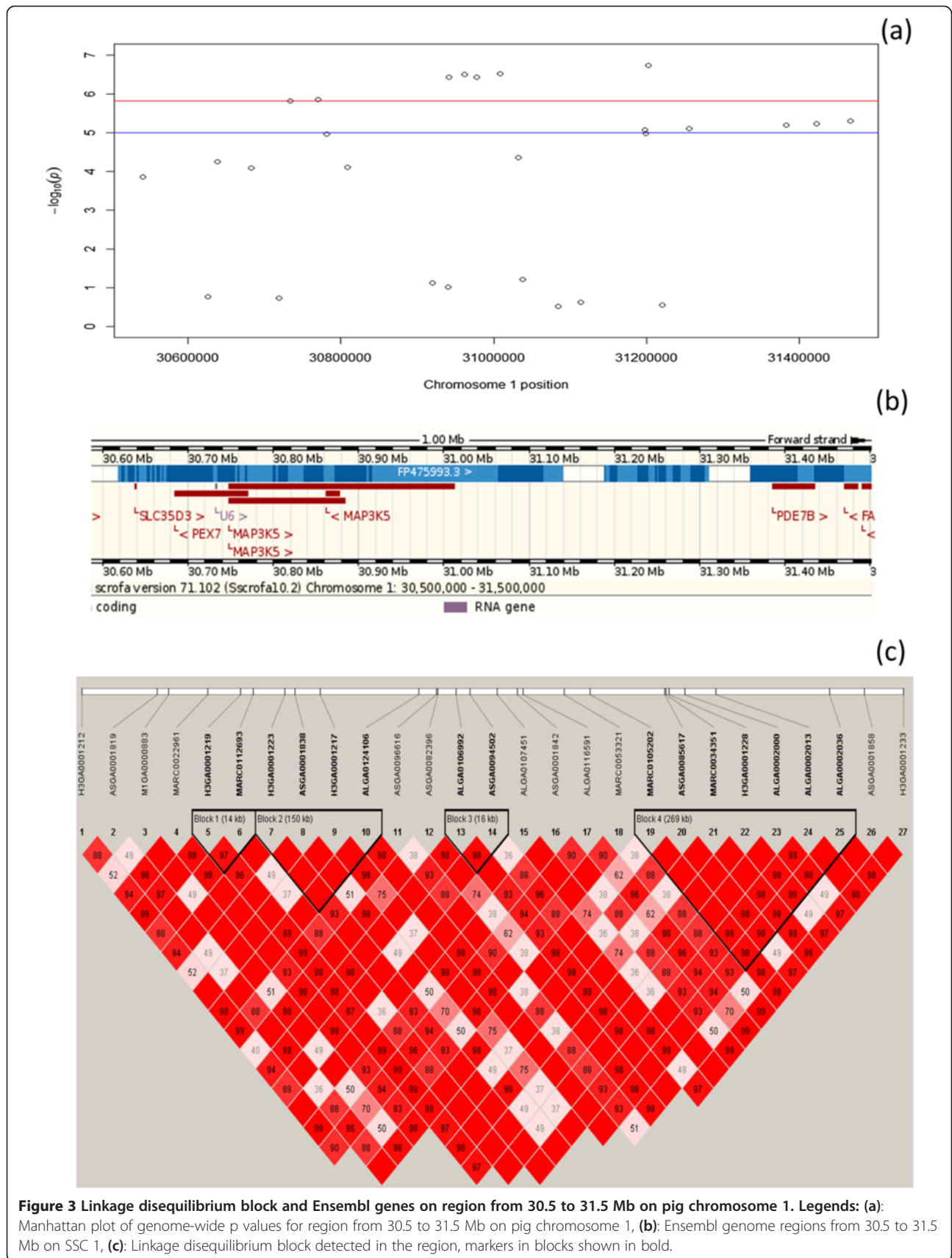
Traits <sup>1</sup>	SNP <sup>2</sup>	SSC <sup>3</sup>	Pos	MAF	p-value	Nearest gene	Gene name
DFI	ALGA0003690	1	64,094,344	0.28	$6.10 \times 10^{-7}$	<i>GABRR2</i>	Gamma-aminobutyric acid (GABA) A receptor, rho 2
ADG*	ASGA0077977	17	63,740,625	0.32	$1.67 \times 10^{-6}$	<i>CBLN4</i>	Cerebellin 4 precursor
BF	ALGA0014061	2	84,789,103	0.07	$9.06 \times 10^{-7}$	<i>ENSSSCG00000024586</i>	Novel gene
BF	ALGA0014028	2	82,276,435	0.08	$1.06 \times 10^{-6}$	<i>RGS14</i>	Regulator of G-protein signaling 14
BF	ASGA0010625	2	86,139,077	0.08	$1.17 \times 10^{-6}$	<i>COL4A3BP</i>	Collagen type IV alpha-3-binding protein
BF	ALGA0014098	2	85,830,426	0.07	$1.27 \times 10^{-6}$	<i>ANKRD31</i>	Ankyrin repeat domain 31
BF	DRGA0003063	2	85,710,507	0.07	$1.45 \times 10^{-6}$	<i>GCNT4</i>	Glucosaminyl (N-acetyl) transferase 4, core 2
BF	MARC0074327	2	85,539,255	0.07	$1.46 \times 10^{-6}$	<i>FAM169A</i>	Family with sequence similarity 169, member A

<sup>1</sup>: DFI: Daily feed intake, ADG: Average daily gain, BF: Backfat <sup>2</sup>: Pig Chromosome ; <sup>3</sup>: Minor Allele Frequency.

\*: The tighest associated SNP for ADG have not pass GW significant threshold ( $1.52 \times 10^{-6}$ ).



