Genome-wide scan for runs of homozygosity identifies potential candidate regions associated with growth traits in beef cattle

Guoyao Zhao, Tianliu Zhang, Yuqiang Liu, Zezhao Wang, Lei Xu, Bo Zhu, Xue Gao, Lupei Zhang, Huijiang Gao, Junya Li, Lingyang Xu

Guoyao Zhao
Institute of animal science Chinese academy of Agricultural Sciences

Tianliu Zhang
Institute of Animal Science Chinese Academy of Agricultural Sciences

Yuqiang Liu
Institute of animal Science Chinese Academy of Agricultural Sciences

Zezhao Wang
Institute of Animal Science Chinese Academy of agricultural Sciences

Lei Xu
Chinese Academy of Agricultural Sciences Institute of Animal Science

Bo Zhu
Institute of animal science Chinese Academy of Agricultural Sciences

Xue Gao
Institute of Animal Science Chinese Academy of Agricultural Sciences

Lupei Zhang
Institute of Animal Science Chinese Academy of Agricultural Sciences

Huijiang Gao
Institute of Animal Science Chinese Academy of Agricultural Sciences

Junya Li
Institute of Animal Science Chinese academy of Agricultural Sciences

Lingyang Xu
Chinese Academy of Agricultural Sciences Institute of Animal Science

Corresponding Author
xulingyang@163.com

Prescreen
Abstract

Background: Runs of homozygosity (ROH) are continuous homozygous regions that generally exist in the DNA sequence of diploid organisms. Identifications of regions of the genome lead to reduction in performance can provide valuable insight into the genetic architecture of complex traits. Here, we evaluated genome-wide patterns of homozygosity and their association with growth traits in a commercial beef cattle population.

Results: We identified a total of 29,271 ROH segments with an average number of 63.36 and an average length of 0.98 Mb in this commercial beef cattle population, representing ~2.53% (~63.36Mb) of the genome. To evaluate the enrichment of ROH across genomes, we initially identified 280 ROH regions by merging ROH events identified across all individuals. Of these, nine regions were significantly associated with six growth phenotype traits (body height, chest circumference, fat coverage, backfat thickness, ribeye area, carcass length; P<0.01), which contain 187 candidate genes. Furthermore, we found 26 consensus ROH regions with frequencies exceeding 10%, and several of these consensus overlapped with QTLs which are associated with weight gain, calving difficulty and stillbirth. To precisely locate locus within each ROH for every studied trait, we further utilized loci-based methods for association analysis among these identified regions. Totally, we obtained 9,360 loci within ROH, and 1,631 loci displaying significant association (P<0.01) for eight traits. In addition, we found that 67 genes embedded with homozygous loci. Several identified candidate genes, including EBF2, SLC20A2, SH3BGR2, HMGA1 and ACSL1, were related to growth traits.

Conclusions: This study assessed genome-wide autozygosity pattern and inbreeding level in a commercial beef cattle population. Our study identified many candidate regions and genes with ROH for growth traits in beef cattle, which can provide important insights into investigating homozygosity across genome in other farm animals. Our findings may further be unitized to assist the design of selection mating strategy.

Background

Runs of homozygosity (ROH) are continuous homozygous regions that generally exist in the DNA sequence of diploid organisms [1]. Several studies have suggested that the individual inherits the chromosomal fragment that are identity by descent (IBD) from both parents, resulting in homozygous segments in the offspring’s genome[2, 3]. Many factors can contribute to the formation of ROH including genetic drift, inbreeding and intensive selection [4]. With the development of high density array and whole-genome sequencing technologies, the emergence of high-throughput approaches provided unprecedented opportunities to explore homozygosity segments in high resolution [5].

Initially, Broman et al. [3] proposed that ROH exists widely in the human genome and has an important impact on human health. Gibson et al. [2] further explored different distribution frequencies of ROH with different lengths using high-density SNP arrays, these study promoted the emergence for study of ROH in human genetics. Additional studies found ROH can provide insights into elucidating population structure and demography in human [6–8] and genetic relationships, inbreeding level and selection pressure for farm animals [9–14]. The enrichment of ROH may increase the harmful recessive alleles, and reduce the survival ability of individuals [15]. Several studies found a significant association between ROH and the occurrence of recessive diseases [16–18]. Other studies also suggested ROHs were related to human diseases including schizophrenia [16], rheumatoid arthritis [19] and Alzheimer’s disease[20].

In recent years, ROH has been widely utilized to assess the genome wide inbreeding level and genetic relationship in farm animals [12, 21–30]. ROH can provide a novel perspective for detecting detrimental variation in the genome, assessing inbreeding level and offer important information of genetic and demographic history [31]. In cattle, Ferencakovic et al. investigated genome-wide ROH pattern and revealed inbreeding coefficients calculated by ROH were useful for assessing inbreeding level in cattle [9]. Also, more accurate estimation based on ROH can be obtained for inbreeding coefficient, when compared with the pedigree data [22]. Moreover, ROH can reflect the genetic relationship and inbreeding levels among groups, and affect the selection on the genomic region [27, 32]. Previous studies suggested the formation of ROH may involve the selection of complex traits and some regions of ROH were identified which are related to economical important traits in farm animals [26, 33, 34]. However, there is no systematic assessment of runs homozygosity pattern
and their association with important traits in commercial beef cattle populations. The commercial beef cattle population used on current study is a hybrid population between Wagyu cattle and Fuzhou cattle in China. Based on this population, we have previously identified many candidate genes associated with body measurement traits [35] and fatty acid composition in these cattle [36].

