Genome-Wide Association Study Using Whole-Genome Sequencing Identifies a Genomic Region on Chromosome 6 Associated With Comb Traits in Nandan-Yao Chicken

Zhuliang Yang[1], Leqin Zou[1], Tiantian Sun[1], Wenwen Xu[1], Linghu Zeng[1], Yinhai Jia[2], Jianping Jiang[3], Jixian Deng[2] and Xiurong Yang[1]*

[1] College of Animal Science and Technology, Guangxi University, Nanning, China, [2] Guangxi Institute of Animal Science, Nanning, China, [3] Guangxi Botanical Garden of Medicinal Plants, Nanning, China

Comb traits have potential economic value in the breeding of indigenous chickens in China. Identifying and understanding relevant molecular markers for comb traits can be beneficial for genetic improvement. The purpose of this study was to utilize genome-wide association studies (GWAS) to detect promising loci and candidate genes related to comb traits, namely, comb thickness (CT), comb weight (CW), comb height, comb length (CL), and comb area. Genome-wide single-nucleotide polymorphisms (SNPs) and small insertions/deletions (INDELs) in 300 Nandan-Yao chickens were detected using whole-genome sequencing. In total, we identified 134 SNPs and 25 INDELs that were strongly associated with the five comb traits. A remarkable region spanning from 29.6 to 31.4 Mb on chromosome 6 was found to be significantly associated with comb traits in both SNP- and INDEL-based GWAS. In this region, two lead SNPs (6:30,354,876 for CW and CT and 6:30,264,318 for CL) and one lead INDEL (a deletion from 30,376,404 to 30,376,405 bp for CL and CT) were identified. Additionally, two genes were identified as potential candidates for comb development. The nearby gene fibroblast growth factor receptor 2 (*FGFR2*)—associated with epithelial cell migration and proliferation—and the gene cytochrome b5 reductase 2 (*CYB5R2*)—identified on chromosome 5 from INDEL-based GWAS—are significantly correlated with collagen maturation. The findings of this study could provide promising genes and biomarkers to accelerate genetic improvement of comb development based on molecular marker-assisted breeding in Nandan-Yao chickens.

Keywords: genome-wide association study, Nandan-Yao chicken, whole-genome sequencing, comb traits, SNPs, INDELs

INTRODUCTION

The Nandan-Yao chicken, with a single comb, is a typical indigenous breed in Guangxi Zhuang Autonomous Region (Yang et al., 2020). To meet the particular demand of local customers, chickens with a prominent ornament are preferred breeding targets for the local poultry industry.

The comb is a sexual ornament in chickens, and its morphological characteristics, such as comb shape and size, are important in mating behavior in both sexes (Siegel and Dudley, 1963;

Zuk et al., 1990; Pizzari et al., 2003). Males prefer to invest sperm in females with relatively large comb (Cornwallis and Birkhead, 2006); reciprocally, females preferentially like to copulate with dominant males with large combs (Graves et al., 1985; Zuk et al., 1990). Additionally, the comb is associated with body temperature regulation (van Kampen, 1971), egg production (Wright et al., 2008), and fecundity and sex maturity (Wright et al., 2012). From the above aspects, it can be concluded that the comb has potential economic value in poultry breeding.

Since the chicken genome was published (Wallis et al., 2004), researchers have made full use of this genomic data to identify the genetic differences under the phenotypic variation. Great progress has been made in understanding the genetic associations with some traits, such as yellow skin (Eriksson et al., 2008), silky-feather (Feng et al., 2014), and muffs and beard phenotypes (Guo et al., 2016). Also, numerous studies have explored the genetic mechanism for comb traits, for instance, the identification of variants and genes responsible for comb shape (Rose-comb and Pea-comb), comb mass (Johnsson et al., 2012), and comb color (Dong et al., 2019). However, traits related to comb size, such as comb length (CL) and comb height (CH), are less considered in genetic research.

The genome-wide association study (GWAS) is a powerful method for detecting causal variants of complex traits and has been widely used in livestock for the past decades. Many studies have successfully used GWAS for poultry breeding to identify quantitative trait loci (QTLs) for numerous desired traits, namely, bone size (Guo et al., 2020), internal organ traits (Moreira et al., 2019), and egg production (Liu et al., 2019). Because of the high cost of sequencing large samples, most of these studies were carried out using SNP chips. However, a limitation of SNP chips is that they can only detect pre-existing genetic variants. Compared with SNP chips, whole-genome sequencing (WGS) is more efficient and powerful in detecting and fine-mapping rare variants (Tam et al., 2019). With the decrease in sequencing costs, it is now more beneficial to explore genomic information using WGS.

In this study, to explore the potential genomic regions and candidate genes associated with size-related comb traits in Nandan-Yao chickens, we employed WGS to detect genome-wide genetic variation in a population of 300 Nandan-Yao chickens and performed a genome-wide association analysis on five comb traits.

MATERIALS AND METHODS

Sample Collection and Phenotypic Measurement

The experimental animals used in this study were Nandan-Yao chickens. A total of 300 animals (149 males and 151 females) were randomly collected from Gangfeng Agriculture and Husbandry Co., Ltd., Guigang. All birds were housed in cages. The animals were slaughtered after 120 days, and the combs were cut off for phenotypic measurements. We collected five phenotypes in total, which were comb area (CA), comb thickness (CT), CL, CH, and comb weight (CW), respectively. CT was measured at the thickest

point with a Vernier caliper, and CW was measured using a scale. For measurement of CA, CH, and CL, we delineated the contour of the comb on A4 paper along the comb's edge and scanned the comb profile into the digital photo by the scanner; we then used the Photoshop Measurement feature to measure CA, CH, and CL. CH was measured from the point at which the comb met the head to the top of the highest spike, and CL was defined from end to end (Navara et al., 2012). To reduce the false positives caused by outliers in the association analysis, the records over triple standard deviation were removed. The phenotypic correlations among five comb traits were calculated using the Pearson method by "psych," a package in R (version 3.6.2).

