

# SCIENTIFIC REPORTS



OPEN

## Association study and expression analysis of *CYP4A11* gene copy number variation in Chinese cattle

Mingjuan Yang<sup>1,2</sup>, Jingqiao Lv<sup>3</sup>, Liangzhi Zhang<sup>1</sup>, Mingxun Li<sup>1</sup>, Yang Zhou<sup>1</sup>, Xianyong Lan<sup>1</sup>, Chuzhao Lei<sup>1</sup> & Hong Chen<sup>1</sup>

Received: 08 July 2016

Accepted: 22 March 2017

Published: 11 May 2017

The identification of copy number variations (CNVs) allow us to explore genomic polymorphisms. In recent years, significant progress in understanding CNVs has been made in studies of human and animals, however, association and expression studies of CNVs are still in the early stage. It was previously reported that the Cytochrome P-450 4A11 (*CYP4A11*) gene is located within a copy number variable region (CNVR) that encompasses quantitative trait loci (QTLs) for economic traits like meat quality and milk production. So, this study was performed to determine the presence of *CYP4A11* CNV in six distinct cattle breeds, identify its relationship with growth, and explore the biological effects of gene expression. For three *CYP4A11* CNV types, Normal was more frequent than Gain or Loss. Association analysis revealed a positive effect of *CYP4A11* copy number on growth traits ( $P < 0.05$ ). One-way analysis of variance (ANOVA) analysis revealed that more *CYP4A11* copies increased the gene expression level. Moreover, overexpression of *CYP4A11* *in vitro* revealed its effect on lipid deposit. The data provide evidence for the functional role of *CYP4A11* CNV and provide the basis for future applications in cattle breeding.

After publication of two cattle genome maps (Btau\_4 and UMD3)<sup>1,2</sup>, single nucleotide polymorphisms (SNPs) were identified as a key factor causing genetic mutation in cattle, and abundant SNP information has been collected<sup>3–5</sup>. Simultaneously, many SNPs were reported to be associated with milk yield, milk protein percentage and milk fat percentage in *Bos taurus* dairy cattle<sup>6</sup>, and average daily gain, net feed intake, residual food intake and intramuscular fat percentage in *Bos taurus* beef cattle<sup>6,7</sup>. In addition, the distribution of quantitative trait loci (QTLs) were revealed, and artificial selection studies have been performed in dairy cattle to explore effects of genetic changes<sup>8,9</sup>. However, most SNP variants have only small effects on phenotypes, making it a priority to explain the missing heritability.

Two studies revealed that copy number variations (CNVs) that represent a significant source of genetic variation are widespread in human genomes<sup>10,11</sup>. Later, CNVs were redefined as gains or losses (insertions or deletions) larger than 50 bp in genomic sequence between two individuals of a species<sup>12</sup>. Soon, progress in understanding CNVs were made in human<sup>12,13</sup>, mice<sup>14,15</sup>, pig<sup>16,17</sup>, horse<sup>18,19</sup>, cattle<sup>20,21</sup>, goat<sup>22</sup> and chicken<sup>23</sup>, using different methods including whole genome sequencing, array-based comparative genomic hybridization (aCGH), and SNP array. However, despite the significant number of investigations identifying copy number variable regions (CNVRs), few have correlated this information with association and expression analysis.

Although individual SNPs may occur more frequently, CNVs are of larger size, and thus occupy a higher percentage of genomic sequence and have potentially larger effects on phenotypic variation, including changing gene structure and dosage, altering gene regulation, and exposing recessive alleles<sup>24</sup>. For example, changes in the copy number of cis-regulatory elements, particularly those related to developmental genes, have been shown to influence phenotype<sup>25</sup>. In some cases, deleting the normal allele in a heterozygous background led to exposure of a recessive allele that led to disease<sup>26</sup>. CNVs that include dosage-sensitive genes can alter gene expression levels resulting in clinical phenotypes. For example, in human, *PMP22* is located within the 1.4-Mb CMT1A (Charcot-Marie-Tooth 1A disease) region. Duplication of this region can lead to *PMP22* overexpression and the CMT1A

<sup>1</sup>College of Animal Science and Technology, Northwest A&F University, Shaanxi Key Laboratory of Molecular Biology for Agriculture, Yangling, Shaanxi, 712100, China. <sup>2</sup>Department of Forensic Medicine, Xinxiang Medical University, Xinxiang, Henan, 453003, China. <sup>3</sup>Library, Xinxiang Medical University, Xinxiang, Henan, 453003, China. Correspondence and requests for materials should be addressed to H.C. (email: chenhong1212@263.net)

Breeds	Sample size	CNV type		
		Gain	Normal	Loss
Jiaxian	246	40 (16.26%)	187 (76.02%)	19 (7.72%)
Qinchuan	187	34 (18.18%)	133 (71.12%)	20 (10.70%)
Nanyang	104	32 (30.77%)	61 (58.65%)	11 (10.58%)
Jinnan	102	4 (3.92%)	77 (75.49%)	21 (20.59%)
Luxi	83	19 (22.89%)	44 (53.01%)	20 (24.10%)
Chinese Red Steppe	66	19 (28.79%)	41 (62.12%)	6 (9.09%)