**Figure 2** Manhattan plot of genome-wide  $p$ -values of association for component traits of residual feed intake. Legends: (a): Manhattan plot for Daily feed intake (DFI), (b): Manhattan plot for Average daily gain (ADG), (c): Manhattan plot for backfat (BF). The horizontal red and blue lines represent the genome-wide significance threshold at  $p < 1.52 \times 10^{-6}$  and  $p < 5 \times 10^{-5}$ , respectively.



**Figure 3** Linkage disequilibrium block and Ensembl genes on region from 30.5 to 31.5 Mb on pig chromosome 1. **Legends:** (a): Manhattan plot of genome-wide p values for region from 30.5 to 31.5 Mb on pig chromosome 1, (b): Ensembl genome regions from 30.5 to 31.5 Mb on SSC 1, (c): Linkage disequilibrium block detected in the region, markers in blocks shown in bold.



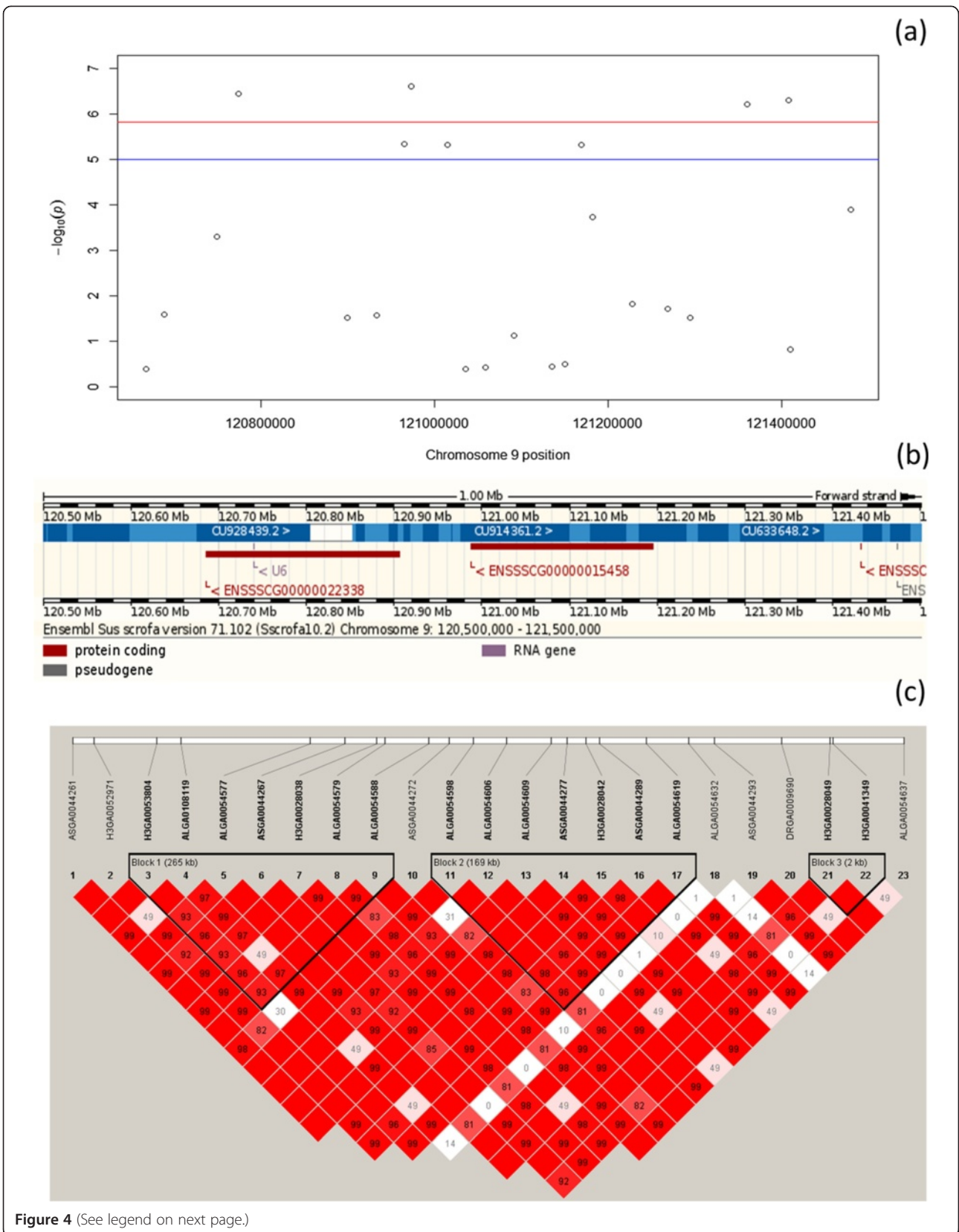


Figure 4 (See legend on next page.)

(See figure on previous page.)

**Figure 4 Linkage disequilibrium block and Ensembl genes in the region from 120.5 to 121.5 Mb on pig chromosome 9.** Legends: (a): Manhattan plot of genome-wide *p*-values for RFI 1 in the region from 120.5 to 121.5 Mb on pig chromosome 9, (b): Ensembl genome region from 30.5 to 31.5 Mb on pig chromosome 9, (c): Linkage disequilibrium block detected in the region, markers in these blocks were shown in bold.

validation for *MAP3K5* as a candidate gene for RFI needs to also take into account the different haplotypes and linkage phases. Close to *MAP3K5* gene, the variant MARC0112693 was also significantly associated with RFI. The variant was located in the intronic region on *PEX7* gene, which encodes the cytosolic receptor for the set of peroxisomal matrix enzymes targeted to the organelle by the peroxisome targeting signal 2. The gene plays an essential role in peroxisomal protein import and defects in this gene cause peroxisome biogenesis disorders (PBDs), which are characterized by multiple defects in peroxisome function in human [46]. Moreover, the chromosomal region is homologous with human chromosomal region (HSA) 6q.23. The HSA6q.23 contained *MAP3K5*–*PEX7* haplotype block which was found associated with age at onset in Huntington's disease [47,48] and with modulation of fetal hemoglobin levels in sickle cell anemia [49]. Therefore, RFI might not be controlled by a single gene but by LD block in the region.