The aims of this study were to (i) assess genome-wide autozygosity patterns and estimate inbreeding level based on ROH in this commercial beef cattle population; (ii) characterize profiles of ROH with different sizes and their gene contents with ROH hotspots; and (iii) identify associated ROH regions and loci for economical important traits.

Results

Genomic ROH distribution

We identified a total of 29271 ROHs in 462 Chinese commercial beef cattle population. ROH were identified with an average number of 63.36 segments and an average length of 0.98 Mb. We found most ROH lengths are relatively short, and most of ROH with lengths ranging from 0.5 Mb to 5 Mb, which accounting for 98.50% of total number, while the proportion of the genome covered by them was relatively small (Fig. 1a). The size of ROH varies from 500.02 kb to 48554.99 kb, we found the longest ROH was located on BTA10 (48554.99 kb, with 10015 SNPs), while the shortest ROH was located on BTA3 (500.02 kb, with 130 SNPs). In addition, we observed the number of ROH vary across chromosomes, while the highest number of ROH was observed on BTA1 (2463 ROH segments) and the minimum number of ROH was observed on BTA25 (Fig. 1b).

Roh Region And Inbreeding Coefficients

ROH regions were determined by merging ROH identified across all individuals using previously published protocols implemented in BEDTools [37]. In this study, we identified 280 nonredundant ROH regions in 462 individuals. Then, we estimated inbreeding coefficients of the commercial beef cattle population using two different approaches including F_HOM and F_ROH. The inbreeding coefficient of F_HOM varied from −0.24 to 0.27, and the values of F_ROH varied from 0.01 to 0.33. Our results also show that F_HOM and F_ROH were highly correlated (r = 0.96, P < 2.2e-16).

Genomic Patterns Of Homozygosity

To evaluate the ROH pattern of the commercial beef cattle population, we divided the length of ROH into three size classes: a. Small (500 kb to 1 Mb), b. Medium (1 Mb to 5 Mb), and c. Large (> 5 Mb), as described in the previous study [6]. The ROH distributions by total number and length for studied population were shown in Fig. 2a. We observed the total of 23390 Small (500 kb to 1 Mb) with length of 15.88 Mb, while the total of 440 Large (> 5 Mb) with length of 4.80 Mb. Meanwhile, our study revealed ~ 2.81% (70.30 Mb) of the genome is autozygous in the population. Also, we found the total lengths of ROH and the number of ROH were highly correlated (r = 0.91, P < 2.2e-16), and the number of ROH has increasing trend with the increase of the length of autosomes (Fig. 2b).

The Consensus Of Roh Across Population

To explore the occurrence of the consensus of ROH across population, the distribution of ROH frequencies for consensus ROH was estimated using PLINK command (-homozyg -group) across genome. We found the distribution of ROH frequencies among different chromosomes is uneven (Fig. 3). For instance, the highest frequency of ROH (> 36%) was observed in the middle part of BTA23, and we also found two ROH region with frequency of 28% and 25% which are located at BTA5 and BTA12, respectively. We totally detected 26 regions with ROH frequency exceeding 10% among studied population, and these regions was overlapped with 16 refgenes based on UMD 3.1. Moreover, based on cattle QTLdb, 35 QTLs were found that overlapped with 26 regions. Notably, we found the high frequency regions were overlapped with QTLs which are associated with weight gain (at BTA6) and calving difficulty (at BTA23). Moreover, it was noted that there are several important QTLs associated with several important traits including Gestation length (at BTA5), Lean meat yield (at BTA12)
and Stillbirth (at BTA23).

Roh Regions Associated With Phenotype

ROH frequently can be formed by selection pressure at specific positions on the genome, which may also indicate association with complex traits [38]. In our study, we totally identified 280 regions and performed association analysis using linear model for eight economical important traits. The summary statistics of ROH region were presented in (Additional file 1: Table S1). For body height, two significant ROHs were found at BTA23 and BTA7, which spanned around 1.2 and 3.6 Mb, these two ROH contain 34 and 56 candidate genes. Moreover, we identified several significant ROH at BTA8, BTA27, BTA12 and BTA28 for the chest circumference, the backfat thickness, the ribeye area and the fat coverage, respectively, which spanned a region 1.9 Mb, 0.6 Mb, 2.1 MB, 1 Mb (Additional file 2: Table S2). These identified regions contained 33, 22, 0, 4 candidate genes (P < 0.01). We generated genome wide ROH regions plot using Circos (http://circos.ca/) which illustrating ROH regions and association analysis results for eight traits (as shown in Fig. 4). For the body height and body length, we detected several regions on BTA23 with 1.2 Mb which contained 34 candidate genes. In addition, we also identified several significant ROH (with significant level P < 0.05) at BTA13, BTA14, BTA22, BTA23, BTA9 overlapping with 4, 62, 1, 34, 9 genes for body length. For slaughter weight, we also identified several significant ROH (P < 0.05) at BTA12, BTA14, BTA18, BTA8, which overlapped with 0, 7, 37, 107 candidate genes. Several candidate ROH regions and candidate genes for growth related traits using region based association analysis was presented in Table 2.