**Table 1.** Distribution of the *CYP4A11* gene CNV types among breeds.

Breeds	Growth traits	CNV Type		
		Gain (Mean $\pm$ SE)	Normal (Mean $\pm$ SE)	Loss (Mean $\pm$ SE)
Jiaxian	Adult HW (cm)	26.76 <sup>a</sup> $\pm$ 0.54	26.08 <sup>ab</sup> $\pm$ 0.35	24.30 <sup>b</sup> $\pm$ 1.05
Qinchuan	Adult HG (cm)	182.39 <sup>a</sup> $\pm$ 1.83	177.67 <sup>ab</sup> $\pm$ 7.61	174.90 <sup>b</sup> $\pm$ 2.95
	Adult CD (cm)	65.06 <sup>a</sup> $\pm$ 0.97	62.33 <sup>ab</sup> $\pm$ 4.06	61.05 <sup>b</sup> $\pm$ 1.57
Nanyang	12 months BL (cm)	119.14 <sup>a</sup> $\pm$ 2.59	116.55 <sup>ab</sup> $\pm$ 1.11	111.17 <sup>b</sup> $\pm$ 2.80
Jinnan	Adult BW (kg)	489.02 <sup>a</sup> $\pm$ 10.82	488.25 <sup>a</sup> $\pm$ 45.91	437.15 <sup>b</sup> $\pm$ 21.07
	Adult HG (cm)	186.01 <sup>a</sup> $\pm$ 1.48	183.25 <sup>ab</sup> $\pm$ 6.29	179.53 <sup>b</sup> $\pm$ 2.89
Chinese Red Steppe	6 months HG (cm)	133.92 <sup>a</sup> $\pm$ 2.81	133.62 <sup>a</sup> $\pm$ 1.47	127.70 <sup>b</sup> $\pm$ 1.84
	6months CG (cm)	12.58 <sup>a</sup> $\pm$ 0.15	12.47 <sup>a</sup> $\pm$ 0.12	11.83 <sup>b</sup> $\pm$ 0.23

**Table 2.** Significant associations of *CYP4A11* CNV types with growth traits. Values with different superscripts within the same line differ significantly at  $P < 0.05$ . <sup>a,b</sup>SE: standard error. HW, Hucklebone width; HG, heart girth; CD, Chest depth; BL, body length; BW, body weight; CG, Cannon girth.

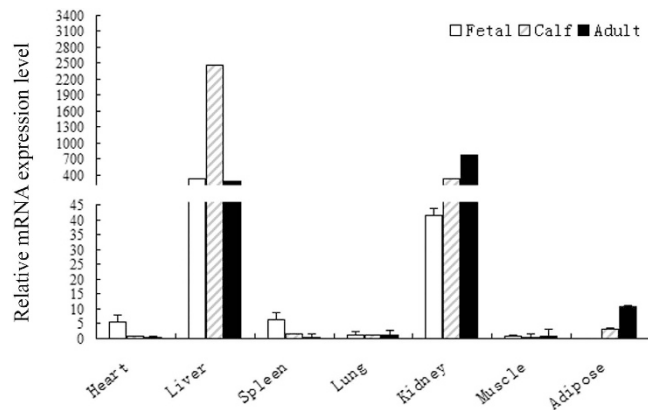
disease phenotype, and deletion of this region can result in insufficient levels of *PMP22* and the HNPP (hereditary neuropathy with liability to pressure palsies)<sup>27</sup>.

Jiaxian (JX), Qinchuan (QC), Nanyang (NY), Jinnan (JN), and Luxi (LX) cattle are five of the eight representative native bovine (*Bos taurus*) breeds in China. They are work and beef dual-purpose breeds and designated as national protection resources due to their excellent performance traits and good meat production characteristics. Chinese Red Steppe (CRS) is a dairy and beef type generated by the cross of shorthorn bulls with Mongolian cows. Some artificial selection and breeding from draft usage to beef production have been applied to this breed for a relatively long time, which might have resulted in different variation. Among isoforms in the *CYP4A* sub-family, cytochrome P-450 4A11 (*CYP4A11*) is a major lauric acid (medium-chain fatty acid) omega hydroxylase in human liver and kidney<sup>28–30</sup>. *CYP4A11* functions to convert arachidonic acid to 20-hydroxyecosatetraenoic acid (20-HETE), and is involved in fatty-acid metabolism and blood pressure regulation<sup>30,31</sup>. However, there have been no reports about the function of the *CYP4A11* gene in cattle. We previously confirmed that the cattle *CYP4A11* gene was located within CNVR44 that encompasses QTLs that alter economic traits like meat quality and milk production<sup>21</sup>. This location suggests that the copy number of *CYP4A11* must be variable, which we hypothesized may have genetic effects that can be conveyed to phenotypes. This study was performed to confirm the widespread existence of *CYP4A11* CNV in different Chinese cattle breeds, determine its relationship with growth, and explore the biological effects of its expression.

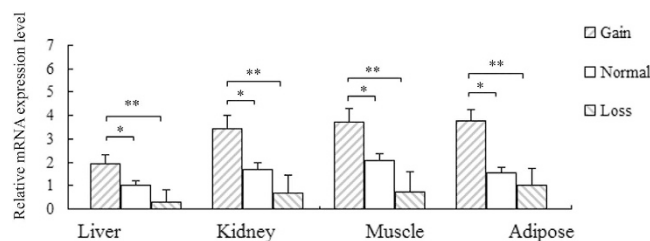
## Results

**Distribution of *CYP4A11* CNV in populations.** To validate the presence and investigate the distribution of *CYP4A11* CNV, the relative gene copy number was detected and estimated in six Chinese cattle breeds (Supplementary Table S1). As shown in Table 1, there was a similar distribution of CNV types in JX, QC, NY, JN, LX and CRS populations, with Normal more common than Gain or Loss. Gain represents more than two gene copies, Loss means less than two DNA copies, and Normal means two gene copies. Additionally, there were differences in the distribution frequency among breeds (Table 1).

**Association between *CYP4A11* CNV and growth.** Association analysis between *CYP4A11* CNV and growth were carried out for samples from JX, QC, NY, JN, LX, and CRS by one-way analysis of variance (ANOVA). As shown in Table 2, the *CYP4A11* CNV was significantly associated with some growth traits. Specifically, in JX, the adult hucklebone width of Gain type was higher than that of Loss type ( $P < 0.05$ ), while in QC, adult heart girth and chest depth were higher for Gain type than the Loss type ( $P < 0.05$ ); very similarly, NY showed a higher 12-month body length for Gain type than Loss type ( $P < 0.05$ ). Significant effects were also found in JN adults, body weight and heart girth were higher in Gain individuals than in Loss ones ( $P < 0.05$ ). Also, the



**Figure 1.** The expression profiling of *CYP4A11* gene at different stages in Qinchuan. Data are presented as means  $\pm$  standard error (SE,  $n = 3$ ). Error bars represent standard errors from three independent experiments.



