Copy number variation of *WBP1L* gene revealed its association with growth traits across Chinese cattle populations

Jiaqiang Zhang[1],[*], Zijing Zhang[2],[*], Xian Liu[3], Yanan Chai[2], Peng Yang[1], Jungang Li[4], Yajun Huang[4], Lijuan Li[4], Weihong Huang[4], Guojie Yang[4], Zengfang Qi[4], Fuying Chen[2], Zhiming Li[3], Chuzhao Lei[1], Qiaoting Shi[2], Baorui Ru[3], Eryao Wang[2] and Yongzhen Huang[1] (iD)

[1]College of Animal Science and Technology, Northwest A&F University, Yangling, Shaanxi 712100, People's Republic of China; [2]Institute of Animal Husbandry and Veterinary Science, Henan Academy of Agricultural Sciences, Zhengzhou, Henan 450002, People's Republic of China; [3]Henan Provincial Animal Husbandry General Station, Zhengzhou, Henan 450008, People's Republic of China and [4]Jiaxian Animal Husbandry Bureau, Jiaxian, Henan 467100, People's Republic of China

Animal Research Paper

*These authors contributed equally to this work.

Abstract

Copy number variation (CNV) is a very common type of gene variation with high frequency. In recent years, CNV has been paid more attention in various fields, especially in livestock and poultry breeding, which has promoted the progress of breeding. WW domain binding protein 1-like (*WBP1L*) is a protein coding gene, which plays an important role in cattle populations, and its function has been extensively studied, but it is not clear whether the copy number of the gene can affect the growth and development of cattle populations. In this study, CNV of *WBP1L* gene was detected in 732 cattle of seven breeds (Qinchuan cattle, QC; Pinan cattle, PN; Yuengling cattle, YL; Xianan cattle, XN; Jiaxian cattle, JX; natural Guyuan cattle, NGY; Jian cattle, JA). In addition, the relationship between CNV and growth phenotype of cattle was studied. The experimental data indicate that the copy number of *WBP1L* was obviously correlated with heart girth of PN cattle (**$P < 0.01$), rump length (RL) and body weight (BW) of PN cattle (*$P < 0.05$), withers height (WH), RL, body length, chest depth and BW of JX cattle (*$P < 0.05$), WH of NGY cattle (*$P < 0.05$) and WH of JA cattle (*$P < 0.05$). It was proved that CNV of *WBP1L* gene could be used as molecular marker locus for genetic breeding of the above four cattle breeds.

Introduction

The WW domain is composed of 38–40 semi-conserved amino acid residues that mediate protein–protein interactions and are functionally similar to the SH3 domain (Chen *et al.*, 1997; Meiyappan *et al.*, 2007), a model of protein–protein interactions. WW domain binding protein 1-like (*WBP1L*), a transmembrane adaptor protein (Guan *et al.*, 2020), can act as a regulator to identify a prognostic factor known as acute leukaemia 1 (*OPAL 1*) (Mosquera-Caro *et al.*, 2003; Neveu *et al.*, 2016), binding to a variety of receptors in the body to regulate the expression level of related materials. Sequence analysis showed that *WBP1L* is a transmembrane aptamer protein. It has short extracellular and luminal region, and can form a small compact domain with a single transmembrane domain and a large cytoplasmic tail (Pei and Grishin, 2012). It has been found that mice deficient in *WBP1L* regulate B cell development and improve the efficiency of bone marrow cell transplantation (Borna *et al.*, 2020). In addition, smoking causes *WBP1L* methylation in a pregnant woman's placenta, which regulates birth weight of a baby (Cardenas *et al.*, 2019). Therefore, we hypothesized that this gene might affect cattle growth traits. Most of the research on *WBP1L* gene is focused on humans. However, the detailed function of *WBP1L* gene in livestock and poultry has not been reported.

China has a vast territory, complex climate and terrain, a wide variety of organisms, and extremely rich genetic resources. However, the shortcomings of local varieties are particularly obvious, which are highlighted in slow growth rate, low feed utilization rate, low lean meat rate and high fat content. Therefore, in order to meet people's higher consumption demand and improve the economic benefits of breeding industry, we must breed superior individuals using some technical method and improve their growth performance.

As a kind of gene mutation, copy number variations (CNVs) have been widely used in the genetic improvement of livestock and poultry (Cheng *et al.*, 2020). CNVs are chromosomal structural variations that are greater than 50 bp in length in the genome (Iafrate *et al.*, 2004), including duplicates, deletions, transformations, inversions and large-scale insertions,

of which the most common type is duplicates (Feuk *et al.*, 2006; MacDonald *et al.*, 2014). CNVs are characterized by high resolution, large genome coverage and stable genetic performance. In recent years, a large number of gene CNVs have been found to be related to the growth characteristics of livestock and poultry, which has great significance for livestock breeding and genetic breeding. However, the relationship between CNVs of *WBP1L* gene and beef cattle growth traits has not been studied so far.