Genomic Analysis

The genomic DNA was isolated from blood using the standard phenol–chloroform method, and the WGS data were generated in Novogene Bioinformatics Technology Co., Ltd., by Illumina PE150 platform. We selected the Gallus_gallus-5.0 (GCA_000002315.3) as the reference genome (Zerbino et al., 2018) and utilized BWA (version 0.7.8) for genome alignment (Li and Durbin, 2009) with parameters "mem -t 4 -k 32 -M." Subsequently, the "mpileup" module from SAMtools (version 1.3.1) was used to detect genome-wide genetic variation, namely, single-nucleotide polymorphisms (SNPs) and small insertions/deletions (INDELs; Li et al., 2009). To obtain reliable SNPs, we further filtered the data with the following thresholds: the average depth per SNP > 5, minor allele frequency > 0.05, missing rate per individual < 0.1, missing rate per SNP < 0.1, and p-value of Hardy–Weinberg Equilibrium $> 10^{-6}$. The filtered process was performed using VCFtools (version 0.1.17) (Danecek et al., 2011) and PLINK (version 1.90) (Purcell et al., 2007). The remaining SNPs with missing genotypes were imputed using BIMBAM (version 1.0) software (Scheet and Stephens, 2006). Finally, all eligible SNPs derived from autosomes (GGA) were used for GWAS analysis. The filter parameters of INDELs were similar to those of SNPs but without imputation.

Genome-Wide Association Studies

We performed a principal component analysis (PCA) using all autosomal SNPs before the association tests to estimate population stratification (Price et al., 2006). The top three PCs accounted for 2.28% of the total variance, suggesting that the differentiation of experimental chickens was not obvious. Besides, considering the effects of hormones on combs (Eitan et al., 1998; Symeon et al., 2012), we both fit the top three PCs and sex as covariates while executing association analysis. All association tests were carried out using GEMMA software (version 0.96) (Zhou and Stephens, 2012). The univariate mixed linear model was as follows:

$$y = W\alpha + x\beta + u + e$$

where y is a vector of phenotypic values; W is a matrix of covariates (fixed effects) including sex and the first three PCs; α is a vector of the corresponding coefficients for fixed effects; x is a vector of genotypes; β is the effect size of the maker; u is a vector of random effects with a covariance structure as $u \sim N(0, G\sigma_u^2)$, where G is the genetic relatedness matrix calculated from all SNP

Trait[a]	Number[b]	Males					Females				
		Min	Max	Mean	SD	CV (%)	Min	Max	Mean	SD	CV (%)
CT (mm)	300	10.20	24.23	17.42	2.80	16.06	3.14	12.05	6.04	1.85	30.63
CW (g)	298	4.26	32.40	15.48	5.71	36.90	0.48	6.92	2.20	1.10	49.86
CA (cm^2)	286	10.47	58.55	32.20	9.52	29.58	1.01	17.42	6.98	2.89	41.35
CL (cm)	280	5.55	12.40	9.09	1.21	13.31	2.69	7.47	4.75	0.84	17.63
CH (cm)	280	2.67	7.07	5.05	0.82	16.27	0.90	3.50	2.28	0.52	22.78

[a]CT, comb thickness; CW, comb weight; CA, comb area; CL, comb length; CH, comb height.
[b]The remaining measurements after removing over triple standard deviation.

markers and σ_u^2 is the polygenic additive variance; e is a vector of residual errors with e $\sim$ N(0, Iσ_e^2), where I is the identity matrix and σ_e^2 is the residual variance.

For SNP-based GWAS, we selected the simpleM method (Gao et al., 2008; Gao, 2011) to adjust the threshold of the genome-wide significance p-value. After the simpleM test, a total of 6,971,926 effective independent tests were obtained. Hence, the significant and suggestive thresholds for SNP-based GWAS were set at 7.17 $\times$ 10^{-9} (0.05/6,971,926) and 1.43 $\times$ 10^{-7} (1/6,971,926), respectively. The SNPs that reached the suggestive genome-wide threshold would be annotated using ANNOVAR software (Wang et al., 2010).

We also carried out the INDEL-based GWAS for comb traits using the same analytical process as that used for SNP-based GWAS. The significant and suggestive thresholds for INDEL-based GWAS were set at 7.31 $\times$ 10^{-8} (0.05/684,401) and 1.46 $\times$ 10^{-6} (1/684,401), respectively, where 684,401 is the number of INDELs we used for GWAS. Finally, we used the "vt" tool to make INDELs normalized (Tan et al., 2015) for annotation.

The Manhattan and quantile–quantile plots for GWAS results were performed using the "CMplot" package,[1] and the diagram for regional plot on GGA 6 was created by "karyoploteR" package (Gel and Serra, 2017) in the R (version 3.6.2). The linkage disequilibrium (LD) correlation (r^2) between the associated SNPs (genome-wide suggestive and significant loci on GGA 6) was estimated using PLINK and visualized by LDBlockShow release 1.35.[2] The conditional GWAS for lead SNP was performed by GCTA software (version 1.26) (Yang et al., 2011).