On chromosome 1, two SNPs were significantly associated with only RFI1. Interestingly, the ALGA0003690 polymorphism were also found to be significantly associated with DFI in the same population [31]. The mutation is located in intronic region of *GABRR2* gene. The gene encodes a receptor of gamma-aminobutyric acid (GABA), the major inhibitory neurotransmitter in the vertebrate brain. Therefore, *GABRR2* might be an interesting candidate gene for both daily feed intake and RFI1. Another possible candidate gene for RFI1 is *NT5E*, which encodes for a protein that catalyzes the conversion of extracellular nucleotides to membrane-permeable nucleosides in human [50]. Due to *NT5E* using AMP as a substrate, the involvement of their gene with residual feed intake might be via energy balance.

The second interesting region for RFI is 120.5–121.5 Mb on SSC 9. Four highly significant SNPs for RFI1 and three for RFI2 were found in the region. The SNPs were located in all different LD blocks (Figure 3), and were highly linked to several suggestive SNPs (Additional files 2 and 3). However, both of these

nearest genes found in the region have not been functionally studied.

A significant SNP for RFI1 and two significant SNPs for RFI2 were found on SSC 13. Because these SNPs are very closely located to each other, the region might contain a QTL for both traits. Notably, a marker H3GA0038097, which was significantly associated with RFI1, was also significantly associated with RFI2 ( $p = 1.62 \times 10^{-6}$ ). The nearest gene, *DSCAM* encodes for Down syndrome cell adhesion molecule, which plays a role in neuronal self-avoidance as discovered in a mice model [51]. Recently, Garrett *et al.* [52] reviewed the role of *DSCAMs* proteins and suggested that their importance of balancing these developmental mechanisms. Close to the *DSCAM* gene, the *HLCS* (holocarboxylase synthetase) gene plays an important role in gluconeogenesis, fatty acid synthesis and branched chain amino acid catabolism in human [53]. Both of *DSCAM* and *HLCS* have not been functionally characterized in pigs, however, since their function involves developmental balance and glucose/fatty acid regulations, they might be important candidate genes for RFI in pigs.

A suggestive QTL spanned a region from 83–92 Mb on SSC 8 and contained five SNPs for RFI1 and 26 suggestively significant SNPs for RFI2 which may be interesting. Sahana *et al.* [26] also found a significant SNP for FCR in the same region in the same Duroc population. The other suggestive QTL regions are 54–56 Mb on SSC 12 containing eight SNPs for both RFI and 26–36 Mb on SSC 18 containing 24 SNPs for RFI1 and 20 SNPs for RFI2. These QTL regions also contained a number of potential candidate genes for RFI. The QTL on SSC12 for RFI was close to QTL for FCR in Meshan  $\times$  Large White cross populations previously recorded by [54] and QTL on SSC 18 was coincided with QTL for FCR on chromosome in the genetically diverse founder groups Meishan, Pietrain and European Wild Boar previously by [55]. However, more analyses are needed to confirm if they are true QTL for RFI.

**Table 5 The results of Gene Ontology (GO) analysis including genes in 0.5 Mb flanking size to SNPs with  $p < 5.0 \times 10^{-5}$**

Go sub-ontology	GO term accession	GO term description	Involved genes	David p-value
Biological process	GO:0051246	Regulation of protein metabolic process	<i>UBE2L3, PRKAR1A, UBE2J1</i>	0.03
Biological process	GO:0044255	Cellular lipid metabolic process	<i>SGMS1, PTEN, ALOX12, PRKAG3</i>	0.04
Biological process	GO:0006629	Lipid metabolic process	<i>SGMS1, PTEN, ALOX12, PRKAG3, PLCD4</i>	0.05

**Table 6 Results of functional annotation on KEGG pathway including genes 0.5 Mb flanking size to SNPs with  $p < 5.0 \times 10^{-5}$**

Term	Pathway name	Number of involved genes	Involved genes	David p-value
ssc05213	Endometrial cancer	5	<i>TCF7L2, ERK2, PTEN, PIK3R5, TP53</i>	0.001
ssc04540	Gap junction	6	<i>TUBA4A, ERK2, ADRB1, PRKG1, TUBA3D, TUBA1A</i>	0.002
ssc04070	Phosphatidylinositol signaling system	4	<i>PTEN, PIK3R5, PLCD4, PI4KA</i>	0.013
ssc00562	Inositol phosphate metabolism	4	<i>PTEN, PIK3R5, PLCD4, PI4KA</i>	0.015
ssc05214	Glioma	4	<i>ERK2, PTEN, PIK3R5, TP53</i>	0.026
ssc05215	Prostate cancer	5	<i>ERK2, PTEN, PIK3R5, TP53, TCF7L2</i>	0.027
ssc04910	Insulin signaling pathway	5	<i>ERK2, SLC2A4, PIK3R5, PRKAR1A, PRKAG3</i>	0.031
ssc05218	Melanoma	4	<i>ERK2, PTEN, PIK3R5, TP53</i>	0.033

### Systems genetics of daily feed intake, average daily gain and backfat

#### QTLs and candidate genes for daily feed intake

We have identified ALGA0003690 (G/A) significantly associated with DFI based on genotype records from 2008 to 2011 in the same population [31]. Although, we have added 300 genotyped animals (recorded in 2012), we still detected only this SNP passed genome-wide significant threshold in the current study. However, we also detected more suggestive loci on SSC 3, 14 and 16 in the current study. The most interesting putative candidate gene for DFI might be Gamma-aminobutyric acid receptor subunit rho-2 (*GABRR2*) gene. The gene encode for the most important inhibitory neurotransmitter in the vertebrate central nervous system and it plays function in feed/food intake as discussed in [31]. Some other nearest genes may be interesting are G protein pathway suppressor 2 (*GPS2*) gene on SSC 3, alpha-2A receptor gene (*ADRA2A*) on SSC 14 and Nipped-B homolog gene (*NIPBL*) in SSC 16 (Table S). For instance, *ADRA2A* is one of candidate genes for obesity and diabetes [56] and variants in the gene was associated with satiation [57]. Because the pig is a model for human obesity research [31], *ADRA2A* is worthy to functionally investigate.