The *WBP1L* gene is located on chromosome 26 of cattle and has a total length of 69 985 bp. The known CNV was detected by quantitative polymerase chain reaction (qPCR) and its copy number was determined by the $2 \times 2^{-\Delta\Delta C_t}$ method. In the current study, CNV of the *WBP1L* gene was correlated with growth traits in 732 cattle from seven breeds (Qinchuan cattle, QC; Pinan cattle, PN; Yuengling cattle, YL; Xianan cattle, XN; Jiaxian cattle, JX; natural Guyuan cattle, NGY; Jian cattle, JA). However, the main subjects in this study were PN, JX, NGY and JA cattle, while the other three breeds were used as control experiments for data comparison. The results showed that CNV of *WBP1L* gene could be used as an important marker for PN, JX, NGY and JA molecular breeding.

Materials and methods

Cattle and growth trait measurements

Previous sequencing revealed that CNV of *WBP1L* gene existed in cattle population. In this study, 732 individuals from seven Chinese cattle breeds were selected to analyse the correlation between growth characteristics of cattle breeds and CNVs of *WBP1L* gene. They are QC ($n = 85$), PN ($n = 129$), YL ($n = 101$), XN ($n = 121$), JX ($n = 122$), NGY ($n = 113$) and JA ($n = 61$). QC are one of the five big cattle breeds in China, large in size, strong in service and good in producing meat. It is named for the eight-hund-li Qin Chuan, which is produced in Guanzhong area of Shaanxi province. PN is a breed of hybrid cattle between Piedmont cattle and Nanyang cattle. It has delicious meat and high nutritional value. YL is the first breed of Chinese cattle bred by three-way crossbreeding in China, which has wide adaptability and excellent performance. XN is a new breed of Chinese cattle with Charolai cattle as its male parent and Nanyang cattle as its female parent. It is also the first Chinese cattle breed with independent intellectual property rights in China. JX originated in Jiaxian county, Henan province, and is a famous local cattle breed with both service and meat. NGY, a national geographical indication product of China, is produced in Guyuan city, Ningxia Hui autonomous region. JA is one of the three local fine breed cattle in Jiangxi province, which has good endurance and resistance to adversity.

All animal experiments complied with the ARRIVE guidelines and were carried out in accordance with the UK Animals (Scientific Procedures) Act, 1986 and associated guidelines, EU Directive 2010/63/EU for animal experiments, or the National Institutes of Health guide for the care and use of Laboratory animals (NIH Publications No. 8023, revised 1978). In addition, the Animal Experiment Review Committee of Northwest A&F University approved all of our experiments. Based on the International Animal Guidelines, we strictly adhere to animal testing policies and safeguard animal welfare. During the same breeding period, our experimental cattle were fed with the same nutritional conditions and the same management methods of feeding and drinking water. The correlation between CNV of

WBP1L gene and the growth traits of 732 cattle of seven breeds was studied in a randomized experimental population. It should be noted that none of the cattle in the study were related.

Sampling and DNA extraction of cattle tissue

We put the extracted ear sample into the sample tube filled with alcohol and shipped it immediately back to the laboratory. During the extraction of genomic DNA, the ear samples collected were first homogenized and fully lysed with cell lysate, then the proteins were removed by protease, chloroform was extracted once, preparation solution was precipitated, RNA was removed by RNase and ethanol was precipitated to obtain the complete DNA. We diluted the DNA to 25 ng/µl and stored in a refrigerator at −40°C over a long period of time. Prior to the experiment, genomic DNA integrity was determined by agarose gel electrophoresis and its concentration was determined by a Nanodroplet 2000 ultramicrophotometer (Thermoscience, Waltham, MA, USA).

Primer design and amplification detection

After genome sequencing and analysis of the CNV area of *WBP1L* gene, the cattle *WBP1L* gene sequence (AC_000183.1) issued in NCBI was used as the reference sequence in this study, and the CNVR location between 23 641 732 and 23 643 331 bp was determined, with a total length of 1600 bp (Huang *et al.*, 2021) (Fig. 1). Primers were designed in CNVR according to the restriction conditions, and multiple pairs of primers were designed for PCR experiment. Primer information is shown in Table 1.

Agarose gel electrophoresis was used to detect their specificity; next the best pair of primers was selected. We selected *BTF3* gene as our experiment internal reference gene. Using the same method, our primers were designed. The reaction system was as follows: 5 µl 2× Taq PCR Master Mix, 0.4 µl for each primer, 3 µl ddH$_2$O and 1 µl DNA sample to form a 10 µl premixture system. PCR procedures for: 95°C 2 min, 95°C 10 s, 62°C 10 s, 72°C 10 s, 72°C for 10 min, 10°C up, extending stage circulation 35 times, and then use 1.5% agarose gel electrophoresis to detect primers for the specificity of the amplification products.