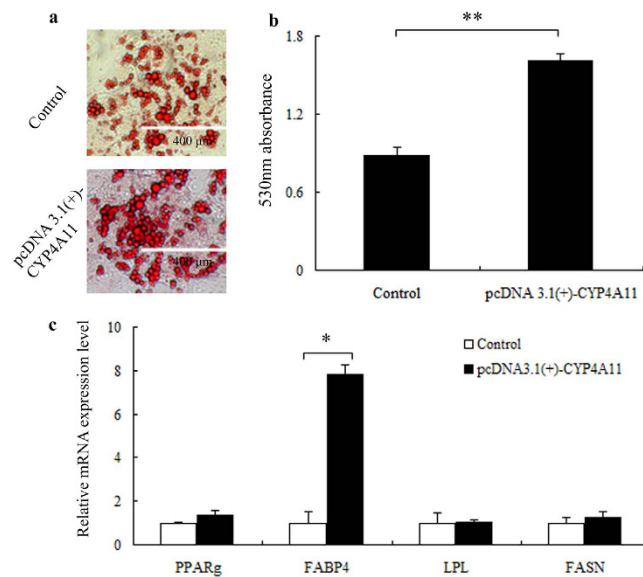
**Figure 2.** Influence of *CYP4A11* CNV type on gene expression in four types of tissues. Data are presented as means  $\pm$  standard error (SE,  $n = 3$ ). Error bars represent standard errors from three independent experiments. Asterisks (\*) and (\*\*) indicate statistically significant differences ( $P < 0.05$ ) and ( $P < 0.01$ ), respectively.

cannon girth and heart girth in CRS at 6 months were higher in Gain or Normal than in Loss type ( $P < 0.05$ ). Except for the aforementioned results, the other traits were nonsignificantly correlated with *CYP4A11* CNV type but showed a tendency that individuals with Gain or Normal types possessed higher growth values than Loss type ( $P > 0.05$ ), including LX breed (Supplementary Tables S2, S3, S4, S5, S6 and S7).

**Expression differences of *CYP4A11* among different tissues.** Before analyzing the effect of *CYP4A11* copy number on mRNA level, gene expression profiling was necessary, as the expression pattern of cattle *CYP4A11* had not been examined previously. We determined the expression of *CYP4A11* in seven different tissues in QC from three growth stages (fetal, calf, and adult). The results showed that the *CYP4A11* expression level varied (Fig. 1). In fetal and calf samples, *CYP4A11* mRNA was expressed in all tissues, with a high level in the liver, moderate level in kidney, and low level in the heart, spleen, lung, muscle, and adipose samples. In adults, the mRNA level was the highest in kidney, followed by liver, and the other tissues showed low levels. Additionally, compared with calf, the level of *CYP4A11* expression in the adipose tissue was increased in the adult samples.

**Influence of *CYP4A11* CNV type on mRNA expression.** Based on the expression profiling, we next analyzed the differences of *CYP4A11* expression level for the different CNV types. As shown in Supplementary Fig. S1, 30 individuals were tested. Among the tested individuals, 21 were of Normal type, 6 were Gain type, and 3 were Loss type. Supplementary Fig. S2 shows the expression in liver, kidney, muscle, and adipose tissues. The CNV type-expression analysis is shown in Fig. 2, which showed that the mRNA level of Gain was significantly higher than that of Normal or Loss type ( $P < 0.01$  or  $P < 0.05$ ). No significant difference was found between Normal and Loss, but the mRNA level in Normal was somewhat higher than in Loss. These results indicated that more *CYP4A11* copies increased the gene expression level, and fewer copies decreased gene expression.

**Promotion of overexpressed *CYP4A11* on lipogenesis.** Expression profiling revealed that there was low expression of the *CYP4A11* gene in adipose tissue. To examine the role of *CYP4A11* in fatty-acid metabolism, the protein was overexpressed from a plasmid in 3T3L1 cells. Before differentiation, the mRNA level of *CYP4A11* gene after overexpression was found to be significantly higher than in the control group ( $P = 0.008$ , Supplementary Fig. S3). The overexpressing cells were exposed to differentiation medium. At the 12<sup>th</sup> day of induction, Oil Red O (ORO) staining was performed and showed that in the overexpressing cells, there were both larger fat droplets and more droplets than in the cells of the control group (Fig. 3a). The absorbance at 530 nm for the pcDNA 3.1(+)-*CYP4A11* group was 1.606, significantly higher than that of the control group ( $P = 0.001$ , Fig. 3b). Furthermore, the mRNA level of marker genes important in adipose differentiation, such as



**Figure 3. The overexpressed *CYP4A11* in 3T3L1 cells promoted adipogenesis.** (a) Lipid staining with Oil Red O (ORO). The “Control” indicates the 3T3L1 cell group transfected with pcDNA 3.1(+) plasmids. (b) Quantification of ORO staining. Data are presented as means  $\pm$  standard error (SE,  $n = 3$ ). Error bars represent standard errors from three independent experiments. “Control” represents the 3T3L1 cell group transfected with pcDNA 3.1(+) plasmids. Asterisks (\*\*) indicate statistically significant differences ( $P < 0.01$ ). (c) Comparison of candidate genes expression. Data are presented as means  $\pm$  standard error (SE,  $n = 3$ ). Error bars represent standard errors from three independent experiments. “Control” represents the 3T3L1 cell group transfected with pcDNA 3.1(+) plasmids. Asterisks (\*) indicate statistically significant differences ( $P < 0.05$ ).

peroxisome proliferator-activated receptor gamma (*PPARG*), fatty acid binding protein 4 (*FABP4*), Lipoprotein lipase (*LPL*), and Fatty acid synthase (*FASN*) were 1.34, 7.81, 1.02, and 1.28-fold higher in the overexpression group, respectively, than in the control group, with a significant difference in *FABP4* levels between the two groups ( $P = 0.047$ , Fig. 3c). These results, together with the previous discovery that additional *CYP4A11* copies increased gene expression level, indicated that *CYP4A11* copy numbers affect preadipocyte differentiation. Based on these results, it could be inferred that the increase in *CYP4A11* copies may promote lipogenesis in cattle.