#### QTLs and candidate genes for average daily gain

Understanding the genomics controlling components traits of RFI helps to better prioritize candidate SNP/genes for further investigation. Amongst suggestive regions found associated with ADG, none of them overlapped with QTL regions detected for RFI. Therefore, markers assisted selection based on candidate genes for RFI (identified in this population) would not have influence on daily gains of pigs. The ASGA007797 marker was tightly associated with ADG ( $p = 1.67 \times 10^{-6}$ ) (Table 5) and was located within brain functioned *CBLN4* gene. The gene encodes for new transneuronal cytokines [58] which have been highly involved in insulin secretion in rats [59]. On chromosome 17, we also

identified a region from 53.4-54.2 Mb which contains 8 suggestive SNPs for ADG. Several genes in this region might be interesting such as *NCOA5* (Nuclear receptor coactivator 5), *SLC35C2* (Solute carrier family 35, member C2) and *CD40* (TNF receptor superfamily member 5) might interesting for further investigation. Moreover, on chromosome 6, a suggestive SNP was found in QTL regions detected for ADG in several resource populations [60,61]. This SNP was located close to Ras-related GTP binding C (*RRAGC*). *RRAGC* encodes for a members of Rag small GTPases, which regulate the growth-controlling TOR signaling pathway (reviewed in [62]). However, many nearest genes for ADG have not been functionally characterized in pigs.

#### QTLs and candidate genes for backfat

In pigs, backfat is one of the phenotypes that have been studied in many resource populations. The QTLs for BF have been mapped in every pig chromosome. Several GWAS studies have been also performed for BF such as those reported by Fontanesi *et al.* [63]; they reported candidate genes on SSC 7 and 18 associated with Italian heavy pigs. Other studies include Onteru *et al.* [24] who reported fat metabolism genes on SSC 3, 7 and 18 for BF in Yorkshire pigs and Okumura *et al.* [64] who reported QTL on SSC 6 for backfat thickness in Japanese Duroc pig population. Because all significant SNPs (Table 5 and more than 70 suggestive QTLs (Additional file 1) for BF were located in the region of 81-86 Mb on SSC 2, we assumed that a QTL in this region was affecting BF. In this region, many genes have been shown to have important role in fat deposition and lipid metabolism. For instance, 3-hydroxy-3-methylglutaryl-CoA reductase (*HMGCR*) gene spanned from 85,967 to 85,990 kb on SSC 2 which encodes for the well-known enzyme regulating the synthesis of cholesterol in humans and other species. Pigs with divergent backfat thickness also expressed different *HMGCR* activity in liver [65]. Canovas *et al.* [66] found that the mutation in pig *HMGCR* gene (807A > G) was associated with not

only cholesterol but also intramuscular fat content in commercial Duroc pig line. It is also worthy to note that we also detected suggestive effect on BF in the chromosomal region at 63-64 Mb on SSC 1 where a QTL for DFI and RFI was reported. Therefore, the characterization of these regions needs to consider the effects of pleiotropic QTL. Moreover, the suggestive QTL at 116.3 Mb on SSC 7 was also very close to the region where Orteru *et al.* [24] detected QTL for BF at 112 Mb position. *GALA* gene encodes to galactosylceramidase might be interesting candidate gene because its protein (an enzyme) breaks down galactolipids and plays role in lipid metabolism in kidney and epithelial cells of small intestine and colon [67].

#### **Inferring pathways and biological functions of nearby genes for residual feed intake**

Based on biological function, several nearby genes of significant SNPs for RFI (*SGMS1*, *PTEN*, *ALOX12*, *PRKAG3*, *PLCD4*) were also clustered in lipid metabolic process (GO:0044255 and GO:0006629) in the current study (Table 5). The relation between lipid metabolism and residual feed intake has been shown in pigs [23] and cattle [68]. Moreover, we also recorded nearby genes involved in regulation of protein metabolic process (*UBE2L3*, *PRKARIA*, *UBE2J1*). *PRKARIA* encodes for protein kinase A (PKA, aka cAMP-dependent protein kinase) which is involved in the regulation of lipid and glucose metabolism and is a component of the signal transduction mechanism of certain G protein-coupled receptors in humans [69]. Malek *et al.* [70] characterized the porcine prepro-orexin gene and found high linkage among *PRKARIA*, *GHI* and *BRCA1* genes. The same authors speculated that *PRKARIA* is a candidate gene for feed intake. Nevertheless, lipid metabolic process is a very broad term, and therefore it might be worthy to further investigate which sub-terms in the process are involved in RFI metabolism.

Interestingly, we also found that the genes clustered in insulin signaling pathway. In the insulin signaling pathway, insulin binds to its receptor resulting in the tyrosine phosphorylation of insulin receptor substrates (IRS) by the insulin receptor tyrosine kinase (INSR). This allows association of IRSs with the regulatory subunit of phosphoinositide 3-kinase (PI3K). PI3K activates 3-phosphoinositide-dependent protein kinase 1 (*PDK1*), which activates Akt, a serine kinase. Akt in turn deactivates glycogen synthase kinase 3 (GSK-3), leading to activation of glycogen synthase and thus glycogen synthesis (KEGG pathway, term: ssc04910). Several studies also mentioned that insulin signaling pathway plays important roles in controlling residual feed intake in cattle [12,71] and in pigs [72].