Experimental process and method

We repeated individual quantitative experiments three times for each animal, and then sent to company to obtain measurement data, viewing the data provided by CFX software; the same method was used to perform a simple comparison between individual's three sets of data, by removing two other significant difference of the data, and then calculated the average. The body size data of individual animals were matched with the analysis results. We used Excel software combined with $2 \times 2^{-\Delta\Delta C_t}$ to probe into the copy number of *WBP1L* gene in bovine genome. The correlation between CNV of *WBP1L* gene and growth traits of each animal was analysed using SPSSV23.0 software (SPSS, Inc., Chicago, IL, USA) with CNV type as the fixed factor and body size traits as the dependent variable.

Statistical analysis

This study tests the CNV of *WBP1L* gene and the correlation between bovine growth characters, and uses the $2 \times 2^{-\Delta\Delta C_t}$ method to calculate *WBP1L* gene in the bovine genome copy
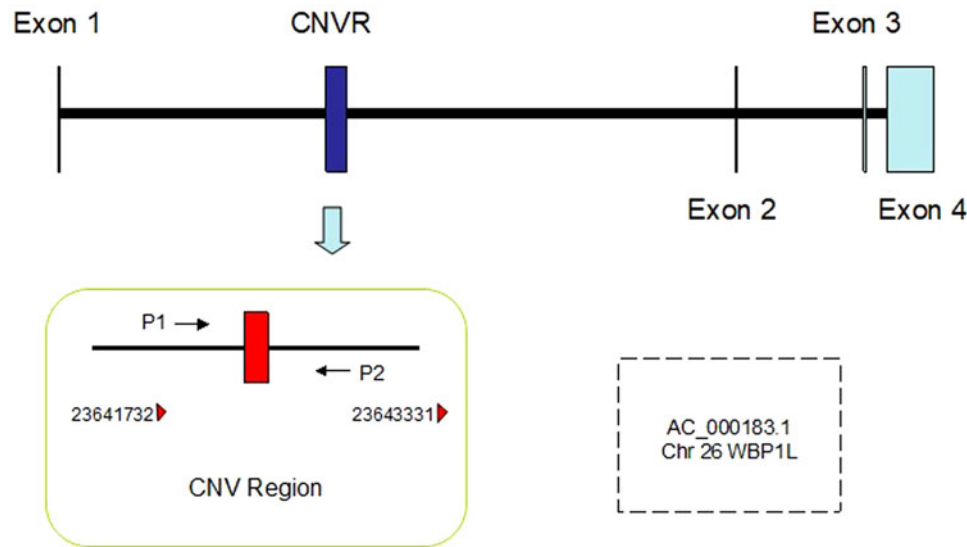
Fig. 1. (Colour online) Region of *WBP1L* gene CNV in cattle. Blue area stands for the coding area. The CNV area is from 23 641 732 to 23 643 331 in Chr26. Primer matching begins from 23 371 659 to 23 371 803 and the detected target sequence length of CNV is 145 bp. In addition, the total length of the CNV region is 1600 bp.

Table 1. Primers in correlation experiments

Primer pairs	Genes	Primer sequence (5′–3′)	Amplification length (bp)
WBP1L-CNV	*WBP1L*	CCAGAGCTGGACTTCGTGGG	145
		TCGTGACATATCTCAGACGCAG	
BTF3-CNV	*BTF3*	AACCAGGAGAAACTCGCCAA	166
		TTCGGTGAAATGCCCTCTCG	

number; the final result for '0' or '1' was represented with the letter 'Deletion', the results for '2' with the letter 'Normal', the results of larger than 2 were expressed with the letter 'Duplication' (for more than 20 sets of data, no follow-up analysis was performed). In order to analyse the influence of CNV of *WBP1L* gene on the phenotypic value of the trait, and to ensure the minimum effect of various factors on this experiment, we established the following simplified mathematical model:

$$Z_{ijk} = \mu + \text{CNV}_i + X_j + Y_k + e_{ijk} \qquad (1)$$

where Z_{ijk} is the phenotype value of traits of experimental cattle; μ is the average value of each experimental cattle trait, CNV_i is the effect of copy number variation type of *WBP1L* gene, X_i is the influencing factor of gender, Y_i is the influencing factor of age and e_{ijk} is the random residual error. Furthermore, the experimental results were expressed as mean ± S.E. values.
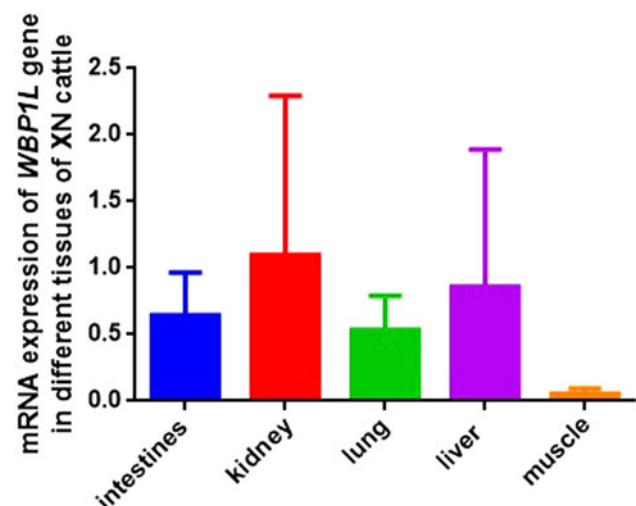
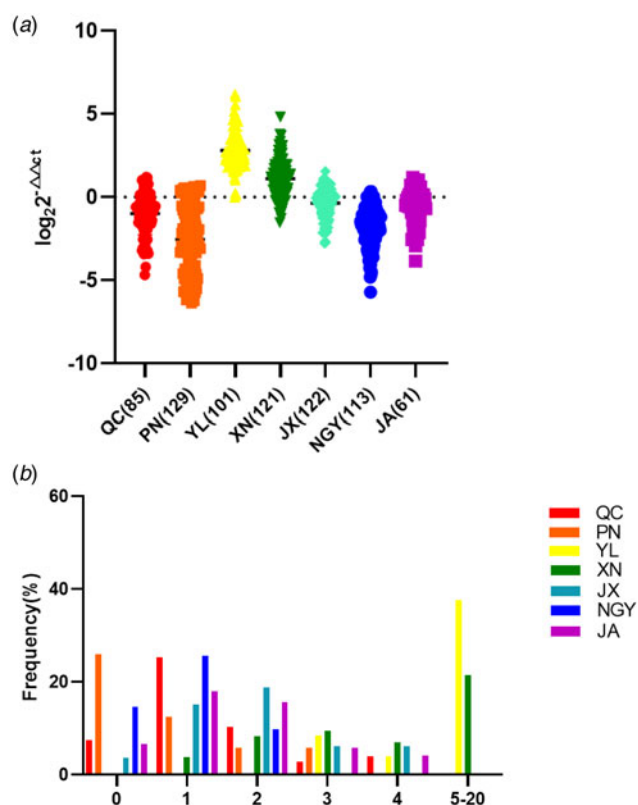## Results

Expression of *WBP1L* gene in XN cattle

To detect the expression level of *WBP1L* gene in Chinese cattle, XN cattle breed was selected for expression profile analysis (Fig. 2). The research data show that the expression level of *WBP1L* gene was highest in the kidney, followed by the liver, intestines and lung, which had similar and low expression levels.

Distribution of CNV types in seven Chinese cattle breeds

Previous studies have confirmed that the gene CNV and livestock growth traits are correlated; in order to study this correlation, the copy number distribution of seven cattle breeds was examined, including QC, PN, YL, XN, JX, NGY and JA. According to the

Fig. 2. (Colour online) Expression profiling of *WBP1L* in XN cattle.

Fig. 3. (Colour online) Distribution of *WBP1L* gene copy variation in seven cattle breeds (*a*). The percentage of *WBP1L* gene CNV copy variation in seven cattle breeds (*b*).

Correlation analysis between CNV and growth traits of experimental cattle

With the development of molecular breeding technology, CNV molecular markers have gradually attracted people's attention. According to previous studies, CNV is closely related to the growth traits of livestock and poultry. In this study, SPSSV23.0 software (SPSS, Inc., Chicago, IL, USA) was used to analyse the correlation between different copy number types and different cattle growth traits. The results showed that the copy number of *WBP1L* was significantly correlated with heart girth (HG) of PN cattle (**$P < 0.01$), and the growth traits of deletion type cattle were significantly better than those of normal type and duplication type cattle (Table 2). The withers height (WH), body length (BL), rump length (RL), chest depth (CD) and body weight (BW) of JX cattle were significantly correlated with the copy number types (*$P < 0.05$), and the growth traits of normal type and duplication type cattle were significantly better than deletion type ones, but there was no significant difference in the growth traits between normal type and duplication type cattle (Table 3). WH of NGY cattle were significantly correlated with copy number types (*$P < 0.05$), and the growth traits of deletion type individuals were the worst (Table 4). WH of JA cattle were significantly correlated with copy number types (*$P < 0.05$), and the growth traits of normal type individuals were the best, while those of duplication type individuals were the worst (Table 5). Surprisingly, no significant correlation was found between the copy number types of QC, YL and XN and growth traits in this study.

Discussion

At present, people's understanding of *WBP1L* gene is not deep enough, and the research field focuses on the study of human diseases. For example, Andreassen *et al.* found that *WBP1L* gene is related to human schizophrenia (Andreassen *et al.*, 2013). Zhang *et al.* found a certain association between *WBP1L* gene and human hypertension (Zhang *et al.*, 2019). However, no related studies have been found, and we hope to discover the potential role of the gene in animal studies.

CNV is caused by genome rearrangement, which refers to the increase or decrease of the copy number of large segments (Zhang *et al.*, 2011; Vissers and Stankiewicz, 2012). It is mainly manifested as deletion or duplication at the submicroscopic level.