RESULTS

Phenotypic Statistics

Descriptive statistics of the five comb traits are presented in **Table 1**, namely, minimum, maximum, mean, standard deviation, and coefficient of variation (CV). We observed that CW had the highest CV in both sexes (36.90% for males and 48.96% for females), followed by CA (29.58% for males and 41.35% for females). The phenotypic correlations for the five

[1]https://github.com/YinLiLin/CMplot
[2]https://github.com/BGI-shenzhen/LDBlockShow

comb traits ranged from 0.93 to 0.97, while CW and CA had the strongest correlation (**Supplementary Figure 1**).

Summary of Sequencing and Population Stratification

We obtained approximately 2,999 Gb of clean data from 300 Nandan-Yao chickens. The average alignment rate and depth were 98.75% and 8.38×, respectively. The coverage of at least 1× per sample was 82.7–93.49%. Detailed information on sequencing is shown in **Supplementary Table 1**, and the PCA plot is presented in **Supplementary Figure 1**.

SNP-Based GWAS

After filtering the SNPs, a total of 9,080,580 autosomal SNPs and 299 Nandan-Yao chickens were left for SNP-based GWAS. The distribution of all eligible SNPs on GGA 1–33 is shown in **Supplementary Figure 1**. The univariate GWAS was performed on CT, CW, CA, CL, and CH, respectively. The results showed that a prominent genomic region spanning from 29.6 to 31.4 Mb on GGA 6 was associated with three comb traits, namely, CT, CW, and CL (**Figure 1**). Precisely, in this region, we identified 132 SNPs suggestively associated with CL (including 47 significant SNPs), 39 SNPs suggestively associated with CT (including 3 significant SNPs), and 3 SNPs suggestively associated with CW, respectively (**Table 2**). The SNPs that reached the suggestive threshold for CL contained all associated SNPs identified for CT and CW (**Supplementary Figure 2**). Besides, an SNP located on GGA 4 (**Figure 1**) was found to be significantly associated with CH ($-\log_{10}(p)$ = 8.22). Unfortunately, no SNP reached the suggestive threshold for CA.

As shown in **Table 2**, the same lead SNP located at 30,354,876 bp on GGA 6 was observed in CW and CT. Meanwhile, the lead SNP for CL was located at 30,264,318 bp on GGA 6. These two loci were both located in the intergenic region between the *MIR7472* and *MCMBP* genes. Interestingly, the two SNPs had a low LD correlation (r^2 = 0.3817). For CH, the most significantly associated SNP (4:89,933,310) was located within an intron of the *GFRA4* gene. The detailed information of all associated SNPs is provided in **Supplementary Table 2**.

INDEL-Based GWAS

We carried out the INDEL-based GWAS for five comb traits with 684,401 INDELs and identified 25 associated INDELs (**Table 3**).
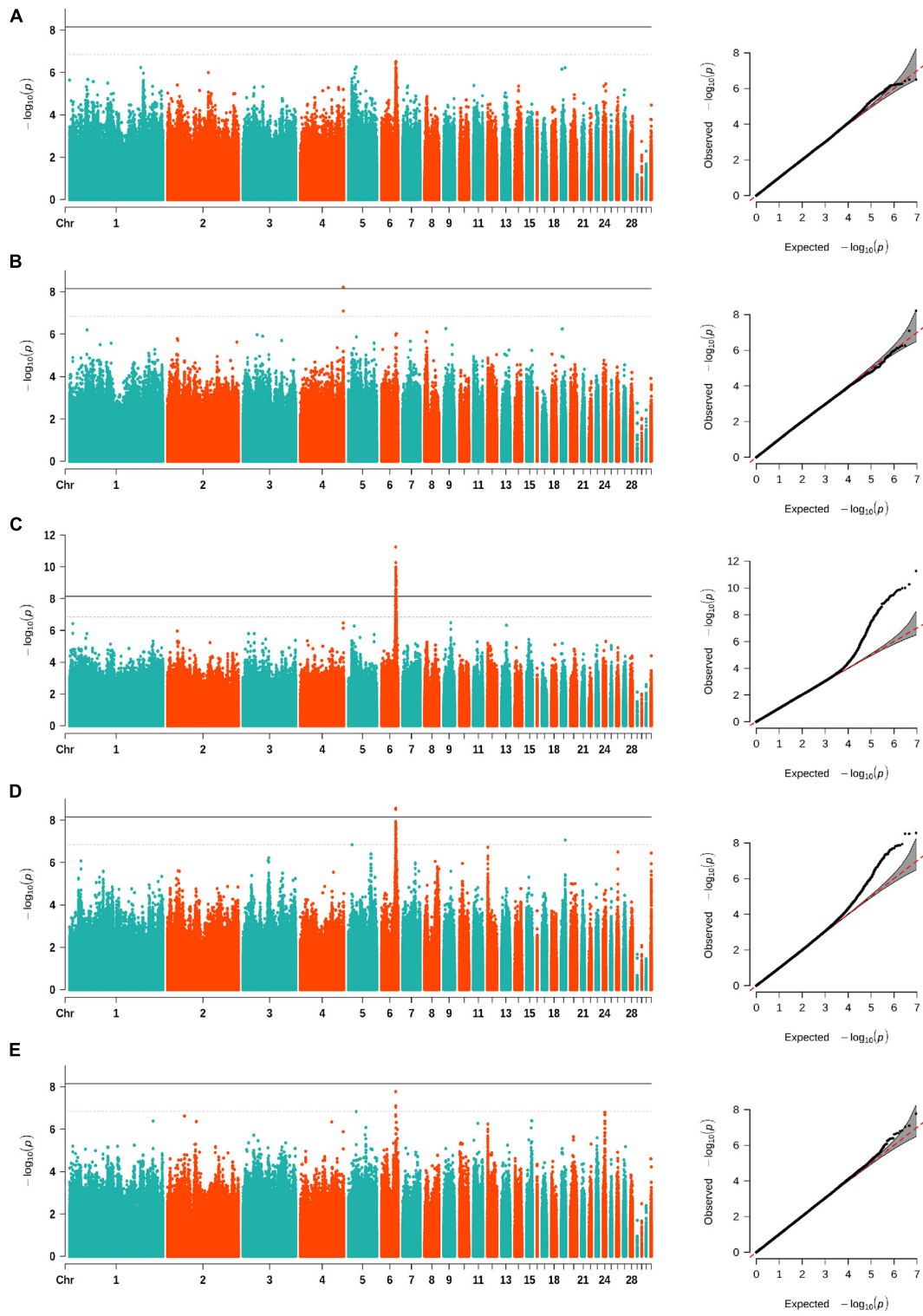
FIGURE 1 | Manhattan (left) and quantile–quantile (Q–Q) (right) plots of SNP-based GWAS for comb area **(A)**, comb height **(B)**, comb length **(C)**, comb thickness **(D)**, and comb weight **(E)**. The Manhattan plots show $-\log_{10}$(observed p) for SNPs (y-axis) against their corresponding positions on each chromosome (x-axis), while the Q–Q plots show the observed and expected $-\log_{10}(p)$. The horizontal dashed and solid lines in Manhattan plots represent the genome-wide suggestive ($-\log_{10}(p) = 6.84$) and significant ($-\log_{10}(p) = 8.14$) thresholds, respectively.