### Table 1
Summary of the identified ROHs and their overlapping candidate QTLs and related traits. The consensus ROH across population were sorted by the rate of occurrence of ROH, only ROH regions overlapped with QTLs (frequency exceeding 10%) were listed.

| BTA | ROH-BP1 | ROH-BP2 | Rate  | QTL ID                  | Traits                      |
|-----|---------|---------|-------|-------------------------|-----------------------------|
| 23  | 43444   | 1017051 | 0.36345 | 51867,51877,51879      | Calving difficulty, stillbirth |
| 5   | 58882448| 59642942| 0.275154 | 146791,146817          | Gestation length            |
| 12  | 41161447| 41983455| 0.250513 | 115982                  | Milk glycosylated kappa-casein |
| 6   | 80839528| 80859578| 0.162218 | 115098                  | Milk glycosylated kappa-casein |
| 6   | 33396376| 33893065| 0.13347 | 66375,66376,66377,66378,67264 | Body weight (birth), Body weight (yearling), Body weight gain |
| 4   | 45165956| 45671950| 0.119097 | 67016                   | Body weight gain            |
| 12  | 25997505| 26782472| 0.119097 | 37105                   | Lean meat yield             |
| 6   | 80611486| 80829499| 0.110883 | 42160,37399,37401       | Calving ease, Facial pigmentation, Eye area pigmentation |
Table 2
Summary of the candidate associated ROH regions and candidate genes for growth related traits using region based association analysis.

| BTA | ROH_left | ROH_right | Length | P-value | Traits               |
|-----|-----------|-----------|--------|---------|----------------------|
| 7   | 12040557  | 15624148  | 3583591| 0.00835 | Body heigh           |
| 8   | 75872649  | 77763089  | 1890440| 0.000273| Chest circumference  |
| 8   | 56872985  | 70416332  | 13543347| 0.017717| Slaughter weight     |
| 9   | 71225006  | 74479423  | 3254417 | 0.044654| Carcass length       |
| 9   | 19317920  | 21949240  | 2631320 | 0.091261| Body length          |
| 12  | 39693710  | 41835914  | 2142204 | 0.001376| Ribeye area          |
| 12  | 39693710  | 41835914  | 2142204 | 0.001376| Slaughter weight     |
| 13  | 15924488  | 17361646  | 1437158 | 0.047526| Body length          |
| 14  | 15437144  | 31577321  | 16140177| 0.023444| Body length          |
| 14  | 63411295  | 64996863  | 1585568 | 0.013867| Slaughter weight     |
| 18  | 61430317  | 63172836  | 1742519 | 0.018138| Slaughter weight     |
| 22  | 36579562  | 37111027  | 531465  | 0.03848  | Body length          |
| 23  | 7177811   | 8413638   | 1235827 | 0.000394| Body length          |
| 23  | 7177811   | 8413638   | 1235827 | 0.000171| Carcass length       |
| 23  | 7177811   | 8413638   | 1235827 | 0.015061| Body length          |
| 27  | 36930734  | 37557668  | 626934  | 0.000614| Backfat thickness    |
| 27  | 13977424  | 15257628  | 1280204 | 0.009542| Carcass length       |
| 27  | 40886907  | 41889675  | 1002768 | 0.001264| Fat coverage         |

Loci Associated With Phenotype

To precisely locate the candidate loci within ROH region associated for important traits, we further carried out association analysis for eight traits in the 462 commercial beef cattle population based each locus with homozygous states. Among the identified candidate ROH region, we totally identified 9360 loci for 8 traits and found 1,631 significant loci corresponding to 67 candidate genes. Among them, 37 loci were significantly associated (P < 0.01) with the fat coverage and only one candidate gene (EBF2) was found on BTA8 (Table 3). Also, 27 loci were significantly associated (P < 0.01) with the Slaughter weight, and two candidate genes ODF1 and UBR5 were identified at BTA14. However, no associated locus was detected for body length. In addition, a total of 109 significant loci (at P < 0.01) was found for chest circumference at BTA8 and BTA14, which contain 22 candidate genes (Additional file3: Table S3).

Table 3
Summary of the candidate associated loci and candidate genes for growth related traits using locus based association analysis.

| BTA | SNP1   | SNP2   | P-value | Traits     | Candidate genes |
|-----|--------|--------|---------|------------|-----------------|
| 8   | 74120149| 74120150| 4.07E-03| Fat coverage| EBF2            |
| 9   | 19552365| 19552366| 5.33E-03| Carcass length| SH3BGRIL2       |
| 14  | 64026130| 64026131| 7.00E-03| Slaughter weight| ODF1           |
| 14  | 64170699| 64170700| 7.00E-03| Slaughter weight| UBR5           |
| 23  | 8261693 | 8261694 | 1.13E-03| Carcass length| HMGA1           |
| 27  | 14223693| 14223694| 1.02E-03| Carcass length| ACSL1           |
| 27  | 36961090| 36961091| 1.24E-05| Backfat thickness| SLC20A2        |

Discussion

Many studies have explored the ROH pattern and inbreeding depression at the genomic level in multiple cattle...
populations [39–41]. To our knowledge, our study is the first attempt to present the occurrence and distribution of ROH in commercial beef cattle population using high-density SNP arrays. Previous study suggested that high-density SNP arrays are more sensitive to determine the small segments, while Bovine SNP50K arrays may underestimate the number of fragments with length of 1–4 Mb [42]. In present study, our results showed the most abundance of ROH segments were between 0.5 and 1 Mb, which implied the power of identification of small ROH using high density SNP array.