$2 \times 2^{-\Delta\Delta C_t}$ values, the types of CNV were classified into deletion type, normal type and duplication type. As shown in Fig. 3(*a*), the CNV of *WBP1L* gene of seven cattle breeds is concentrated in less than 0. As shown in Fig. 3(*b*), the CNV of *WBP1L* gene was more between 0 and 4 copies and less between 5 and 20 copies. The results showed that the deletion type individuals of PN (77%), NGY (81%) and JA (49%) were significantly higher than the other two types. Surprisingly, YL without deletion and normal type individuals, duplication type individuals were up to 100%, NGY without duplication type individuals, deletion type individuals were up to 81%.

Table 2. Statistical correlation analysis of growth traits in PN cattle with WBP1L-CNV

	CNV types (average value ± standard error)				
	Deletion		Normal	Duplication	
Growth traits	CN = 0 (*n* = 67)	CN = 1 (*n* = 32)	CN = 2 (*n* = 15)	CN ≥ 3 (*n* = 15)	*P* value
WH (cm)	125 ± 6.3	123 ± 5.7	123 ± 4.3	121 ± 7.5	0.078
BL (cm)	148 ± 12.7	146 ± 10.5	146 ± 10.1	1400 ± 14.3	0.154
HAHC (cm)	132 ± 6.7	130 ± 5.8	130 ± 4.0	129 ± 7.4	0.192
HG (cm)	175 ± 12.7	171 ± 12.1	166 ± 7.5	165 ± 15.1	0.004
HW (cm)	47 ± 4.3	46 ± 4.1	46 ± 3.3	45 ± 5.0	0.268
RL (cm)	49 ± 3.8	47 ± 3.6	47 ± 2.9	46 ± 4.5	0.028
BW (kg)	428 ± 94.0	399 ± 80.4	375 ± 53.0	360 ± 105.8	0.021

HAHC, height at hip cross; HW, hip width.
Note: The BW is the estimated value, and the estimation method is the square of HG × BL/10 800.

Table 3. Statistical correlation analysis of growth traits in JX cattle with WBP1L-CNV

| | CNV types (average value ± standard error) | | | | | |
| Growth traits | Deletion | | Normal | Duplication | | |
	CN = 0 (n = 9)	CN = 1 (n = 37)	CN = 2 (n = 46)	CN = 3 (n = 15)	CN $\geq$ 4 (n = 15)	P value
WH (cm)	121 ± 6.3	121 ± 7.1	124 ± 6.3	127 ± 6.5	123 ± 8.9	0.031
BL (cm)	144 ± 9.9	140 ± 11.4	144 ± 10.1	150 ± 10.1	139 ± 13.5	0.024
HG (cm)	170 ± 13.5	170 ± 16.5	174 ± 11.3	180 ± 9.9	168 ± 15.9	0.082
HW (cm)	45 ± 4.8	44 ± 4.7	45 ± 4.5	47 ± 4.2	43 ± 5.2	0.104
HLW (cm)	25 ± 5.9	23 ± 5.9	23 ± 5.0	26 ± 8.0	22 ± 5.0	0.387
RL (cm)	46 ± 4.6	45 ± 3.9	46 ± 4.5	48 ± 3.1	43 ± 5.2	0.033
HAHC (cm)	123 ± 6.7	122 ± 6.6	125 ± 5.8	127 ± 7.2	124 ± 8.1	0.081
CD (cm)	63 ± 3.6	61 ± 5.2	64 ± 4.9	66 ± 4.3	63 ± 6.8	0.041
CW (cm)	36 ± 3.6	37 ± 4.6	37 ± 3.3	38 ± 2.4	37 ± 6.4	0.871
BW (kg)	390 ± 79.3	382 ± 93.8	405 ± 74.5	454 ± 74.9	368 ± 94.3	0.038

HLW, hucklebone width; CW, chest width.

These mutations are not only individual normal polymorphisms, but also pathogenic mutations, which are widely distributed in livestock and poultry genomes (Gheyas *et al.*, 2013; Zhao *et al.*, 2020). Single-nucleotide polymorphisms (SNPs) and CNVs are two important potential sources of phenotypic variation in livestock and poultry, and the mutation rate of CNV sites is higher than that of SNPs. It is one of the most common types of variation. However, it should be pointed out that there are still many problems in the study of CNV: (1) due to the limitation of detection technology level, the detection effectiveness of small CNVs is limited; (2) due to the immaturity of relevant technologies and the lack of research data, the discovery rate of CNVs is much lower than SNP at present and (3) the CNV location reported at present is the location of a region containing CNV, and there are many possible mutations. There is still no good way to measure the exact mutation location or gene. All these will affect the research progress of CNV.