TABLE 2 | A descriptive summary of associated SNPs on GGA 6.

Chromosome	Position	N_Sugg[a]	N_Sig[b]	Lead variant	−log$_{10}$(p)[c]	Beta estimates[d]	Trait[e]
6	30,237,969–30,946,600	39	3	6:30,354,876	8.56	−1.96	CT
6	30,352,191–30,354,876	3	0	6:30,354,876	7.77	−3.25	CW
6	29,620,815–31,440,202	132	47	6:30,264,318	11.26	−1.22	CL

[a] The number of SNPs with −log$_{10}$(p) > 6.84.
[b] The number of significant SNPs with −log$_{10}$(p) > 8.14.
[c] The −log$_{10}$(p) of lead variant.
[d] The beta estimates of lead SNP.
[e] CT, comb thickness; CW, comb weight; CL, comb length.

TABLE 3 | Genes linked to the 25 associated INDELs for comb traits in Nandan-Yao chickens.

Chromosome	Position[a] (bp)	Ref[b]	Alt[c]	Trait[d]	−log$_{10}$(p)	Annotation	Nearby gene
1	154,683,824	–	A	CA	6.28	Intergenic	*ACOD1*; *UCHL3*
1	154,805,766	–	CAAGGAGG	CA	6.84	Intergenic	*ACOD1*; *UCHL3*
2	127,710,350	–	C	CW	6.07	Intergenic	*CPQ*; *RPL30*
5	7,403,641–7,403,642	TC	–	CW, CL, CA	6.13–7.58	Intronic	*CYB5R2*
5	13,977,826	A	–	CL, CA	6.44–7.83	Intergenic	*IGF2*; *MRPL23*
5	17,679,480	–	T	CW	7.09	Intergenic	*FGF3*; *CTTN*
6	29,621,960–29,621,961	CA	–	CL	5.99	Intergenic	*VAX1*; *PRLHR*
6	29,923,004–29,923,010	CCCCCTC	–	CL	5.84	Intergenic	*CACUL1*; *FAM45A*
6	29,944,104	–	A	CL	6.08	Intergenic	*CACUL1*; *FAM45A*
6	29,948,947	–	AT	CL	6.65	Intergenic	*CACUL1*; *FAM45A*
6	29,961,899–29,961,902	AAGG	–	CL	6.28	Intergenic	*CACUL1*; *FAM45A*
6	30,237,444	A	–	CT, CL	6.49–8.07	Intergenic	*MIR7472*; *MCMBP*
6	30,271,529	C	–	CA, CT, CL	5.84–8.20	Intergenic	*MIR7472*; *MCMBP*
6	30,310,013	–	G	CT, CL	6.88–8.19	Intergenic	*MIR7472*; *MCMBP*
6	30,376,404–30,376,405	TA	–	CT, CL	7.85–8.48	Intergenic	*MIR7472*; *MCMBP*
6	30,571,297	–	T	CT, CL	6.81–7.11	Intergenic	*MCMBP*; *FGFR2*
6	30,878,266	T	–	CL	6.1	Intergenic	*MCMBP*; *FGFR2*
6	30,889,738	C	–	CL	8.07	Intergenic	*MCMBP*; *FGFR2*
6	30,911,206–30,911,207	AT	–	CT, CL	6.33–8.23	Intronic	*FGFR2*
6	30,911307	–	CATT	CL	7.98	Intronic	*FGFR2*
6	30,943,119	–	TGTG	CL	6.41	Intronic	*FGFR2*
6	31,073,199–31,073,201	TTG	–	CL	8.39	Intronic	*ATE1*
6	31,083,785	G	–	CL	6.46	Intronic	*ATE1*
6	31,088,999	–	G	CL	7.18	Intronic	*ATE1*
6	31,440,157	–	C	CL	6.50	Intergenic	*ATE1*; *IKZF5*

[a] The normalized position of INDELs.
[b,c] Insertion or deletion.
[d] CT, comb thickness; CW, comb weight; CA, comb area; CL, comb length; CH, comb height.