Another potentially relevant pathway is gap junction which consists of 6 nearby genes of significant SNPs for

RFI (Table 5). Gap junctions contain intercellular channels that allow direct communication between the cytosolic compartments of adjacent cells [73]. Regulation of feed intake is a complex process, which not only happens inside the cells but also in interactions among cells. *PRKG1* is one of the genes involved in gap junction and is also a candidate gene for RFI in cattle [11] and for intramuscular fat content in pigs [74]. Nevertheless, implying pathways based on GWAS data analyses alone needs caution because many other factors can have an effect on the results such as hidden confounders, threshold for significant p-value of SNP from GWAS data, length of flanking region to get gene list, the statistical test methods and so on [75] and a systems biology approach using additional more or less independent data to verify or add information to the findings would be one of the best approaches to profile pathways underlying complex traits [8].

In summary, this study used comprehensive GWAS and pathway approaches to reveal genetic variants, and genes that control feed efficiency (RFI) and the related traits in pigs and possible biological pathways in which these genes are exerting their effects. This study confirmed highly similar genetic mechanisms underlying different measurement of RFI; however, it could not find distinct genetic markers for RFI2. Therefore, including back fat in the RFI models was not important for this particular data and analyses. In the context of genomic selection for feed efficiency, the estimated magnitude and direction of SNP effects could potentially be useful for specifying more informative prior information in genomic prediction/selection models to increase genetic gain.

#### **Conclusion**

This study revealed possible genetic architecture and potential biological pathways for a feed efficiency measure, RFI in pigs. We identified two important genomic regions including 30.5-31.5 Mb on pig SSC 1 and 120.5-121.5 Mb on SSC 9 for RFI. We also conclude that there is a high similarity of genetic architecture between RFI1 and RFI2. Key positional candidate genes have been found: *MAP3K5*, *PEX7*, *ENSSSCG0000022338* and *DSCAM* for both RFI measures. We also detected several novel QTLs for other production traits including DFI, ADG and BF that were components of RFI measure. Systems genetic analyses and functional annotation of nearby genes confirmed an important role of insulin signaling pathway in regulation of RFI and revealed some other possible pathways such as gap junction or inositol phosphate metabolism. Therefore, this study offered important knowledge of the potential candidate genes, biomarkers, genetic architecture and biological pathways for feed efficiency measures.

## Additional files

**Additional file 1: Suggestive SNPs associated with RFI and its component traits, respectively.**

**Additional file 2: Frequency of each haplotype for different LD blocks on pig chromosome 1.**

**Additional file 3: Frequency of each haplotype for different LD blocks on pig chromosome 9.**

## Abbreviations

EBV: Estimated breeding values; dEBV: Deregressed estimated breeding values; GWAS: Genome-wide association study; LD: Linkage disequilibrium; SNP: Single nucleotide polymorphism; RFI: Residual feed intake; Bp: Base pairs; kb: Kilo base pairs; Mb: Mega base pairs; MAF: Minor allele frequency; QTL: Quantitative trait locus.

## Competing interest

No conflict of interest.

## Authors' contributions

DND did the analysis with the help of HNK, TO and ABS. DND wrote the first draft of the manuscript. HNK conceived and designed this project and provided supervision for DND in all aspects including GWAS, systems genetics and bioinformatics work. JJ and TM provided useful suggestions on analyses. All authors contributed to writing of this manuscript, read and approved the final manuscript.

## Acknowledgements

We would like to thank the Department of Breeding and Genetics of the Danish Pig Research Centre for providing all data for the research reported in this study. We thank Dr Goutam Sahana at Aarhus University for help with some computational steps. Duy Ngoc Do is a PhD student funded by the Department of Clinical Veterinary and Animal Sciences, Faculty of Health and Medical Sciences, University of Copenhagen. Haja N. Kadarmideen thanks EU-FP7 Marie Curie Actions – Career Integration Grant (CIG-293511) for partially funding his time spent on this research.

## Author details

<sup>1</sup>Section of Animal Genetics, Bioinformatics and Breeding, Department of Veterinary Clinical and Animal Sciences, Faculty of Health and Medical Sciences, University of Copenhagen, 1870 Frederiksberg C, Denmark. <sup>2</sup>Danish Agriculture & Food Council, Pig Research Centre, Axelstorv 3, 1609 Copenhagen V, Denmark. <sup>3</sup>Department of Molecular Biology and Genetics, Aarhus University, Blichers Alle 20, 8830 Tjele, Denmark.