While studying the correlation between CNVs and phenotypes, some studies indicated that SNP genotypes and CNV measurements were statistically significantly correlated with 83 and 18% of gene-expression traits, respectively, but this underestimated the role of CNVs. Therefore, more complete and accurate detection techniques need to be proposed. Although CNV research is still incomplete, it has made many achievements: in human diseases (Henrichsen *et al.*, 2009), CNV has been used to study multiple myeloma (Yang *et al.*, 2017*a*, 2017*b*); leprosy (Ali *et al.*, 2013) and autism (Glessner *et al.*, 2009). In animal breeding, such as *IGF1R* gene of CNV and Jin Na salience (Jacqueline Nottingham) cow's weight and the height of QC and bone was significantly associated with wide (Ma *et al.*, 2019), the *CADM2* gene CNV of Guizhou black goats was obviously correlated with its shoulder height and BL, which lack individual in WH and BL of replication and normal individuals (Xu *et al.*, 2020), the CNV of *MLLT 10* gene had significant effects on the HW and RL of cattle (*P* < 0.05) (Yang *et al.*, 2020). By inserting a 19 bp repeat sequence into the promoter region of *CDKN3* gene, it was found that the growth and carcase traits of chicken were significantly affected (Li *et al.*, 2019). The insertion mutation of *lncFAM* 97 bp was closely related to the growth traits of chickens at different growth stages (Li *et al.*, 2020). In addition, CNV has been used for breeding in sheep (Jenkins *et al.*, 2016; Yang *et al.*, 2018) and monkeys (Boyd *et al.*, 1980). Our study found that CNV of *WBP1L* gene was significantly correlated with growth traits of a variety of Chinese yellow cattle.

In order to detect the relationship between different copy numbers of *WBP1L* gene and growth performance of QC, a total of 732 Chinese cattle breeds were tested. Of all the experimental results, there are four varieties showed significant correlation, and only QC HG a character showed significant correlation with the gene copy number, four varieties of nine characters showed significant correlation, but only a HG of Jiaxian red cattle breed attained up to five traits that showed significant correlation, WH was significantly correlated with JX, NGY and JA, and BL and BW were also significantly correlated with PN and JX populations. In addition, different from other populations, BL and CD were correlated in JX population. Among the seven herds, some had the best CNV of the normal type, some had the best CNV, while others had the best deletion type.

Table 4. Statistical correlation analysis of growth traits in NGY cattle with WBP1L-CNV

| | CNV types (average value ± standard error) | | | |
| Growth traits | Deletion | | Normal | |
	CN = 0 (n = 33)	CN = 1 (n = 58)	CN = 2 (n = 22)	P value
WH (cm)	115. ± 19.3	122 ± 6.0	121 ± 6.5	0.038
HAHC (cm)	120 ± 19.4	123 ± 15.4	124 ± 6.9	0.641
BL (cm)	142 ± 11.5	147 ± 10.0	145 ± 9.5	0.124
HG (cm)	173 ± 12.1	176 ± 10.6	174 ± 12.1	0.504
BW (kg)	401 ± 77.3	425 ± 72.7	409 ± 77.9	0.313

Note: The BW is the estimated value, and the estimation method is the square of HG × BL/10 800.

Table 5. Statistical correlation analysis of growth traits in JA cattle with WBP1L-CNV

Growth traits	CNV types (average value ± standard error)					P value
	Deletion		Normal	Duplication		
	CN = 0 (n = 8)	CN = 1 (n = 22)	CN = 2 (n = 19)	CN = 3 (n = 7)	CN ≥ 4 (n = 5)	
WH (cm)	103 ± 3.5	107 ± 5.9	110 ± 8.0	104 ± 2.5	103 ± 4.8	0.035
HAHC (cm)	106 ± 3.0	109 ± 5.6	111 ± 6.8	106 ± 3.2	106 ± 3.4	0.134
BL (cm)	115 ± 5.4	119 ± 10.7	120 ± 11.2	113 ± 7.1	113 ± 3.6	0.342
HG (cm)	150 ± 4.5	150 ± 10.6	154 ± 15.2	141 ± 3.1	146 ± 4.0	0.118
CBC (cm)	13.8 ± 0.53	14 ± 1.0	14.2 ± 0.65	14.5 ± 0.41	14 ± 1.0	0.339
BW (kg)	239 ± 21.1	251 ± 52.8	268 ± 74.1	207 ± 18.2	225 ± 16.8	0.115

CBC, cannon bone circumference.

Note: The BW is the estimated value, and the estimation method is the square of HG × BL/10 800.

Conclusions

In conclusion, our study found that the CNV of *WBP1L* gene was obviously correlated with ten traits of four breeds of cattle, which in a sense indicated that these traits could be used for cattle seed selection, promoting breed improvement and promoting genetic and breeding work.

References

Ali S, Srivastava AK, Chopra R, Aggarwal S, Garg VK, Bhattacharya SN and Bamezai RN (2013) IL12B SNPs and copy number variation in IL23R gene associated with susceptibility to leprosy. *Journal of Medical Genetics* **50**, 34–42.