Interestingly, we found that the same region on GGA 6 was strongly associated with CL, CA, and CT (**Figure 2**), which implicated 19 INDELs (**Supplementary Figure 2**). The most significantly associated variant was a 2-bp deletion, located from 30,376,404 to 30,376,405 bp in the intergenic region between the *MIR7472* and *MCMBP* genes. We also annotated several genes from the associated INDELs on GGA 1, 2, and 5, namely, *ACOD1*, *UCHL3*, *CPQ*, *CYB5R2*, *IGF2*, *MRPL23*, and *FGF3* (**Table 3**).

Analysis of LD and Conditional GWAS

We performed LD analysis on prominent signals for SNP-based GWAS; the LD correlation for all suggestive loci on GGA 6 from 29.6 to 31.4 Mb is shown in **Figure 3**. Unexpectedly,
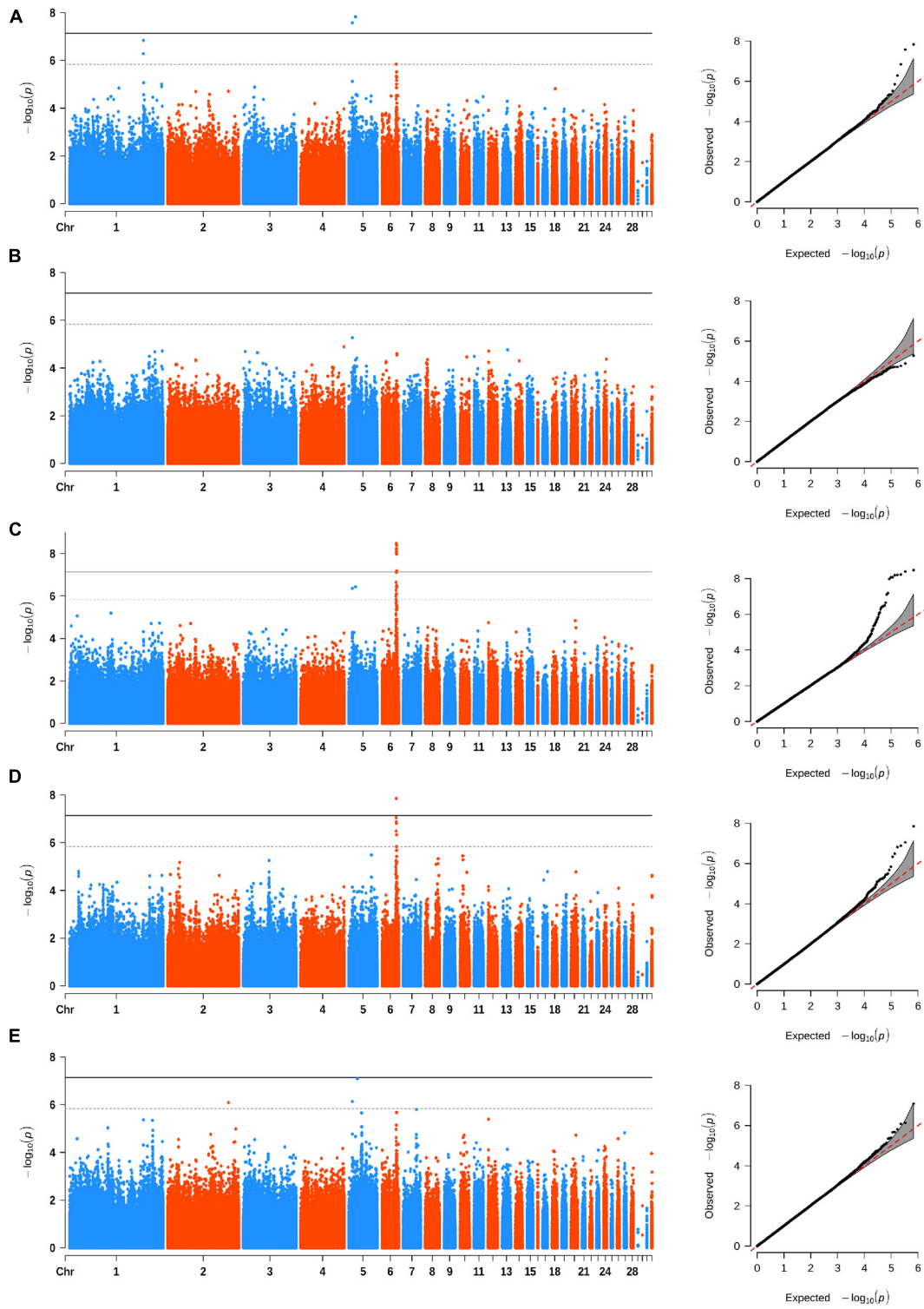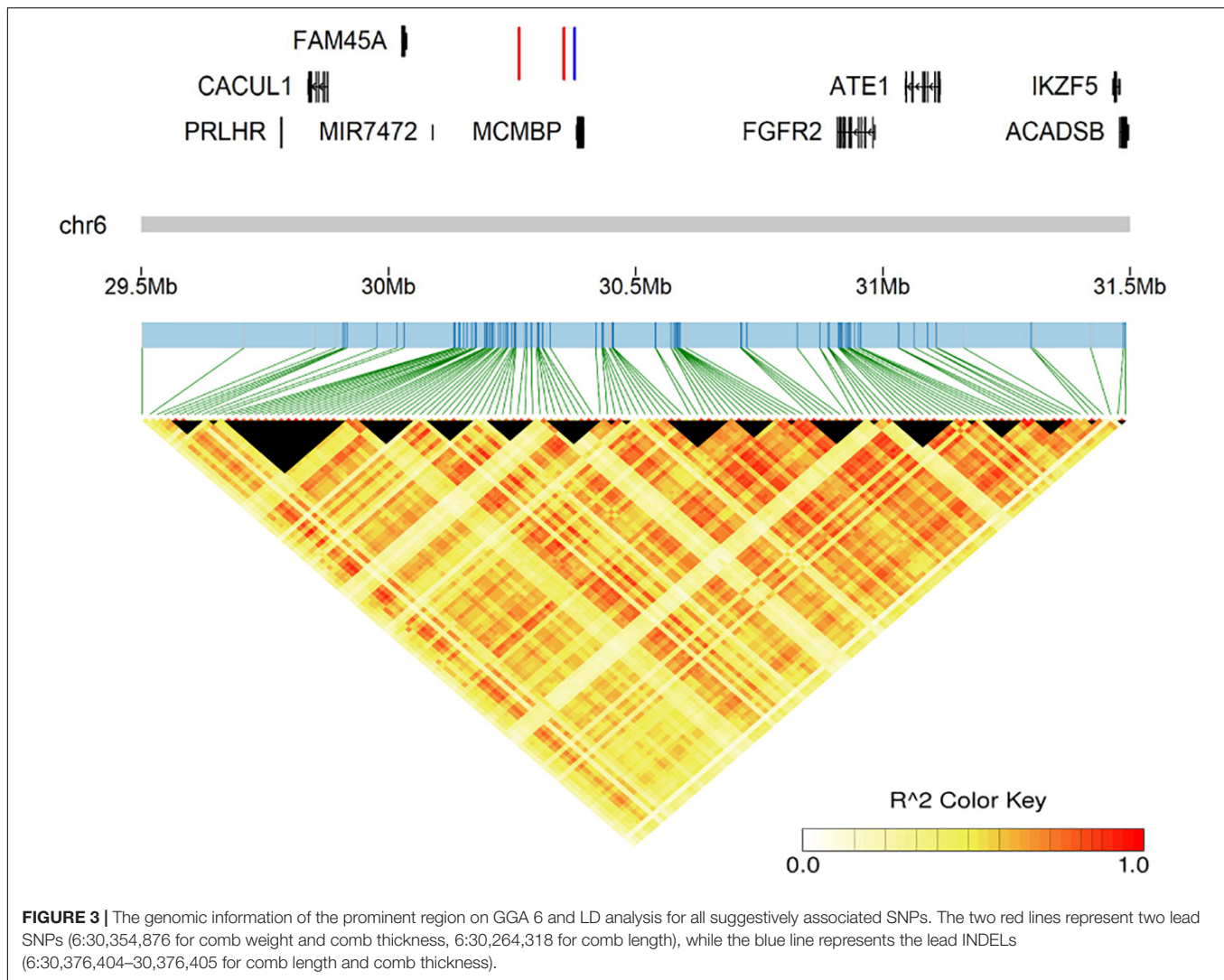
FIGURE 2 | Manhattan (left) and quantile–quantile (Q–Q) (right) plots of INDEL-based GWAS for comb area **(A)**, comb height **(B)**, comb length **(C)**, comb thickness **(D)**, and comb weight **(E)**. The Manhattan plots indicate $-\log_{10}$(observed p) for INDELs (y-axis) against their corresponding positions on each chromosome (x-axis), while the Q–Q plots show the observed and expected $-\log_{10}(p)$. The horizontal dashed and solid lines in Manhattan plots represent the genome-wide suggestive ($-\log_{10}(p)$ = 5.84) and significant ($-\log_{10}(p)$ = 7.14) thresholds, respectively.

FIGURE 3 | The genomic information of the prominent region on GGA 6 and LD analysis for all suggestively associated SNPs. The two red lines represent two lead SNPs (6:30,354,876 for comb weight and comb thickness, 6:30,264,318 for comb length), while the blue line represents the lead INDELs (6:30,376,404–30,376,405 for comb length and comb thickness).

we found that the two lead SNPs (6:30,354,876 for CT and CW, 6:30,264,318 for CL) had a low level of LD correlation ($r^2 = 0.3817$). Therefore, to test if the independently associated SNPs existed in the signal region for relevant comb traits, we subsequently carried out an approximately conditional GWAS based on the lead SNPs for CT, CW and CL. As the results showed, the level of significant and suggestive loci around the lead SNPs was all decreased below the suggestive genome-wide threshold ($-\log_{10}(p) = 6.84$) after conditional analysis (**Supplementary Figure 3**), indicating that no multiple independently associated variants existed in the signal region on GGA 6.