We estimated inbreeding coefficients in the commercial beef cattle population using two methods including $F_{ROH}$ and $F_{HOM}$. In general, the correlation between FROH and FHOM varies across many studies. Previous study found the correlation range 0.78 to 0.85, indicating that FROH has a high correlation with the FHOM, thus $F_{ROH}$ can be used as an accurate estimate of the proportion of IBD genomes [43]. In this study, the significant correlation ($p < 2.2e-16$) were observed between FROH and FHOM, which is consistent with previous report [44].

Many studies have shown significant differences in the total number and length of ROH in cattle, as well as the genetic relationship between individuals [7, 45]. The pattern of ROH count and length may indicate the differences of breed formation and recent breed management [7]. For instance, analysis of the British Isles breeds including Hereford, Guernsey, Angus and Jersey cattle displayed the highest sum of ROH events per animal compared other breeds, while African breeds displayed high variability in total number of ROH among breeds [7]. A recent study revealed that the average total length of ROH was 106 Mb and 371 Mb for Piedmontese cattle and Brown cattle, and the study also showed that ROHs identified in dairy breeds was longer and larger than beef and dual-purpose cattle. Consistent with previous findings, our results show ROH sizes range from 500 kb to several megabytes and also length and number of ROH are vary in commercial beef cattle population. Several studies suggest that short ROH reflect ancient inbreeding, and long ROH segments reveal recent inbreeding [31, 46, 47]. Therefore, our study revealed that most of ROH belong to the short and Medium, and this also indicate that the commercial population has low inbreeding level, which are agree with population history that this population have undergo hybrid process for recent selection.

Our study identified many consensuses of ROH across genome among population, and the distribution of ROH across genomic regions can imply functional effect for traits and help to identify functional pathways affected by inbreeding. In this study, we totally detected 26 regions with ROH frequency exceeding 10% among population, these regions were overlapped with 16 QTLs related to important traits including weight gain, calving difficulty, Gestation length, Lean meat yield and Stillbirth, which may potentially imply their impacts on reduction in performance for important trait [38].

Many previous studies have been investigated to detect the regions of homozygosity and their impact on complex traits in human [16, 19, 20, 48]. For instance, several methods have been utilized to explore the association between ROH region and complex disease [14, 45, 49, 50]. In current study, we proposed the region-based and loci-based approach to investigate the candidate the ROH region and loci associated with important traits in cattle. These approaches can also be extended to identify the loci of ROH related to important traits in other farm animals.

In present study, we firstly carried the association test based on ROH region. Totally, we obtained 280 regions by merging all individual ROH events among population, and then we utilized proportion of ROH coverage for each region among individual as variable. Using linear model, we identified 16 significant ROH regions, we found 371 candidate genes been located in nonredundant regions in the studied population. Based on the significant region, we obtained 153 genes, while other region with 218 genes. Since many genes were identified for the association analysis based on ROH region, therefore, it is difficult to pinpoint the region of ROH associated with studied trait.

Furthermore, we performed loci-based association to precisely locate the region of ROH for important trait. Totally, we found 1631 non redundant loci and 67 candidate genes for 8 traits. For the fat coverage, we found only one gene (EBF2) within 37 loci for fat coverage, this gene was found related to fat formation [51, 52]. Two genes (ODF1 and UBR5) with 4 loci and 23 loci were detected for slaughter Weight. One study found UBR5 gene is involved with glucose-dependent degradation of PEPCK1 [53]. For backfat thickness, 103 candidate loci within 9 candidate genes were identified in current study. However, only one gene (SLC20A2) was associated with
carcass trait [54]. For carcass length, 188 loci and 31 candidate genes were identified. Among them, there most promising candidate genes including SH3BGRL2, HMGA1 and ACSL1 have been reported from previous report. For instance, SH3BGRL2 gene had a significant effect on fatty acid metabolism [55], and HMGA1 gene was related to the growth, fertility and lean meat content in commercial pigs [56], and ACSL1 gene was a candidate gene for the function of fatty acid composition in bovine skeletal muscle [57]. For the body height, we found 24 candidate genes embedded with 119 loci, of these gene, HMGA1 and ACSL1 have been implied that can promote the growth and development of animal body [57, 58].

Conclusions

Our study investigated the homozygosity patterns and population inbreeding level in commercial beef cattle population. Both region-based and loci-based association were proposed and used to assess the association between ROH and important traits, which can be extended to explored genetic basis involved ROH for other farm animals. Our finding also implied some candidate ROH regions which may have potential impacts on reduction in performance for important trait in beef cattle.

Methods

Ethics Statement

No ethics statement was required for the collection of genetic material. The dataset from animals included in this study were from previous analyses that obtained specific permissions [36].

Genotyped Samples

The original data used were the Bovine HD chip from commercial beef cattle (Hybrid population from Wagyu and Fuzhou cattle), which was extracted from our pervious publication [36]. The population was established in Dalian XueLong Co. Ltd, Liaoning Province, China. Only autosomal SNP markers (735293) were analyzed for subsequently analysis. Then, individuals with missing SNPs exceeding 5% were removed. PLINK v1.07 [59] software was used to check the quality of dataset, and the standards were set as follows: minimum allele frequency (MAF) < 0.05, individual call rate < 10%, marker genotype deletion rate < 10%. After quality control, 503579 SNPs and 462 individuals were used for subsequently analyses.