Andreassen OA, Djurovic S, Thompson WK, Schork AJ, Kendler KS, O'Donovan MC, Rujescu D, Werge T, van de Bunt M, Morris AP, McCarthy MI, International Consortium for Blood Pressure GWAS, Diabetes Genetics Replication and Meta-analysis Consortium, Psychiatric Genomics Consortium Schizophrenia Working Group, Roddey JC, McEvoy LK, Desikan RS and Dale AM (2013) Improved detection of common variants associated with schizophrenia by leveraging pleiotropy with cardiovascular-disease risk factors. *American Journal of Human Genetics* **92**, 197–209.

Borna S, Drobek A, Kralova J, Glatzova D, Splichalova I, Fabisik M, Pokorna J, Skopcova T, Angelisova P, Kanderova V, Starkova J, Stanek P, Matveichuk OV, Pavliuchenko N, Kwiatkowska K, Protty MB, Tomlinson MG, Alberich-Jorda M, Korinek V and Brdicka T (2020) Transmembrane adaptor protein WBP1L regulates CXCR4 signalling and murine haematopoiesis. *Journal of Cellular and Molecular Medicine* **24**, 1980–1992.

Boyd ES, Boyd EH and Brown LE (1980) The M-wave and CNV in the squirrel monkey: generality of cue modality and of reward. *Electroencephalography and Clinical Neurophysiology* **49**, 66–80.

Cardenas A, Lutz SM, Everson TM, Perron P, Bouchard L and Hivert MF (2019) Mediation by placental DNA methylation of the association of prenatal maternal smoking and birth weight. *American Journal of Epidemiology* **188**, 1878–1886.

Chen HI, Einbond A, Kwak SJ, Linn H, Koepf E, Peterson S, Kelly JW and Sudol M (1997) Characterization of the WW domain of human yes-associated protein and its polyproline-containing ligands. *The Journal of Biological Chemistry* **272**, 17070–17077.

Cheng J, Jiang R, Cao X, Huang Y, Lan X, Lei C, Hu L and Chen H (2020) Association analysis of KMT2D copy number variation as positional candidate for sheep stature. *Gene* **2020**, 144799.

Feuk L, Carson AR and Scherer SW (2006) Structural variation in the human genome. *Nature Reviews Genetics* **7**, 85–97.

Gheyas AA and Burt DW (2013) Microarray resources for genetic and genomic studies in chicken: a review. *Genesis (New York, N.Y.: 2000)* **51**, 337–356.

Glessner JT, Wang K, Cai G, Korvatska O, Kim CE, Wood S, Zhang H, Estes A, Brune CW, Bradfield JP, Imielinski M, Frackelton EC, Reichert J, Crawford EL, Munson J, Sleiman PM, Chiavacci R, Annaiah K, Thomas K, Hou C, Glaberson W, Flory J, Otieno F, Garris M, Soorya L, Klei L, Piven J, Meyer KJ, Anagnostou E, Sakurai T, Game RM, Rudd DS, Zurawiecki D, McDougle CJ, Davis LK, Miller J, Posey DJ, Michaels S, Kolevzon A, Silverman JM, Bernier R, Levy SE, Schultz RT, Dawson G, Owley T, McMahon WM, Wassink TH, Sweeney JA, Nurnberger JI, Coon H, Sutcliffe JS, Minshew NJ, Grant SF, Bucan M, Cook EH, Buxbaum JD, Devlin B, Schellenberg GD and Hakonarson H (2009) Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. *Nature* **459**, 569–573.

Guan F, Ni T, Han W, Lin H, Zhang B, Chen G, Zhu L, Liu D and Zhang T (2020) Evaluation of the relationships of the WBP1L gene with schizophrenia and the general psychopathology scale based on a case-control study. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics* **183**, 164–171.

Henrichsen CN, Chaignat E and Reymond A (2009) Copy number variants, diseases and gene expression. *Human Molecular Genetics* **18**, R1–R8.

Huang Y, Li Y, Wang X, Yu J, Cai Y, Zheng Z, Li R, Zhang S, Chen N, Asadollahpour Nanaei H, Hanif Q, Chen Q, Fu W, Li C, Cao X, Zhou G, Liu S, He S, Li W, Chen Y, Chen H, Lei C, Liu M and Jiang Y

(2021) An atlas of CNV maps in cattle, goat and sheep. *Science China Life Sciences* **64**, 1747–1764.

Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, Scherer SW and Lee C (2004) Detection of large-scale variation in the human genome. *Nature Genetics* **36**, 949–951.

Jenkins GM, Goddard ME, Black MA, Brauning R, Auvray B, Dodds KG, Kijas JW, Cockett N and McEwan JC (2016) Copy number variants in the sheep genome detected using multiple approaches. *BMC Genomics* **17**, 441.