DISCUSSION

To date, relatively few GWAS have been performed on comb traits, among which only two studies are established on the quantitative characters of combs. One was conducted on CL, CH, and CW using 600 K SNP arrays (Shen et al., 2016) and

the other on CA using a 60 K SNP chip (Lien et al., 2017). Compared with these two studies, the magnitude of SNPs used in our study was much larger. Furthermore, our association tests were conducted on genome-wide INDELs, which are rarely used in GWAS research for chickens. However, a limitation is that our sample size was small for GWAS, which led us to remove variants with low frequency. Further GWAS with a larger sample size might result in the identification of additional variants in the future. Overall, the WGS data generated in this study can be used in conjunction with subsequent studies to provide more useful information.

Benefits for Local Poultry Industries

The Nandan-Yao chicken is one of the most popular indigenous breeds in the Guangxi Zhuang Autonomous Region. However, local poultry industries mostly breed Nandan-Yao chickens based on phenotypes, resulting in slow genetic progress. The appropriate use of molecular marker-assisted breeding can accelerate the selection of Nandan-Yao chickens to obtain the

desired phenotypes, allowing poultry companies to increase profits more quickly. In our study, we identified 134 SNPs and 25 INDELs that were strongly associated with comb size, which could be helpful in the selection of Nandan-Yao chickens. For example, the lead SNP (6:30,354,876, rs737686019) identified from CT and CW could be considered a priority for selection. From the SNP-based GWAS, we could find that the C allele of rs737686019 had negative effects on CW and CT (**Table 2**). Therefore, by removing animals with the C allele, the comb size of the population might increase modestly. This would be much faster than screening based on the phenotypes directly. In addition, due to the associated SNPs on GGA 6 are highly correlated, breeders could choose other SNPs as candidate markers for comb traits by selecting the loci with more obvious effects.

Candidate Genes Associated With Comb Traits

Two genes were considered as candidate genes for comb development, one is *FGFR2* (fibroblast growth factor receptor 2) and the other is *CYB5R2* (cytochrome b5 reductase 2). For *FGFR2*, there were several reasons: (1) we identified this gene from both SNP- and INDEL-based GWAS, suggesting that it was highly associated with comb traits; (2) it is known that the comb comprises three layers, namely, epidermis, dermis, and central connective tissue (Nakano et al., 1996); meanwhile, a previous study reported that *FGFR2* can be used as a gene marker for epithelial cell migration and proliferation (Carre et al., 2018); therefore, we speculated that *FGFR2* might be involved in the structural formation of the comb; (3) some studies have reported that, in human, a mutation in the 2c splice variant of the *FGFR2* gene results in Crouzon syndrome (Reardon et al., 1994), as well as regulates coronal suture development (Pfaff et al., 2016); additionally, in chicken, a previous study (Wilke et al., 1997) showed that *FGFR2* transcripts were expressed throughout the head mesenchyme before budding out of facial prominences (from Hamburger–Hamilton stage 9 to 17 of chick embryo) and expressed throughout the frontonasal mass during early facial prominence formation (Hamburger–Hamilton stage 20); collectively, these studies indicate that *FGFR2* is critical for facial or craniofacial development, which could further indicate an indirect influence on comb development. Indeed, the receptor protein encoded by *FGFR2* is a receptor tyrosine kinase (RTK), suggesting that *FGFR2* may be involved in comb development through the RTK/Ras/MAPK signaling pathway. However, this inference needs to be supported by additional evidence. For *CYB5R2* gene, we identified it from a deletion (5:7,403,641–7,403,642) in INDEL-based GWAS. It is a member of flavoprotein pyridine nucleotide cytochrome reductase family and is involved in many physiological processes (O'Leary et al., 2016). Collagen is known to be one of the major chemical compositions of comb (Nakano et al., 1996); meanwhile, a previous study reported that the expression of *CYB5R2* was significantly correlated with collagen maturation (Liu et al., 2015). Thus, we speculate that *CYB5R2* has a potential role in comb development. In addition, the region that contains *CYB5R2* on GGA 5 has been reported

to be associated with head width (Gao et al., 2011), carcass weight, and body weight (Wang et al., 2012), implying that this genomic region is closely associated with chicken growth. However, because we selected these two genes as candidate genes primarily based on our inferences from previous research, more experiments and evidence are needed to verify their function in comb development.

CONCLUSION

Using GWAS for comb traits with autosomal SNPs and INDELs derived from the WGS, the current study identified 134 SNPs and 25 INDELs strongly associated with comb traits, and two genes, *FGFR2* and *CYR5B2*, could be considered candidate genes for comb traits. The markers identified in our study could provide genomic targets for genetic improvement of comb traits in Nandan-Yao chickens. Additionally, continued investigation of these two candidate genes will help researchers further understand the genetic mechanism and identify causal variants underlying comb development.

DATA AVAILABILITY STATEMENT

The SNP and INDEL data generated from this study can be found in the European Bioinformatics Institute with the accession number PRJEB46210.

ETHICS STATEMENT

The animal study was reviewed and approved by Guangxi University's Animal Care and Use Committee.

AUTHOR CONTRIBUTIONS

ZY performed the GWAS and wrote the manuscript. LZo performed the statistical analysis of the phenotype. JD, YJ, and JJ contributed to the collection of samples and data. TS, WX, and LZe extracted the genomic DNA. XY contributed to the design of the study and the review of the manuscript. All authors read and approved the final manuscript.