Received: 27 September 2013 Accepted: 5 February 2014

Published: 17 February 2014

## References

1. Cai W, Casey DS, Dekkers JCM: Selection response and genetic parameters for residual feed intake in Yorkshire swine. *J Anim Sci* 2008, **86**(2):287–298.
2. Kennedy B, Van der Werf J, Meuwissen T: Genetic and statistical properties of residual feed intake. *J Anim Sci* 1993, **71**(12):3239–3250.
3. Gilbert H, Bidanel JP, Gruand J, Caritez JC, Billon Y, Guillouet P, Lagant H, Noblet J, Sellier P: Genetic parameters for residual feed intake in growing pigs, with emphasis on genetic relationships with carcass and meat quality traits. *J Anim Sci* 2007, **85**(12):3182–3188.
4. Saintilan R, Mérour I, Brossard L, Tribout T, Dourmad JY, Sellier P, Bidanel J, van Milgen J, Gilbert H: Genetics of residual feed intake in growing pigs: relationships with production traits, and nitrogen and phosphorus excretion traits. *J Anim Sci* 2013, **91**(6):2542–2554.
5. Muro-Reyes A, Gutierrez-Banuelos H, Diaz-Garcia LH, Gutierrez-Pina FJ, Escareno-Sanchez LM, Banuelos-Valenzuela R, Medina-Flores CA, Corral Luna A: Potential environmental benefits of residual feed intake as strategy to mitigate methane emissions in sheep. *J Anim Vet Adv* 2011, **10**(12):1551–1556.
6. Carberry CA, Kenny DA, Han S, McCabe MS, Waters SM: Effect of phenotypic residual feed intake and dietary forage content on the rumen microbial community of beef cattle. *Appl Environ Microbiol* 2012, **78**(14):4949–4958.
7. Do DN, Strathe AB, Jensen J, Mark T, Kadarmideen HN: Genetic parameters for different measures of feed efficiency and related traits in boars of three pig breeds. *J Anim Sci* 2013, **91**(9):4069–4079.
8. Kadarmideen HN: Genetical systems biology in livestock: application to gonadotrophin releasing hormone and reproduction. *IET Syst Biol* 2008, **2**(6):423–441.
9. Kadarmideen HN, von Rohr P, Janss LL: From genetical genomics to systems genetics: potential applications in quantitative genomics and animal breeding. *Mamm Genome* 2006, **17**(6):548–564.
10. Barendse W, Reverter A, Bunch RJ, Harrison BE, Barris W, Thomas MB: A validated whole-genome association study of efficient food conversion in cattle. *Genetics* 2007, **176**(3):1893–1905.
11. Sherman E, Nkrumah J, Moore S: Whole genome single nucleotide polymorphism associations with feed intake and feed efficiency in beef cattle. *J Anim Sci* 2010, **88**(1):16–22.
12. Rolf M, Taylor J, Schnabel R, McKay S, McClure M, Northcutt S, Kerley M, Weaber R: Genome-wide association analysis for feed efficiency in Angus cattle. *Anim Genet* 2012, **43**(4):367–374.
13. Lindholm-Perry A, Kuehn L, Smith T, Ferrell C, Jenkins T, Freetly H, Snelling W: A region on BTA14 that includes the positional candidate genes LYPLA1, XKR4 and TMEM68 is associated with feed intake and growth phenotypes in cattle. *Anim Genet* 2012, **43**(2):216–219.
14. Karisa BK, Thomson J, Wang Z, Stothard P, Moore SS, Plastow GS: Candidate genes and single nucleotide polymorphisms associated with variation in residual feed intake in beef cattle. *J Anim Sci* 2013, **91**(8):3502–3513.
15. Sherman EL, Nkrumah JD, Li C, Bartusiak R, Murdoch B, Moore SS: Fine mapping quantitative trait loci for feed intake and feed efficiency in beef cattle. *J Anim Sci* 2009, **87**(1):37–45.
16. Sherman EL, Nkrumah JD, Murdoch BM, Moore SS: Identification of polymorphisms influencing feed intake and efficiency in beef cattle. *Anim Genet* 2008, **39**(3):225–231.
17. Lindholm-Perry AK, Kuehn LA, Snelling WM, Smith TPL, Ferrell CL, Jenkins TG, Andy King D, Shackelford SD, Wheeler TL, Freetly HC: Genetic markers on BTA14 predictive for residual feed intake in beef steers and their effects on carcass and meat quality traits. *Anim Genet* 2012, **43**(5):599–603.
18. Lu D, Miller S, Sargolzaei M, Kelly M, Vander Voort G, Caldwell T, Wang Z, Plastow G, Moore S: Genome-wide association analyses for growth and feed efficiency traits in beef cattle. *J Anim Sci* 2013, **91**(8):3612–3633.
19. Yao C, Spurlack DM, Armentano LE, Page CD, VandeHaar MJ, Bickhart DM, Weigel KA: Random Forests approach for identifying additive and epistatic single nucleotide polymorphisms associated with residual feed intake in dairy cattle. *J Dairy Sci* 2013, **90**(10):6716–6729.
20. Chen Y, Gondro C, Quinn K, Herd RM, Parnell PF, Vanselow B: Global gene expression profiling reveals genes expressed differentially in cattle with high and low residual feed intake. *Anim Genet* 2011, **42**(5):475–490.
21. Gilbert H, Riquet J, Gruand J, Billon Y, Fève K, Sellier P, Noblet J, Bidanel JP: Detecting QTL for feed intake traits and other performance traits in growing pigs in a Piétrain–Large White backcross. *Anim* 2010, **4**(08):1308–1318.
22. Fan B, Lkhagvadorj S, Cai W, Young J, Smith RM, Dekkers JCM, Huff-Lonergan E, Lonergan SM, Rothschild MF: Identification of genetic markers associated with residual feed intake and meat quality traits in the pig. *Meat Sci* 2010, **84**(4):645–650.
23. Lkhagvadorj S, Qu L, Cai W, Couture OP, Barb CR, Hausman GJ, Nettleton D, Anderson LL, Dekkers JC, Tuggle CK: Gene expression profiling of the short-term adaptive response to acute caloric restriction in liver and adipose tissues of pigs differing in feed efficiency. *Am J Physiol Regul Integr Comp Physiol* 2010, **298**(2):R494–507.
24. Onteru SK, Gorbach DM, Young JM, Garrick DJ, Dekkers JCM, Rothschild MF: Whole genome association studies of residual feed intake and related traits in the pig. *PLoS ONE* 2013, **8**(6):e61756.
25. Jiao S, Cassidy J, Maltecca C, Gray K, Holl J: Whole-genome association analysis for feed efficiency traits in Duroc pigs, Abstract no. 124, Proceedings of the ADSA-ASAS Joint Annual Meetings: The American Dairy Science Association, July, 08–12, 2013. Indiana; 2013.
26. Sahana G, Kadlecová V, Hornshøj H, Nielsen B, Christensen OF: A genome-wide association scan in pig identifies novel regions associated with feed efficiency trait. *J Anim Sci* 2013, **91**(3):1041–1050.