Li W, Liu D, Tang S, Li D, Han R, Tian Y, Li H, Li G, Li W, Liu X, Kang X and Li Z (2019) A multiallelic indel in the promoter region of the cyclin-dependent kinase inhibitor 3 gene is significantly associated with body weight and carcass traits in chickens. *Poultry Science* **98**, 556–565.

Li W, Jing Z, Cheng Y, Wang X, Li D, Han R, Li W, Li G, Sun G, Tian Y, Liu X, Kang X and Li Z (2020) Analysis of four complete linkage sequence variants within a novel lncRNA located in a growth QTL on chromosome 1 related to growth traits in chickens. *Journal of Animal Science* **98**, skaa122.

Ma YL, Wen YF, Cao XK, Cheng J, Huang YZ, Ma Y, Hu LY, Lei CZ, Qi XL, Cao H and Chen H (2019) Copy number variation (CNV) in the IGF1R gene across four cattle breeds and its association with economic traits. *Archives Animal Breeding* **62**, 171–179.

MacDonald JR, Ziman R, Yuen RK, Feuk L and Scherer SW (2014) The database of genomic variants: a curated collection of structural variation in the human genome. *Nucleic Acids Research* **42**, D986–D992.

Meiyappan M, Birrane G and Ladias J (2007) Structural basis for polyproline recognition by the FE65 WW domain. *Journal of Molecular Biology* **372**, 970–980.

Mosquera-Caro M, Helman P, Veroff R, Shuster J, Martin S, Davidson G, Potter J, Harvey R, Hromas R, Andries E, Atlas S, Wilson C, Ar K, Xu YX, Chen IM, Carroll A, Camitta B and Willman C (2003) Identification, validation and cloning of a novel gene (OPAL1) and associated genes highly predictive of outcome in pediatric acute lymphoblastic leukemia using gene expression profiling. *Blood* **102**, 4a.

Neveu B, Spinella JF, Richer C, Lagacé K, Cassart P, Lajoie M, Jananji S, Drouin S, Healy J, Hickson GR and Sinnett D (2016) CLIC5: a novel ETV6 target gene in childhood acute lymphoblastic leukemia. *Haematologica* **101**, 1534–1543.

Pei J and Grishin NV (2012) Unexpected diversity in Shisa-like proteins suggests the importance of their roles as transmembrane adaptors. *Cellular Signalling* **24**, 758–769.

Vissers LE and Stankiewicz P (2012) Microdeletion and microduplication syndromes. *Methods in Molecular Biology (Clifton, NJ)* **838**, 29–75.

Xu Z, Wang X, Zhang Z, An Q, Wen Y, Wang D, Liu X, Li Z, Lyu S, Li L, Wang E, Ru B, Xu Z and Huang Y (2020) Copy number variation of CADM2 gene revealed its association with growth traits across Chinese *Capra hircus* (goat) populations. *Gene* **741**, 144519.

Yang M, Lv J, Zhang L, Li M, Zhou Y, Lan X, Lei C and Chen H (2017a) Association study and expression analysis of CYP4A11 gene copy number variation in Chinese cattle. *Scientific Reports* **7**, 46599.

Yang L, Wang YZ, Zhu HH, Chang Y, Li LD, Chen WM, Long LY, Zhang YH, Liu YR, Lu J and Qin YZ (2017b) PRAME gene copy number variation is related to its expression in multiple myeloma. *DNA and Cell Biology* **36**, 1099–1107.

Yang L, Xu L, Zhou Y, Liu M, Wang L, Kijas JW, Zhang H, Li L and Liu GE (2018) Diversity of copy number variation in a worldwide population of sheep. *Genomics* **110**, 143–148.

Yang P, Zhang Z, Xu J, Qu K, Lyv S, Wang X, Cai C, Li Z, Wang E, Xie J, Ru B, Xu Z, Lei C, Chen H, Huang B and Huang Y (2020) The association of the copy number variation of the MLLT10 gene with growth traits of Chinese cattle. *Animals: An Open Access Journal from MDPI* **10**, 250.

Zhang Y, Jiang WK and Gao LZ (2011) Evolution of microRNA genes in *Oryza sativa* and *Arabidopsis thaliana*: an update of the inverted duplication model. *PLoS One* **6**, e28073.

Zhang H, Mo X, Zhou Z, Zhu Z, HuangFu X, Guo Z and Zhang Y (2019) Detection of putative functional single nucleotide polymorphisms in blood pressure loci and validation of association between single nucleotide polymorphism in WBP1L and hypertension in the Chinese Han population. *Journal of Cardiovascular Pharmacology* **73**, 48–55.

Zhao Y, Dong L, Jiang C, Wang X, Xie J, Rashid M, Liu Y, Li M, Bu Z, Wang H, Ma X, Sun S, Wang X, Bo C, Zhou T and Kong L (2020) Distinct nucleotide patterns among three subgenomes of bread wheat and their potential origins during domestication after allopolyploidization. *BMC Biology* **18**, 188.