Roh Estimation

In this study, we used PLINK v1.07 to detect ROH on the autosomes of each individual. Since LD can cause a short and common ROH throughout the genome, ROH was defined to be at least 0.5 MB in this study. Then, the specific parameters were as follows: 50 SNPs sliding windows slid along the chromosome to detect homozygous segments of each individual, and the sliding window allowed no more than 1 heterozygote. Several important parameters of defining ROH are involved: (i) the minimum length > 500 kb; (ii) the proportion of homozygous overlap window is 0.05; (iii) the minimum number of consecutive SNPs included in a ROH was 100; (iv) the minimum SNP density > 50 kb/SNP; (v) the maximum gap between continuous homozygous SNPs > 100 kb; (vi) the missing genotypes of ROH < 2.

Roh Classification And Inbreeding Coefficient

The length of ROH were divided into three classes: Small (0.5-1 Mb), Medium (1-5 Mb) and Large (> 5 Mb). Two methods were used to calculate inbreeding coefficient for each individual including: 1) $F_{HOM}$ was assessed based on the proportion of homozygotes using PLINK v1.07. 2) $F_{ROH}$ was calculated as described by McQuillan et al. (an individual’s summed ROH length was normalized by the length of the autosomal genome covered by SNPs) [60]. Correlations of inbreeding coefficient for three methods were estimated using cor.test function in R v3.2.4.

Qtl Regions And Genomic Regions Within Roh

To explore the potential QTLs loci related to economic traits and diseases in the commercial beef cattle, we
adopted approach plugged in PLINK v1.07 [59] (–homozyg-group) to estimate the consensus region across individuals, which represent ROH pools of overlapping and potentially matching segments. Then, based on these consensus regions, we annotated QTL and reference gene based on both QTLdb and UMD3.1 genome assembly when ROH frequencies exceeding 10%. The function annotation of the identified genes and GO terms were further assessed using DAVID platform [61, 62].

Region Association

ROH regions were generated using BEDtools and custom script [37]. The proportion of ROH coverage for each region among individual were considered as variable. The matrix represents the ratio of each ROH event to ROH regions across each individual. Raw phenotype was adjusted by the general linear model. The covariates included the weight of entry and fattening days, and the fixed effects included field, year, gender and group stratification. After correction of phenotypic data, eight growth traits including fat coverage, slaughter weight, backfat thickness, ribeye area, carcass length, body height, chest circumference and body length were utilized for subsequently analysis. Association analysis between ROH region and the adjusted phenotype were conducted using linear regression model, and candidate regions were considered statistically significant when $P < 0.01$. All statistical analyses were carried out using R programming (https://www.r-project.org/).

Loci Based Association

To further investigate the candidate loci within ROH for studied traits, we carried out the association test based the homozygous status of each locus. Variable for association analysis was generated based on the homozygous status of loci. For locus with the ROH for regions was coded as 1, while the locus was coded as 0 without the ROH. Also, association analysis was carried using the linear regression model between the state of loci (homozygous or non homozygous) and the adjusted phenotypes, candidate loci were determined by the statistically significant level ($P < 0.01$).

Abbreviations

SNP:Single nucleotide polymorphism; ROH:Runs of homozygosity; BTA:Bos Taurus autosome; MAF:Minor allele frequency; QTL:Quantitative trait loci; Linkage disequilibrium; IBD:Identity by descent; GO:Gene ontology; DAVID:The Database for Annotation, Visualization and Integrated Discovery.

Declarations

Ethics approval and consent to participate

No ethics approval was required for any aspect of this study.

Consent to publish

Not applicable

Availability of Data and Materials

The genotype data reported in this article are available upon request for research purpose.

Competing interests

The authors declare that they have no competing interests.

Funding

This study was supported by the Agricultural Science and Technology Innovation Program in Chinese Academy of Agricultural Sciences (CAAS-ZDXT2018006, ASTIP-IAS-TS-16 and ASTIP-IAS03) and National Beef Cattle Industrial Technology System (CARS-37) for the design of the study and data collection. The project was also partly supported by Beijing City Board of Education Foundation (PXM2016_014207_000012) for the data analysis and
interpretation of the study. L.Y.X was supported by the Elite Youth Program in Chinese Academy of Agricultural Sciences.

**Acknowledgement**

Not Applicable

**Authors’ contributions**

Conceived and designed the experiments: LYX, JYL and GYZ. Analyzed the data: GYZ, ZZW, TLZ, BZ and LX. Contributed reagents/materials/analysis tools: YQL, XG, HJG, LPZ and LX. Wrote the paper: LYX, and GYZ. All authors read and approved the final manuscript.

**References**

1. Nandolo W, Utsunomiya YT, Meszaros G, Wurzinger M, Khayadzadeh N, Torrecilha RBP, Mulindwa HA, Gondwe TN, Waldmann P, Ferencakovic M et al: Misidentification of runs of homozygosity islands in cattle caused by interference with copy number variation or large intermarker distances. *Genet Sel Evol* 2018, 50(1):43.

2. Jane G, Morton NE, Andrew C: Extended tracts of homozygosity in outbred human populations. *Human Molecular Genetics* 2006, 15(5):789-795.

3. Broman KW, Weber JL: Long homozygous chromosomal segments in reference families from the centre d’Etude du polymorphisme humain. *American Journal of Human Genetics* 1999, 65(6):1493-1500.

4. Falconer DS, Mackay TFC: *Introduction to quantitative genetics*; 1996.

5. Howrigan DP, Simonson MA, Keller MC: Detecting autozygosity through runs of homozygosity: a comparison of three autozygosity detection algorithms. *BMC genomics* 2011, 12(1):460.