27. Gunsett G, Gunsett FC: **Linear index selection to improve traits defined as ratios.** *J Anim Sci* 1984, **59**:1185–1193.
28. Henderson CR: *Applications of Linear Models in Animal Breeding.* Guelph, Ontario, Canada: University of Guelph; 1984.
29. Garrick DJ, Taylor JF, Fernando RL: **Deregressing estimated breeding values and weighting information for genomic regression analyses.** *Genet Sel Evol* 2009, **41**.
30. Ostensen T, Christensen OF, Henryon M, Nielsen B, Su G, Madsen P: **Deregressed EBV as the response variable yield more reliable genomic predictions than traditional EBV in pure-bred pigs.** *Genet Sel Evol* 2011, **43**:38.
31. Do DN, Strathe AB, Ostensen T, Jensen J, Mark T, Kadarmideen HN: **Genome-wide association study reveals genetic architecture of eating behavior in pigs and its implications for humans obesity by comparative mapping.** *PLoS ONE* 2013, **8**(8):e71509.
32. Madsen P, Sørensen P, Su G, Damgaard LH, Thomsen H, Labouriau R: **DMU-a package for analyzing multivariate mixed models.** In *8th World Congress on Genetics Applied to Livestock Production.* 2006:247.
33. Duggal P, Gillanders E, Holmes T, Bailey-Wilson J: **Establishing an adjusted p-value threshold to control the family-wide type 1 error in genome wide association studies.** *BMC Genomics* 2008, **9**(1):516.
34. Consortium WTCC: **Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls.** *Nature* 2007, **447**(7145):661–678.
35. Barrett JC, Fry B, Maller J, Daly MJ: **Haploview: analysis and visualization of LD and haplotype maps.** *Bioinformatics* 2005, **21**(2):263–265.
36. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, et al: **The structure of haplotype blocks in the human genome.** *Science* 2002, **296**(5576):2225–2229.
37. R Development Team: *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing; 2008.
38. Huang DW, Sherman BT, Lempicki RA: **Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources.** *Nat Protoc* 2009, **4**(1):44–57.
39. Nkrumah JD, Sherman EL, Li C, Marques E, Crews DH Jr, Bartusiak R, Murdoch B, Wang Z, Basarab JA, Moore SS: **Primary genome scan to identify putative quantitative trait loci for feedlot growth rate, feed intake, and feed efficiency of beef cattle.** *J Anim Sci* 2007, **85**(12):3170–3181.
40. Chibon F, Mariani O, Derre J, Mairal A, Coindre JM, Guillou L, Sastre X, Pedeutour F, Aurias A: **ASK1 (MAP3K5) as a potential therapeutic target in malignant fibrous histiocytomas with 12q14-q15 and 6q23 amplifications.** *Genes Chromosomes Cancer* 2004, **40**(1):32–37.
41. Matsuzawa A, Nishitoh H, Tobiume K, Takeda K, Ichijo H: **Physiological roles of ASK1-mediated signal transduction in oxidative stress- and endoplasmic reticulum stress-induced apoptosis: advanced findings from ASK1 knockout mice.** *Antioxid Redox Signal* 2002, **4**(3):415–425.
42. Hayes MR, Lechner TM, Zhao SR, Lee GS, Chowansky A, Zimmer D, De Jonghe BC, Kanoski SE, Grill HJ, Bence KK: **Intracellular signals mediating the food intake-suppressive effects of hindbrain glucagon-like peptide-1 receptor activation.** *Cell Metab* 2011, **13**(3):320–330.
43. Kyriakis JM, Avruch J: **Mammalian mitogen-activated protein kinase signal transduction pathways activated by stress and inflammation.** *Physiol Rev* 2001, **81**(2):807–869.
44. Mogensen MS, Scheibye-Alsing K, Karlskov-Mortensen P, Proschowsky HF, Jensen VF, Bak M, Tommerup N, Kadarmideen HN, Fredholm M: **Validation of genome-wide intervertebral disk calcification associations in Dachshund and further investigation of the chromosome 12 susceptibility locus.** *Front Genet* 2012, **3**.
45. Cardon LR, Abecasis GR: **Using haplotype blocks to map human complex trait loci.** *Trends Genet* 2003, **19**(3):135–140.
46. Braverman N, Steel G, Obie C, Moser A, Moser H, Gould SJ, Valle D: **Human PEX7 encodes the peroxisomal PTS2 receptor and is responsible for rhizomelic chondrodysplasia punctata.** *Nat Genet* 1997, **15**(4):369–376.
47. Arning L, Monté D, Hansen W, Wiczorek S, Jagiello P, Akkad D, Andrich J, Kraus P, Saft C, Epplen J: **ASK1 and MAP2K6 as modifiers of age at onset in Huntington's disease.** *J Mol Med* 2008, **86**(4):485–490.
48. Li J-L, Hayden MR, Almquist EW, Brinkman RR, Durr A, Dodé C, Morrison PJ, Suchowersky O, Ross CA, Margolis RL, et al: **A genome scan for modifiers of Age at onset in huntington disease: the HD MAPS study.** *Am J Hum Genet* 2003, **73**(3):682–687.
49. Wyszynski D, Baldwin C, Cleves M, Amirault Y, Nolan V, Farrell J, Bisbee A, Kutlar A, Farrer L, Steinberg M: **Polymorphisms near a chromosome 6q QTL area are associated with modulation of fetal hemoglobin levels in sickle cell anemia.** *Cell Mol Biol (Noisy-le-grand)* 2004, **50**(1):23.
50. Hansen KR, Resta R, Webb CF, Thompson LF: **Isolation and characterization of the promoter of the human 5'-nucleotidase (CD73)-encoding gene.** *Gene* 1995, **167**(1–2):307–312.
51. Fuerst PG, Bruce F, Tian M, Wei W, Elstrott J, Feller MB, Erskine L, Singer JH, Burgess RW: **DSCAM and DSCAML1 function in self-avoidance in multiple cell types in the developing mouse retina.** *Neuron* 2009, **64**(4):484–497.