## Discussion

In livestock such as cattle, most CNV studies have been limited to CNV discovery and enumeration using various platforms, including aCGH, SNP array, and next generation sequencing<sup>20,32–34</sup>. Even though these studies have identified many copy number variable regions in different species, there have been few attempts to associate copy number with expression effects. The investigation of the diversity and origin of CNVs, the characterization of their genetic properties, and the determination of the functional impacts of CNVs are promising active areas of research<sup>35</sup>.

The quantitative real-time PCR (qPCR) assay used here was a highly effective method of CNV detection<sup>33,36</sup>. Through qPCR analysis, the *CYP4A11* CNV types showed similar distributions in JX, QC, NY, JN, LX, and CRS populations. The Normal type was more common than Gain or Loss types. This indicated the genetic similarity of these breeds after long-term artificial selection and breeding efforts. The six breeds showed differences in CNV frequency that may be caused by the diversity in their genetic backgrounds<sup>37</sup>. JX, QC, NY, JN, and LX cattle are five native Chinese breeds, and CRS is a dairy and beef breed generated by crossing shorthorn bulls with Mongolian cows.

Association analysis revealed that the *CYP4A11* CNVs were significantly associated with some growth traits, however, most of the studied phenotypes showed no significant differences among CNV types, just a tendency of higher growth for Gain or Normal individuals compared to Loss individuals. This reflected the positive effect of multiple copies of *CYP4A11* on growth, which may be attributed to the location of *CYP4A11* in the CNVR44 which includes QTLs that are linked to economic traits<sup>21</sup>. As previous reports showed, some genes were located within CNVRs that directly overlapped with QTLs for body weight, marbling score, and somatic cell count<sup>38</sup>, or were partly encompassed by QTLs for body weight, carcass weight, and breast muscle weight<sup>39,40</sup>. *PLA2G2D*, a gene related to milk production and meat quality, was reported to have multiple copies in beef cattle<sup>5</sup>. Consistently, *PLA2G2D* was found to be located in CNVR 22, whose number of repeats can significantly reduce the mRNA levels of *PLA2G2D* and correspondingly cause growth retardation in Chinese cattle<sup>21</sup>. Yue *et al.*<sup>41,42</sup> studied Holstein bulls and found that the *PRAMEY* copy number is negatively correlated with scrotal circumference, relative scrotal circumference, percentage of normal sperm, and nonreturn rate, and the CNVs of both *HSPY* and *ZNF280BY* are negatively correlated with testis size and positively correlated with sire conception rate. This suggests that some phenotypic differences could be attributed to genes encompassed by CNVRs.

CNVs mainly influence phenotypes in three ways: gene dosage, expression regulation changes, and recessive allele exposure. Of these, gene dosage is the major contributor to phenotypes associated with CNVs<sup>43</sup>. Functional genes that are partly encompassed by CNVRs could affect gene transcription or expression to a certain degree<sup>44,45</sup>, but genes entirely contained by CNVRs may show a positive relationship with their transcription and expression activity, because the multiple copies could all be actively transcribed and expressed<sup>46</sup>. For instance, the duplication of the *CCL3L1* gene significantly increased its expression level<sup>47</sup>. Similarly, the increase in *CYP4A11* copy numbers increased its expression level, suggesting that the *CYP4A11* CNV may affect cattle phenotypes by altering gene dosage.

To test whether *CYP4A11* CNV functions by gene dosage effects, the *CYP4A11* gene was overexpressed in 3T3-L1 cells. The *CYP4A11* overexpression dramatically increased the formation of fat droplets. At the same time, the expression level of genes related to adipose and lipid accumulation, *PPARG*, *FABP4*, *LPL*, and *FASN*, were all increased. *PPARG* is a positive regulator of adipose accretion<sup>48,49</sup>. The expression and activity of *LPL* is positively correlated with adipose tissue mass<sup>50</sup>, like *FASN*, which plays a critical role in determining the fatty acid profile<sup>51</sup>. The differences of *PPARG*, *LPL*, and *FASN* levels between the two groups were not significant. The lack of a dramatic effect may be caused by the strikingly higher *FABP4* level in the pcDNA 3.1(+)-*CYP4A11* group, which increased significantly following lipid accumulation to increase fatty acid degradation, and subsequently reduced the expression levels of *PPARG*, *LPL*, and *FASN*<sup>52</sup>. These results, together with the finding that the increase in *CYP4A11* copy numbers could increase its expression level, suggest that the *CYP4A11* multicopy may promote 3T3-L1 cell differentiation from preadipocytes to adipocytes and may increase adipose deposits in cattle.

The increase in *CYP4A11* copy numbers promoted cattle growth and increased the expression level, confirming the positive gene dosage effect by which the *CYP4A11* CNV could affect phenotype. This information may be used in future molecular marker-assisted selection for cattle breeding. The *in vitro* overexpression revealed that *CYP4A11* duplication could increase adipose deposits. Future *in vivo* study will provide more insights into the effects of CNV. Taken together, these studies revealed the functional role of the *CYP4A11* CNV and provided the basis for its further exploitation in cattle breeding.

## Methods

**Ethics statement.** The experiments and animal care were performed according to the Regulations for the Administration of Affairs Concerning Experimental Animals (Ministry of Science and Technology, China, 2004) and approved by the Institutional Animal Care and Use Committee (College of Animal Science and Technology, Northwest A&F University, China). All the experiments and methods were performed in accordance with the approved guidelines. Cattle were permitted access to feed and water at will and all efforts were made to minimize suffering.