6. Bosse M, Megens HJ, Madsen O, Paudel Y, Frantz LA, Schook LB, Crooijmans RP, Groenen MA: Regions of homozygosity in the porcine genome: consequence of demography and the recombination landscape. *PLoS Genet* 2012, 8(11):e1003100.

7. Purfield DC, Berry DP, McParland S, Bradley DG: Runs of homozygosity and population history in cattle. *BMC genetics* 2012, 13(1):70.

8. Herrero-Medrano JM, Megens H-J, Groenen MA, Ramis G, Bosse M, Pérez-Enciso M, Crooijmans RP: Conservation genomic analysis of domestic and wild pig populations from the Iberian Peninsula. *BMC genetics* 2013, 14(1):106.

9. Ferencakovic M, Hamzic E, Gredler B, Curik I, Sölkner J: Runs of homozygosity reveal genome-wide autozygosity in the Austrian Fleckvieh cattle. *Agriculturae Conspectus Scientificus* 2011, 76(4):325-329.

10. Ferencakovic M, Hamzic E, Gredler B, Solberg T, Klemetsdal G, Curik I, Sölkner J: Estimates of autozygosity derived from runs of homozygosity: empirical evidence from selected cattle populations. *Journal of Animal Breeding and Genetics* 2013, 130(4):286-293.

11. Kim E-S, Cole JB, Huson H, Wiggans GR, Van Tassell CP, Crooker BA, Liu G, Da Y, Sonstegard TS: Effect of artificial selection on runs of homozygosity in US Holstein cattle. *PLoS One* 2013, 8(11):e80813.

12. Zavarez LB, Utsunomiya YT, Carmo AS, Neves HH, Carvalheiro R, Ferenčaković M, Pérez O’Brien AM, Curik I, Cole JB, Van Tassell CP: Assessment of autozygosity in Nellore cows (Bos indicus) through high-density SNP genotypes. *Frontiers in genetics* 2015, 6(5).

13. Karimi S: Runs of homozygosity patterns in taurine and Indicine cattle breeds (master thesis). Vienna: BOKU-University of Natural Resources and Life Sciences 2013:53.

14. Zhang Q, Guldbrandtsen B, Bosse M, Lund MS, Sahana G: Runs of homozygosity and distribution of functional variants in the cattle genome. *BMC genomics* 2015, 16(1):542.

15. Bosse M, Megens HJ, Madsen O, Crooijmans RPMA, Cara MARD: Using genome-wide measures of coancestry to maintain diversity and fitness in endangered and domestic pig populations. *Genome Research* 2015, 25(7):970.

16. Lencz T, Lambert C, DeRosse P, Burdick KE, Morgan TV, Kane JM, Kucherlapati R, Malhotra AK: Runs of
homozygosity reveal highly penetrant recessive loci in schizophrenia. *Proc Natl Acad Sci U S A*, **104**(50):19942-19947.

17. Vine AE, McQuillin A, Bass NJ, Pereira A, Kandaswamy R, Robinson M, Lawrence J, Anjorin A, Sklar P, Gurling HMD: **No evidence for excess runs of homozygosity in bipolar disorder.** *Psychiatric Genetics* 2009.

18. Szpiech Zachary A, Xu J, Pemberton Trevor J, Peng W, Zöllner S, Rosenberg Noah A, Li Jun Z: **Long Runs of Homozygosity Are Enriched for Deleterious Variation.** 2013, **93**(1):90-102.

19. Yang HC, Lun-Ching C, Yu-Jen L, Chien-Hsing L, Pei-Li W, Stacey C: **A Genome-Wide Homozygosity Association Study Identifies Runs of Homozygosity Associated with Rheumatoid Arthritis in the Human Major Histocompatibility Complex.** *Plos One*, **7**(4):e34840-

20. Nalls MA, Guerreiro RJ, Simon-Sanchez J, Bras JT, Traynor BJ, Gibbs JR, Launer L, Hardy J, Singleton AB: **Extended tracts of homozygosity identify novel candidate genes associated with late-onset Alzheimer's disease.** **10**(3):183-190.

21. Ferencakovic, Hamzic M, Gredler E, Curik B, Sölkner J, Johann: **Runs of Homozygosity Reveal Genome-wide Autozygosity in the Austrian Fleckvieh Cattle.** *Agriculturae Conspectus Scientificus* 2011, **76**(4):325-329.

22. Keller MC, Visscher PM, Goddard ME: **Quantification of Inbreeding Due to Distant Ancestors and Its Detection Using Dense Single Nucleotide Polymorphism Data.** *Genetics* 2011, **189**(1):237-249.

23. Silió L, Rodríguez MC, Fernández A, Barragán C, Benítez R, óvilo C, Fernández Al: **Measuring inbreeding and inbreeding depression on pig growth from pedigree or SNP-derived metrics.** *Journal of Animal Breeding & Genetics*, **130**(5):n/a-n/a.

24. Mastrangelo S, Sardina MT, Tolone M, Di Gerlando R, Sutera AM, Fontanesi L, Portolano B: **Genome-wide identification of runs of homozygosity islands and associated genes in local dairy cattle breeds.** *Animal*:1-9.

25. Kim ES, Sonstegard TS, Tassell CPV, Wiggans G, Rothschild MF: **The Relationship between Runs of Homozygosity and Inbreeding in Jersey Cattle under Selection.** *Plos One* 2015, **10**(7):e0129967.