52. Garrett AM, Tadenev ALD, Burgess RW: **DSCAMs: restoring balance to developmental forces.** *Front Mol Neurosci* 2012, **5**:e. 00086.
53. Yang X, Aoki Y, Li X, Sakamoto O, Hiratsuka M, Kure S, Taheri S, Christensen E, Inui K, Kubota M, et al: **Structure of human holocarboxylase synthetase gene and mutation spectrum of holocarboxylase synthetase deficiency.** *Hum Genet* 2001, **109**(5):526–534.
54. Houston RD, Haley CS, Archibald AL, Rance KA: **A QTL affecting daily feed intake maps to Chromosome 2 in pigs.** *Mamm Genome* 2005, **16**(6):464–470.
55. Dragos-Wendrich M, Stratil A, Hojny J, Moser G, Bartenschlager H, Reiner G, Geldermann H: **Linkage and QTL mapping for Sus scrofa chromosome 18.** *J Anim Breed Genet* 2003, **120**:138–143.
56. Langberg EC, Seed Ahmed M, Efendic S, Gu HF, Ostenson CG: **Genetic association of adrenergic receptor alpha 2A with obesity and type 2 diabetes.** *Obesity (Silver Spring)* 2013, **21**(8):1720–1725.
57. Papathanasopoulos A, Camilleri M, Carlson PJ, Vella A, Nord SJ, Burton DD, Odunsi ST, Zinsmeister AR: **A preliminary candidate genotype-intermediate phenotype study of satiation and gastric motor function in obesity.** *Obesity (Silver Spring)* 2010, **18**(6):1201–1211.
58. Yuzaki M: **Cbln and C1q family proteins: new transneuronal cytokines.** *Cell Mol Life Sci* 2008, **65**(11):1698–1705.
59. Strowski MZ, Kaczmarek P, Mergler S, Wiedenmann B, Domin D, Szwalkowski P, Wojciechowicz T, Skrzypski M, Szczepankiewicz D, Szkudelski T, et al: **Insulinostatic activity of cerebellin—evidence from in vivo and in vitro studies in rats.** *Regul Pept* 2009, **157**(1–3):19–24.
60. Bidanel JP, Milan D, Iannuccelli N, Amigues Y, Boscher MY, Bourgeois F, Caritez JC, Gruand J, Le Roy P, Lagant H, et al: **Detection of quantitative trait loci for growth and fatness in pigs.** *Genetics, selection, evolution GSE* 2001, **33**(3):289–309.
61. Ruckert C, Bennewitz J: **Joint QTL analysis of three connected F2-crosses in pigs.** *Genetics, selection, evolution GSE* 2010, **42**:40.
62. Duran RV, Hall MN: **Regulation of TOR by small GTPases.** *EMBO reports* 2012, **13**(2):121–128.
63. Dufour CR, Levasseur M-P, Pham NHH, Eichner LJ, Wilson BJ, Charest-Marcotte A, Duguay D, Poirier-Héon J-F, Cermakian N, Giguère V: **Genomic convergence among ERRA, PROX1, and BMAL1 in the control of metabolic clock outputs.** *PLoS Genet* 2011, **7**(6):e1002143.
64. Okumura N, Matsumoto T, Hayashi T, Hirose K, Fukawa K, Itou T, Uenishi H, Mikawa S, Awata T: **Genomic regions affecting backfat thickness and cannon bone circumference identified by genome-wide association study in a Duroc pig population.** *Anim Genet* 2013, **44**(4):454–457.
65. Pond WG, Mersmann HJ: **Genetically diverse pig models in nutrition research related to lipoprotein and cholesterol metabolism.** In *Advances in swine in biomedical research*, Volume 2. Edited by Tumbleson ME, Schook LB. Springer; 1997:843–863.
66. Cánovas A, Quintanilla R, Gallardo D, Díaz I, Noguera JL, Ramírez O, Pena RN: **Functional and association studies on the pig HMGR gene, a cholesterol-synthesis limiting enzyme.** *Anim* 2010, **4**(02):224–233.
67. Zakharova E, Boukina TM: **Gene symbol: GALC. Disease: krabbe disease.** *Hum Gen* 2008, **124**(3):299.
68. Herd RM, Arthur PF: **Physiological basis for residual feed intake.** *J Anim Sci* 2009, **87**(14):E64–E71.
69. Nagasaki K, Iida T, Sato H, Ogawa Y, Kikuchi T, Saitoh A, Ogata T, Fukami M: **PRKAR1A mutation affecting cAMP-mediated G protein-coupled receptor signaling in a patient with acrodysostosis and hormone resistance.** *J Clin Endocrinol Metab* 2012, **97**(9):E1808–1813.
70. Malek M, Marklund S, Dyer C, Matteri R, Rothschild M: **Linkage and physical mapping of the porcine prepro-orexin gene.** *Mamm Genome* 2000, **11**(4):342–343.
71. Chen Y, Arthur P, Barchia I, Quinn K, Parnell P, Herd R: **Using gene expression information obtained by quantitative real-time PCR to**

- evaluate Angus bulls divergently selected for feed efficiency. *Anim Prod Sci* 2012, **52**(11):1058–1067.
72. Colditz I: Some mechanisms regulating nutrient utilisation in livestock during immune activation: an overview. *Anim Prod Sci* 2004, **44**(5):453–457.
73. Lampe PD, Lau AF: The effects of connexin phosphorylation on gap junctional communication. *Int J Biochem Cell Biol* 2004, **36**(7):1171–1186.
74. Hamill RM, McBryan J, McGee C, Mullen AM, Sweeney T, Talbot A, Cairns MT, Davey GC: Functional analysis of muscle gene expression profiles associated with tenderness and intramuscular fat content in pork. *Meat Sci* 2012, **92**(4):440–450.
75. Wang K, Li M, Hakonarson H: Analysing biological pathways in genome-wide association studies. *Nat Rev Genet* 2010, **11**(12):843–854.

doi:10.1186/1471-2156-15-27

**Cite this article as:** Do et al.: Genome-wide association and systems genetic analyses of residual feed intake, daily feed consumption, backfat and weight gain in pigs. *BMC Genetics* 2014 15:27.

**Submit your next manuscript to BioMed Central  
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