**Animals and samples.** Blood samples of JX (N = 246), QC (N = 187), NY (N = 104), JN (N = 102), LX (N = 83), and CRS (N = 66) were respectively collected from Henan, Shannxi, Henan, Shanxi, Shandong, and Jilin provinces. Growth data, including body weight, average daily gain, withers height, body length, heart girth, hucklebone width, hip width, rump length, chest depth, chest breadth, height of hip cross, cannon girth, and abdominal girth were obtained. Tissue samples of fetal (N = 3), calf (N = 3), and adult (N = 3) of QC, including heart, liver, spleen, lung, kidney, muscle and adipose (no adipose for fetal) were collected in a slaughter house. In addition, liver, kidney, muscle, and adipose tissues were obtained from QC adults (N = 30).

**DNA and RNA isolation.** Genomic DNA was extracted from blood samples and 30 QC adult livers following standard procedures<sup>53</sup>. Total RNA was isolated from tissues and 3T3-L1 cells by the Trizol method (TaKaRa, Otsu, Shiga, Japan) and reverse-transcribed to get the first-strand cDNA using the Prime Script RT Reagent Kit (TaKaRa, Dalian, China) using 500 ng of total RNA as template and oligo-dT primers. Both the genomic DNA and cDNA were diluted to 20 ng/ $\mu$ L.

**Plasmid construction.** The *CYP4A11* coding region was amplified using cDNA as template and inserted into the pcDNA 3.1(+) vector (Invitrogen, Carlsbad, CA). *NheI* and *HindIII* restriction sites were included in the forward and reverse primers to amplify the gene. The primers were as follows, sense: 5'-CTAGCTAGCATGAGTGTCTCTGCACTGAGCCCC-3', and antisense: 5'-CCCAAGCTTCTAAGTCCCTGGATCAGAGAGCTTCCT-3'. The 1563 bp CDS fragments were purified using TIANquick Midi Purification Kit (Tiangen, Beijing, China), then digested with *NheI* and *HindIII*. Then, the ligation was carried out overnight at 16 °C and transformed into Top10 *E. coli* cells (Tiangen, Beijing, China). Finally, recombinant plasmid was isolated and confirmed by double digestion and DNA sequencing.

**Cell culture and transfection.** 3T3-L1 cells (purchased from cellbank.org.cn) were plated into 12-well cell culture dishes and cultured in growth medium comprised of Dulbecco's Modified Eagle's Medium (DMEM) (Hyclone, Logan, UT) supplemented with 10% fetal bovine serum (FBS) (HyClone) and 1  $\times$  antibiotics (100 U/ml penicillin, 100  $\mu$ g/ml streptomycin; Key-GEN, Nanjing, China) and incubated at 37 °C with 5% CO<sub>2</sub>.

pcDNA3.1(+)-*CYP4A11* and pcDNA3.1(+) vector (control group) were separately transfected into confluent 3T3-L1 cells using Lipofectamine 2000 transfection reagent (Invitrogen, Carlsbad, CA). First, plasmid DNA and Lipofectamine 2000 were diluted in DMEM medium separately at a ratio of 1:2.5. After incubation for 5 minutes at room temperature, the diluted DNA and Lipofectamine 2000 were mixed together and incubated for 20 minutes at room temperature. This complex was then added to a culture dish containing cells and growth medium.

**Cell differentiation and lipid staining.** Cell differentiation was induced by addition of differentiation medium. Briefly, six wells of cells were transferred to differentiation medium I (growth medium plus 0.5 mmol/L 3-Isobutyl-1-methylxanthine (IBMX), 1  $\mu$ mol/L dexamethasone (Dex, Dex medium has a weak ability to induce

differentiation of 3T3-L1 cells into adipocytes), and 10  $\mu\text{g}/\text{mL}$  Insulin (Promega, Heidelberg, Germany) 24 hours after transfection. Then, about 72 hours later, cells were rinsed twice with phosphate buffered saline (PBS), and differentiation medium I was replaced with differentiation medium II (growth medium plus 10  $\mu\text{g}/\text{mL}$  Insulin). Differentiation medium II was renewed every 48 hours until the 12<sup>th</sup> day when the induction finished. The control group was treated as the same.

Lipid accumulation in adipocytes was visualized by staining with ORO (TaKaRa, Dalian, China). First, three wells of cells were rinsed twice with phosphate buffered saline (PBS) and fixed with 10% formaldehyde for 1 hour at room temperature. Then, after washing twice again with PBS, the cells were stained for at least 20 minutes in freshly diluted ORO solution (0.3% (w/v) ORO in 60% isopropanol). The staining solution was then removed, and cells were washed twice with PBS. ORO was extracted with isopropanol and the signal was quantified spectrophotometrically (NanoPhotometer, Implen Inc., CA, USA) at 530 nm<sup>54</sup>.

**qPCR.** To evaluate the CNV type, copy number was estimated by DNA qPCR with the CFX 96TM Real Time Detection System (Bio-Rad, Hercules, CA, USA), using the SYBR Premix Ex Taq II kit (TaKaRa, Dalian, China). Genomic DNA was extracted from blood samples following standard procedures<sup>53</sup>, and was diluted to 20 ng/ $\mu\text{L}$ . The primers were designed using Beacon Designer 7 software (Supplementary Table S8), based on the DNA sequence of the *Bos taurus CYP4A11* gene (GenBank accession no. NC\_007301.5), with *BTF3* as an internal reference gene<sup>38</sup>. Positive and negative controls were generated, respectively with the genomic DNA of Angus (a beef breed previously used as a reference because it has two DNA copies) and water as templates. Each reaction was performed in triplicate. The 25  $\mu\text{L}$  qPCR amplification volume contained 20 ng genomic DNA, 1  $\mu\text{L}$  of each primer (10  $\mu\text{mol}/\text{L}$ ), and 12.5  $\mu\text{L}$  SYBR Premix Ex TaqII. The PCR protocol was as follows: 30 s at 95 °C, followed by 40 cycles of 95 °C for 10 s, 60 °C for 20 s, and 68 °C for 20 s.