26. Mastrangelo S, Ciani E, Sardina MT, Sottile G, Portolano B: **Runs of homozygosity reveal genomewide autozygosity in Italian sheep breeds.** *Animal Genetics* 2018, **49**(1).

27. Upadhyay MR, Chen W, Lenstra JA, Goderie CRJ, MacHugh DE, Park SDE, Magee DA, Matassino D, Ciani F, Megens H-J: **Genetic origin, admixture and population history of aurochs (Bos primigenius) and primitive European cattle.** *Heredity* 2017.

28. Bjelland DW, Weigel KA, Vukasinovic N, Nkrumah JD: **Evaluation of inbreeding depression in Holstein cattle using whole-genome SNP markers and alternative measures of genomic inbreeding.** *Journal of Dairy Science*, **96**(7):4697-4706.

29. Gomez-Ray L, Rodriguez C, Barragán C, Silió L: **Genomic inbreeding coefficients based on the distribution of the length of runs of homozygosity in a closed line of Iberian pigs.** *Genetics Selection Evolution*, **47**(1):81.

30. Manunza A, Noce A, Serradilla JM, Goyache F, Martínéz A, Capote J, Delgado JV, Jordana J, Muoz E, Molina A: **A genome-wide perspective about the diversity and demographic history of seven Spanish goat breeds.** *Genetics Selection Evolution*, **48**(1):52.

31. Mirna K, Ruth M, S. FC, Harry C, M. MP, F. WJ, Manfred K: **Genomic Runs of Homozygosity Record Population History and Consanguinity.** *Plos One*, **5**(11):e13996-.

32. Reverter A, Porto-Neto LR, Fortes MRS, Kasarapu P, Lehnert SA: **Genomic inbreeding depression for climatic adaptation of tropical beef cattle.** *Journal of Animal Science* 2017, **95**(9):3809-3821.

33. Pryce JE, Haile-Mariam M, Goddard ME, Hayes BJ: **Identification of genomic regions associated with inbreeding depression in Holstein and Jersey dairy cattle.** *Genetics Selection Evolution*, **46**(1):71.

34. Metzger J, Karwath M, Tonda R, Beltran S, Agueda L, Gut M, Gut IG, Distl O: **Runs of homozygosity reveal signatures of positive selection for reproduction traits in breed and non-breed horses.** *BMC Genomics* 2015, **16**:764.

35. Genome-wide association study reveals candidate genes associated with body measurement traits in Chinese Wagyu beef cattle. *Animal Genetics* 2019.

36. Wang Z, Zhu B, Niu H, Zhang W, Xu L, Xu L, Chen Y, Zhang L, Gao X, Gao H: **Genome wide association study identifies SNPs associated with fatty acid composition in Chinese Wagyu cattle.** *Journal
of Animal Science and Biotechnology 2019.
37. Quinlan AR, Hall IM: BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 2010, 26(6):841-842.
38. Pryce JE, Haile-Mariam M, Goddard ME, Hayes BJ: Identification of genomic regions associated with inbreeding depression in Holstein and Jersey dairy cattle. Genet Sel Evol 2014, 46:71.
39. Peripolli E, Munari DP, Silva M, Lima ALF, Irgang R, Baldi F: Runs of homozygosity: current knowledge and applications in livestock. Anim Genet 2017, 48(3):255-271.
40. Bjelland D, Weigel K, Vukasinovic N, Nkrunum J: Evaluation of inbreeding depression in Holstein cattle using whole-genome SNP markers and alternative measures of genomic inbreeding. Journal of Dairy Science 2013, 96(7):4697-4706.
41. Ferencakovic M, Solkner J, Kaps M, Curik I: Genome-wide mapping and estimation of inbreeding depression of semen quality traits in a cattle population. J Dairy Sci 2017, 100(6):4721-4730.
42. Ferencakovic M, Solkner J, Curik I: Estimating autozygosity from high-throughput information: effects of SNP density and genotyping errors. Genet Sel Evol 2013, 45:42.
43. Peripolli E, Stafuzza NB, Munari DP, Lima ALF, Irgang R, Machado MA, Panetto J, Ventura RV, Baldi F, da Silva M: Assessment of runs of homozygosity islands and estimates of genomic inbreeding in Gyr (Bos indicus) dairy cattle. BMC Genomics 2018, 19(1):34.
44. Zhang Q, Calus MP, GulDBrandtsen B, Lund MS, Sahana G: Estimation of inbreeding using pedigree, 50k SNP chip genotypes and full sequence data in three cattle breeds. BMC Genetics 2015, 16(1):88.
45. Xu L, Zhao G, Yang L, Zhu B, Chen Y, Zhang L, Gao X, Gao H, Liu GE, Li J: Genomic Patterns of Homozygosity in Chinese Local Cattle. Sci Rep 2019, 9(1):16977.
46. Apollonio M, Randi E, TosO S: The systematics of the Wild Boar (Sus scrofa L.) in Italy.
47. Marras G, Gaspa G, Sorbolini S, Dimauro C, Ajmone-Marsan P, Valentini A, Williams Jl, Macciotta NPP: Analysis of runs of homozygosity and their relationship with inbreeding in five cattle breeds farmed in Italy. Animal Genetics, 46(2):110-121.
48. Ku CS, Naidoo N, Teo SM, Pawitan Y: Regions of homozygosity and their impact on complex diseases and traits. Human genetics 2011, 129(1):1-15.
49. Goszczynski D, Molina A, Terán E, Morales-Durand H, Ross P, Cheng H, Giovambattista G, Demyda-Peyrás S: Runs of homozygosity in a selected cattle population with extremely inbred bulls: descriptive and functional analyses revealed highly variable patterns. PloS one 2018, 13(7):e0200069.
50. Szmatola T, Gurgul A, Jasielczuk I, Zabek T, Ropka-Molik K, Litwinczuk Z, Bugno-Poniewierska M: A Comprehensive Analysis of Runs of Homozygosity of Eleven Cattle Breeds Representing Different Production Types. Animals (Basel) 2019, 9(12).
51. Shapira SN, Lim H-W, Rajakumari S, Sakers AP, Ishibashi J, Harms Mj, Won Kj, Seale P: EBF2 transcriptionally regulates brown adipogenesis via the histone reader DPF3 and the BAF chromatin remodeling complex. Genes & development 2017, 31(7):660-673.
52. Stine RR, Shapira SN, Lim Hw, Ishibashi J, Harms M, Won Kj, Seale P: EBF2 promotes the recruitment of beige adipocytes in white adipose tissue. Mol Metab 2016, 5(1):57-65.
53. Jiang W, Wang S, Xiao M, Lin Y, Zhou L, Lei Q, Xiong Y, Guan KL, Zhao S: Acetylation regulates gluconeogenesis by promoting PEPCK1 degradation via recruiting the UBR5 ubiquitin ligase. Mol Cell 2011, 43(1):33-44.
54. Abo-Ismail MK, Vander Voort G, Squires JJ, Swanson KC, Mandell IB, Liao X, Stothard P, Moore S, Plastow G, Miller SP: Single nucleotide polymorphisms for feed efficiency and performance in crossbred beef cattle. BMC Genet 2014, 15:14.
55. Zambrano MFB, Varona L, Glória LS, Lopes PS, Silva MVGB, Arbex W, Lázaro SF, Resende MDVd, Guimarães SFE: Genome association study through nonlinear mixed models revealed new candidate genes for pig growth curves. Scientia Agricola 2017, 74(1):1-7.
56. Yu S, Chung H, Sang B, Bhuiyan M, Yoon D, Kim K, Jeon J, Lee J: Molecular cloning and characterization of bovine HMGA1 gene. Asian-Australasian Journal of Animal Sciences 2007, 20(11):1662-1669.
57. Widmann P, Nuernberg K, Kuehn C, Weikard R: Association of an ACSL1 gene variant with polyunsaturated fatty acids in bovine skeletal muscle. BMC Genet 2011, 12:96.
58. McLean K, Schmutz S: Melanocortin 4 receptor polymorphism is associated with carcass fat in beef cattle. Canadian journal of animal science 2011, 91(1):75-79.

59. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ et al: PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007, 81(3):559-575.

60. McQuillan R, Leutenegger AL, Abdel-Rahman R, Franklin CS, Pericic M, Barac-Lauc L, Smolej-Narancic N, Janicijevic B, Polasek O, Tenesa A et al: Runs of homozygosity in European populations. American journal of human genetics 2008, 83(3):359-372.

61. Huang da W, Sherman BT, Lempicki RA: Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature protocols 2009, 4(1):44-57.

62. Huang da W, Sherman BT, Lempicki RA: Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic acids research 2009, 37(1):1-13.

Supplementary Files Legend

Table S1. Summary of summary statistics of ROH region including chromosome, start, end and frequency of population, average of length of each ROH, and average of SNP count within each region. The ROH region was generated were produced by aggregating overlapping CNVs (by at least 1bp) across samples using BEDTools,

Table S2. Summary of the all candidate associated loci and candidate genes for growth related traits using locus based association analysis. Candidate loci were determined by the statistically significant level (P<0.01).

Table S3. Summary of the candidate associated ROH regions and their candidate genes list for growth related traits using region based association analysis in commercial beef cattle population.

Figures
Figure 1

(a) The distribution of lengths ROH in commercial beef cattle population. (b) The distribution of total number of ROH across chromosomes.
Figure 2

(a) The total length of ROH belonging to three size classes including Small (0.5 to 1 Mb), Medium (1 to 5 Mb) and Large (>5 Mb) size. (b) Evaluation of number of ROH and ROH total length. The number of ROH found for each individual genome (y-axis) is plotted against ROH total length (i.e. the length of Mb covered by ROH in each genome, x-axis).
Figure 3

Distribution characteristics of ROH on chromosome in commercial beef cattle population. The horizontal axis is SNPs, which are ordered by the physical location of the genome; the vertical axis is the sample proportion with ROH in this SNP. The horizontal line in the graph is where the ROH frequency exceeds 10%.
Figure 4

Circos plot illustrating ROH regions and association analysis for eight traits. Association significance test for each of ROH are plotted based on -log10(P) value using histograms plot within the gray inner circle. ROH frequency indicated by occurrence of ROH with each region across total population. Frequency values are shown in the inner circle using circus package. The outermost circle displays the cattle auto chromosomes. The circles from inside to outside represent association result for body height, body length, chest circumference, fat coverage, backfat thickness, ribeye area, carcass length and slaughter weight, respectively.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additional file2 Table S2 Region.xlsx
- Additional file3 Table S3 loci.xlsx
- Additional file1 Table S1 -region_p gene-0302.xlsx