The differences in *CYP4A11* gene expression in the different tissues were detected by cDNA qPCR. Total RNA was isolated from heart, liver, spleen, lung, kidney, muscle and adipose tissues by the Trizol method (TaKaRa, Otsu, Shiga, Japan). The quantity and quality of the RNA were determined using a NanoPhotometer spectrometer (Implen Inc, CA, USA). Next, the RNA was reverse-transcribed to obtain the first-strand cDNA using the Prime Script RT Reagent Kit (TaKaRa, Dalian, China) using 500 ng of total RNA as template and oligo-dT primers. After synthesis, the cDNA template was diluted to 20 ng/ $\mu\text{L}$ . As a housekeeping gene, glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) was selected for the endogenous standard<sup>55</sup>. The specific primers for *CYP4A11* and *GAPDH* gene are listed in Supplementary Table S8. The qPCR amplification volume and protocol was performed the same as the genomic DNA qPCR. Every reaction was performed in triplicate.

To evaluate the influences of *CYP4A11* copy numbers on expression level, genomic DNA and total RNA samples were isolated from 30 QC individuals and qPCR was performed. Genomic DNA was isolated from livers<sup>53</sup> and total RNA was isolated from liver, kidney, muscle, and adipose tissues and reversed-transcribed. The genomic DNA qPCR and cDNA qPCR amplification were respectively performed following methods described above. Every reaction was performed in triplicate.

To assess the effects of multiple *CYP4A11* copies on adipogenesis, the expression levels of *CYP4A11*, *PPARG*, *FABP4*, *LPL*, and *FASN* in 3T3-L1 cells were determined for the two groups (pcDNA3.1(+)-*CYP4A11* group and control group) by qPCR. After transfection, RNA was isolated from three wells of cells and the synthesized cDNA was used to detect the *CYP4A11* expression level. After differentiation, three wells of cells were gathered for RNA isolation and subsequent expression level detection of *PPARG*, *FABP4*, *LPL*, and *FASN* genes. The primer sequences for the tested genes and the internal control gene *GAPDH* are listed in Supplementary Table S9. RNA isolation, cDNA synthesis, and qPCR amplification were performed as described above. Every reaction was performed in triplicate.

**CNV type assay.** The qPCR analysis was performed based on the  $\Delta\Delta\text{Ct}$  method<sup>56</sup>, and  $\Delta\Delta\text{Ct}$  values were determined by comparing tested samples and the Angus reference, using the *BTF3* gene as an internal control. Finally, the relative copy number for each sample was calculated as  $2^{-\Delta\Delta\text{Ct}}$ . By comparing the  $2^{-\Delta\Delta\text{Ct}}$  value of the tested individuals with the value for Angus, the CNV type of the *CYP4A11* gene was defined as Gain (more than two gene copies), Loss (less than two gene copies), or Normal (two gene copies, same as the Angus reference).

**Association analysis.** Association analysis between the CNV types of *CYP4A11* and growth traits was performed using the General Linear Model (GLM) in SPSS (version 16.0).

The linear model was:  $Y_{ijk} = \mu + A_i + T_j + e_{ijk}$ , where  $Y_{ijk}$  is the phenotypic value of the  $ijk$ th individual,  $\mu$  is the overall population mean,  $A_i$  is the effect of the  $i$ th age,  $T_j$  is the effect associated with the  $j$ th CNV type, and  $e_{ijk}$  is the random error.

**Expression analysis.** The relative expression levels of *CYP4A11*, *PPARG*, *FABP4*, *LPL*, and *FASN* were calculated using Gene Expression Macro software (Applied Biosystems, Life Technologies, Carlsbad, CA, USA) by normalizing to the expression of *GAPDH* with the  $\Delta\Delta\text{Ct}$  method<sup>56</sup>.

The differences of mRNA expression level among three *CYP4A11* CNV types were estimated by ANOVA. The results of the multiple comparisons were corrected by Bonferroni correction, and the differences were considered significant if  $P < 0.05$ .

The differences of expression levels in *CYP4A11*, *PPARG*, *FABP4*, *LPL*, and *FASN* genes between pcDNA3.1(+)-*CYP4A11* and the control group were evaluated by the Student's t-test (two-tail). The differences were considered significant if  $P < 0.05$ .

## References

- Elsik, C. G. *et al.* The Genome Sequence of Taurine Cattle: A Window to Ruminant Biology and Evolution. *Science* **324**, 522–528, doi: 10.1126/science.1169588 (2009).
- Zimin, A. V. *et al.* A whole-genome assembly of the domestic cow, *Bos taurus*. *Genome Biology* **10**, doi: 10.1186/gb-2009-10-4-r42 (2009).
- Eck, S. H. *et al.* Whole genome sequencing of a single *Bos taurus* animal for single nucleotide polymorphism discovery. *Genome Biology* **10**, doi: 10.1186/gb-2009-10-8-r82 (2009).
- Matukumalli, L. K. *et al.* Development and Characterization of a High Density SNP Genotyping Assay for Cattle. *PLoS One* **4**, doi: 10.1371/journal.pone.0005350 (2009).
- Stothard, P. *et al.* Whole genome resequencing of Black Angus and Holstein cattle for SNP and CNV discovery. *BMC genomics* **12**, doi: 10.1186/1471-2164-12-559 (2011).
- Barendse, W., Harrison, B. E., Bunch, R. J., Thomas, M. B. & Turner, L. B. Genome wide signatures of positive selection: The comparison of independent samples and the identification of regions associated to traits. *BMC genomics* **10**, doi: 10.1186/1471-2164-10-178 (2009).
- Hayes, B. J. *et al.* A genome map of divergent artificial selection between *Bos taurus* dairy cattle and *Bos taurus* beef cattle. *Animal Genetics* **40**, 176–184, doi: 10.1111/j.1365-2052.2008.01815.x (2009).
- Cole, J. B. *et al.* Distribution and location of genetic effects for dairy traits. *Journal Of Dairy Science* **92**, 2931–2946, doi: 10.3168/jds.2009-92-7-3542 (2009).
- Flori, L. *et al.* The Genome Response to Artificial Selection: A Case Study in Dairy Cattle. *PLoS One* **4**, doi: 10.1371/journal.pone.0006595 (2009).
- Iafraite, A. J. *et al.* Detection of large-scale variation in the human genome. *Nature Genetics* **36**, 949–951, doi: 10.1038/ng1416 (2004).
- Sebat, J. *et al.* Large-scale copy number polymorphism in the human genome. *Science* **305**, 525–528, doi: 10.1126/science.1098918 (2004).
- Mills, R. E. *et al.* Mapping copy number variation by population-scale genome sequencing. *Nature* **470**, 59–65, doi: 10.1038/nature09708 (2011).
- Altshuler, D. M. *et al.* Integrating common and rare genetic variation in diverse human populations. *Nature* **467**, 52–58, doi: 10.1038/nature09298 (2010).
- Guryev, V. *et al.* Distribution and functional impact of DNA copy number variation in the rat. *Nature Genetics* **40**, 538–545, doi: 10.1038/ng.141 (2008).
- Yalcin, B. *et al.* Sequence-based characterization of structural variation in the mouse genome. *Nature* **477**, 326–329, doi: 10.1038/nature10432 (2011).
- Wang, J. Y. *et al.* Identification of Genome-Wide Copy Number Variations among Diverse Pig Breeds Using SNP Genotyping Arrays. *PLoS One* **8**, doi: 10.1371/journal.pone.0068683 (2013).
- Wang, J. Y. *et al.* Enhancing Genome-Wide Copy Number Variation Identification by High Density Array CGH Using Diverse Resources of Pig Breeds. *PLoS One* **9**, doi: 10.1371/journal.pone.0087571 (2014).
- Doan, R. *et al.* Identification of copy number variants in horses. *Genome Research* **22**, 899–907, doi: 10.1101/gr.128991.111 (2012).
- Metzger, J. *et al.* Analysis of copy number variants by three detection algorithms and their association with body size in horses. *BMC genomics* **14**, doi: 10.1186/1471-2164-14-487 (2013).
- Jiang, L. *et al.* Genome-Wide Identification of Copy Number Variations in Chinese Holstein. *PLoS One* **7**, doi: 10.1371/journal.pone.0048732 (2012).
- Zhang, L. Z. *et al.* Detection of copy number variations and their effects in Chinese bulls. *BMC genomics* **15**, doi: 10.1186/1471-2164-15-480 (2014).
- Fontanesi, L. *et al.* An initial comparative map of copy number variations in the goat (*Capra hircus*) genome. *BMC genomics* **11**, doi: 10.1186/1471-2164-11-639 (2010).
- Wang, X. F., Nahashon, S., Feaster, T. K., Bohannon-Stewart, A. & Adefope, N. An initial map of chromosomal segmental copy number variations in the chicken. *BMC genomics* **11**, doi: 10.1186/1471-2164-11-351 (2010).
- Zhang, F., Gu, W. L., Hurles, M. E. & Lupski, J. R. In *Annual Review Of Genomics And Human Genetics Vol. 10 Annual Review Of Genomics and Human Genetics* 451–481 (2009).
- Spielmann, M. & Klopocki, E. CNVs of noncoding cis-regulatory elements in human disease. *Current Opinion In Genetics & Development* **23**, 249–256, doi: 10.1016/j.gde.2013.02.613 (2013).
- Boone, P. M. *et al.* Deletions of recessive disease genes: CNV contribution to carrier states and disease-causing alleles. *Genome Research* **23**, 1383–1394, doi: 10.1101/gr.156075.113 (2013).
- Lupski, J. R. & Chance, P. F. *Hereditary Motor and Sensory Neuropathies Involving Altered Dosage or Mutation of PMP22: The CMT1A Duplication and HNPP Deletion* (2005).
- Adas, F. *et al.* Requirement for omega and (omega-1)-hydroxylations of fatty acids by human cytochromes P450 2E1 and 4A11. *Journal Of Lipid Research* **40**, 1990–1997 (1999).
- Imaoka, S., Ogawa, H., Kimura, S. & Gonzalez, F. J. Complete cDNA sequence and cDNA-directed expression of CYP4A11, a fatty acid omega-hydroxylase expressed in human kidney. *Dna & Cell Biology* **12**, 893–899 (1994).
- Powell, P. K., Wolf, I. & Lasker, J. M. Identification of CYP4A11 as the major lauric acid omega-hydroxylase in human liver microsomes. *Archives Of Biochemistry And Biophysics* **335**, 219–226, doi: 10.1006/abbi.1996.0501 (1996).
- Lasker, J. M. *et al.* Formation of 20-hydroxyeicosatetraenoic acid, a vasoactive and natriuretic eicosanoid, in human kidney - Role of CYP4F2 and CYP4A11. *Journal Of Biological Chemistry* **275**, 4118–4126, doi: 10.1074/jbc.275.6.4118 (2000).
- Hou, Y. L. *et al.* Fine mapping of copy number variations on two cattle genome assemblies using high density SNP array. *BMC genomics* **13**, doi: 10.1186/1471-2164-13-376 (2012).
- Bickhart, D. M. *et al.* Copy number variation of individual cattle genomes using next-generation sequencing. *Genome Research* **22**, 778–790, doi: 10.1101/gr.133967.111 (2012).
- Ciconardi, F. *et al.* Massive screening of copy number population-scale variation in *Bos taurus* genome. *BMC genomics* **14**, doi: 10.1186/1471-2164-14-124 (2013).
- Xu, L. *et al.* Population-genetic properties of differentiated copy number variations in cattle. *Scientific Reports* **6** (2016).
- Ali, S. *et al.* IL12B SNPs and copy number variation in IL23R gene associated with susceptibility to leprosy. *Journal Of Medical Genetics* **50**, 34–42, doi: 10.1136/jmedgenet-2012-101214 (2013).
- Lehnert, S. A. *et al.* Gene expression studies of developing bovine longissimus muscle from two different beef cattle breeds. *Bmc Developmental Biology* **7**, doi: 10.1186/1471-213x-7-95 (2007).
- Liu, G. E. *et al.* Analysis of copy number variations among diverse cattle breeds. *Genome Research* **20**, 693–703, doi: 10.1101/gr.105403.110 (2010).
- Wang, J. Y. *et al.* A genome-wide detection of copy number variations using SNP genotyping arrays in swine. *BMC genomics* **13**, doi: 10.1186/1471-2164-13-273 (2012).
- Wang, Y. *et al.* A genome-wide survey of copy number variation regions in various chicken breeds by array comparative genomic hybridization method. *Animal Genetics* **43**, 282–289, doi: 10.1111/j.1365-2052.2011.02308.x (2012).
- Yue, X. P. *et al.* Copy number variation of PRAMEY across breeds and its association with male fertility in Holstein sires. *Journal Of Dairy Science* **96**, 8024–8034, doi: 10.3168/jds.2013-7037 (2013).

42. Yue, X. P. *et al.* Copy number variations of the extensively amplified Y-linked genes, HSFY and ZNF280BY, in cattle and their association with male reproductive traits in Holstein bulls. *BMC genomics* **15**, doi: 10.1186/1471-2164-15-113 (2014).
43. Bickhart, D. M. & Liu, G. E. The challenges and importance of structural variation detection in livestock. *Frontiers in genetics* **5**, 37, doi: 10.3389/fgene.2014.00037 (2014).
44. Aldred, P. M. R., Hollox, E. J. & Armour, J. A. L. Copy number polymorphism and expression level variation of the human alpha-defensin genes DEFA1 and DEFA3. *Human Molecular Genetics* **14**, 2045–2052, doi: 10.1093/hmg/ddi209 (2005).
45. Perry, G. H. *et al.* Diet and the evolution of human amylase gene copy number variation. *Nature Genetics* **39**, 1256–1260, doi: 10.1038/ng2123 (2007).
46. Wong, K. K. *et al.* A comprehensive analysis of common copy-number variations in the human genome. *American Journal Of Human Genetics* **80**, 91–104, doi: 10.1086/510560 (2007).
47. Lee, H., Bae, S., Choi, B. W., Choi, J. C. & Yoon, Y. Copy number variation of CCL3L1 influences asthma risk by modulating IL-10 expression. *Clinica Chimica Acta* **412**, 2100–2104, doi: 10.1016/j.cca.2011.07.017 (2011).
48. Wang, Y. X. PPARs: diverse regulators in energy metabolism and metabolic diseases. *Cell Research* **20**, 124–137, doi: 10.1038/cr.2010.13 (2010).
49. Lehrke, M. & Lazar, M. A. The many faces of PPAR gamma. *Cell* **123**, 993–999, doi: 10.1016/j.cell.2005.11.026 (2005).
50. Xu, F. *et al.* Lack of SIRT1 (Mammalian Sirtuin 1) Activity Leads to Liver Steatosis in the SIRT1(+/-) Mice: A Role of Lipid Mobilization and Inflammation. *Endocrinology* **151**, 2504–2514, doi: 10.1210/en.2009-1013 (2010).
51. de Souza, F. R. P. *et al.* Associations of FASN gene polymorphisms with economical traits in Nellore cattle (*Bos primigenius indicus*). *Molecular Biology Reports* **39**, 10097–10104, doi: 10.1007/s11033-012-1883-6 (2012).
52. Garin-Shkolnik, T., Rudich, A., Hotamisligil, G. S. & Rubinstein, M. FABP4 Attenuates PPAR gamma and Adipogenesis and Is Inversely Correlated With PPARg in Adipose Tissues. *Diabetes* **63**, 900–911 (2014).
53. Müllenbach, R., Lagoda, P. J. & Welter, C. An efficient salt-chloroform extraction of DNA from blood and tissues. *Trends in Genetics* **5**, 391–391 (1989).
54. Okamura, M. *et al.* COUP-TFII acts downstream of Wnt/beta-catenin signal to silence PPAR gamma gene expression and repress adipogenesis. *Proceedings Of the National Academy Of Sciences Of the United States Of America* **106**, 5819–5824, doi: 10.1073/pnas.0901676106 (2009).
55. Overbergh, L., Valckx, D., Waer, M. & Mathieu, C. Quantification of murine cytokine mRNAs using real time quantitative reverse transcriptase PCR. *Cytokine* **11**, 305–312 (1999).
56. Schmittgen, T. D. & Livak, K. J. Analyzing real-time PCR data by the comparative CT method. *Nature protocols* **3**, 1101–1108 (2008).

## Acknowledgements

This study was supported by the Program of National Beef Cattle and Yak Industrial Technology System (CARS-38), the National Natural Science Foundation of China (No. 31272408), Bio-breeding capacity-building and industry specific projects from the National Development and Reform Commission (2014–2573), Specific Projects of Science and Technology in Henan Province (141100110200), and Science and Technology Coordinator Innovative engineering projects of Shaanxi Province (No. 2014KTZB02-02-02-02, No. 2015KTCL02-08). We thank Dr. Xiaoying Li (Department of Forensic Medicine, Xinxiang Medical University, Xinxiang, Henan, China) for language editing.

## Author Contributions

M.J.Y. designed and supervised this study, M.J.Y., J.Q.L. and L.Z.Z. performed the experiments, M.J.Y., M.X.L. and Y.Z. performed the data analysis, M.J.Y. and J.Q.L. prepared the manuscript, X.Y.L., C.Z.L. and H.C. provided suggestions for the experimental design and provided funding assistance. All authors read and approved the final manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing Interests:** The authors declare no competing financial interests.

**How to cite this article:** Yang, M. J. *et al.* Association study and expression analysis of *CYP4A11* gene copy number variation in Chinese cattle. *Sci. Rep.* **7**, 46599; doi: 10.1038/srep46599 (2017).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017