FUNDING

ACKNOWLEDGMENTS

and the students from the Animal Genetics and Breeding Laboratory, College of Animal Science and Technology, Guangxi University for helping us collect phenotypic information.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.682501/full#supplementary-material

Supplementary Figure 1 | **(A)** Pearson correlation coefficients for five comb traits. **(B)** The principal component plot of Nandan-Yao chickens. **(C)** The SNP distribution in Nandan-Yao chicken, the different colors indicate the number of SNPs within 0.5 Mb window size. **(D)** The INDEL distribution in Nandan-Yao

chicken, the different colors indicate the number of INDELs within 1 Mb window size.

Supplementary Figure 2 | Venn diagrams of associated SNPs and INDELs identified on GGA 6 for comb traits.

Supplementary Figure 3 | Conditional GWAS in the prominent region on GGA 6. Regional association results of before **(left)** and after **(right)** conditional GWAS for CL, CT, and CW. The red dots represent the two lead SNPs for each trait (6:30,264,318 for CL, 6:30,354,876 for CT and CW). The blue dots indicate the $-\log_{10}(P)$ of SNPs greater than 6.84 (suggestive genome-wide threshold) before conditional GWAS.

Supplementary Table 1 | Detailed information about the whole-genome sequencing results.

Supplementary Table 2 | The detailed information of all suggestively associated SNPs. The text in bold represents the lead SNP in each trait.

REFERENCES

Carre, A. L., Hu, M. S., James, A. W., Kawai, K., Galvez, M. G., Longaker, M. T., et al. (2018). beta-catenin-dependent Wnt signaling: a pathway in acute cutaneous wounding. *Plast. Reconstr. Surg.* 141, 669–678. doi: 10.1097/PRS.0000000000004170

Cornwallis, C. K., and Birkhead, T. R. (2006). Social status and availability of females determine patterns of sperm allocation in the fowl. *Evolution* 60, 1486–1493. doi: 10.1111/j.0014-3820.2006.tb01227.x

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., et al. (2011). The variant call format and VCFtools. *Bioinformatics* 27, 2156–2158. doi: 10.1093/bioinformatics/btr330

Dong, X., Li, J., Zhang, Y., Han, D., Hua, G., Wang, J., et al. (2019). Genomic analysis reveals pleiotropic alleles at EDN3 and BMP7 involved in chicken comb color and egg production. *Front. Genet.* 10:612. doi: 10.3389/fgene.2019.00612

Eitan, Y., Soller, M., and Rozenboim, I. (1998). Comb size and estrogen levels toward the onset of lay in broiler and layer strain females under ad libitum and restricted feeding. *Poult. Sci.* 77, 1593–1600. doi: 10.1093/ps/77.11.1593

Eriksson, J., Larson, G., Gunnarsson, U., Bed'hom, B., Tixier-Boichard, M., Stromstedt, L., et al. (2008). Identification of the yellow skin gene reveals a hybrid origin of the domestic chicken. *PLoS Genet.* 4:e1000010. doi: 10.1371/journal.pgen.1000010

Feng, C., Gao, Y., Dorshorst, B., Song, C., Gu, X., Li, Q., et al. (2014). A cis-regulatory mutation of PDSS2 causes silky-feather in chickens. *PLoS Genet.* 10:e1004576. doi: 10.1371/journal.pgen.1004576

Gao, X. (2011). Multiple testing corrections for imputed SNPs. *Genet. Epidemiol.* 35, 154–158. doi: 10.1002/gepi.20563

Gao, X., Starmer, J., and Martin, E. R. (2008). A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. *Genet. Epidemiol.* 32, 361–369. doi: 10.1002/gepi.20310

Gao, Y., Feng, C. G., Song, C., Du, Z. Q., Deng, X. M., Li, N., et al. (2011). Mapping quantitative trait loci affecting chicken body size traits via genome scanning. *Anim. Genet.* 42, 670–674. doi: 10.1111/j.1365-2052.2011.02193.x

Gel, B., and Serra, E. (2017). karyoploteR: an R/bioconductor package to plot customizable genomes displaying arbitrary data. *Bioinformatics* 33, 3088–3090. doi: 10.1093/bioinformatics/btx346

Graves, H. B., Hable, C. P., and Jenkins, T. H. (1985). Sexual selection in gallus: effects of morphology and dominance on female spatial behavior. *Behav. Processes* 11, 189–197. doi: 10.1016/0376-6357(85)90060-9

Guo, J., Qu, L., Dou, T. C., Shen, M. M., Hu, Y. P., Ma, M., et al. (2020). Genome-wide association study provides insights into the genetic architecture of bone size and mass in chickens. *Genome* 63, 133–143. doi: 10.1139/gen-2019-0022

Guo, Y., Gu, X., Sheng, Z., Wang, Y., Luo, C., Liu, R., et al. (2016). A complex structural variation on chromosome 27 leads to the ectopic expression of HOXB8 and the muffs and beard phenotype in chickens. *PLoS Genet.* 12:e1006071. doi: 10.1371/journal.pgen.1006071

Johnsson, M., Gustafson, I., Rubin, C. J., Sahlqvist, A. S., Jonsson, K. B., Kerje, S., et al. (2012). A sexual ornament in chickens is affected by pleiotropic alleles

at HAO1 and BMP2, selected during domestication. *PLoS Genet.* 8:e1002914. doi: 10.1371/journal.pgen.1002914

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Lien, C. Y., Tixier-Boichard, M., Wu, S. W., Wang, W. F., Ng, C. S., and Chen, C. F. (2017). Detection of QTL for traits related to adaptation to sub-optimal climatic conditions in chickens. *Genet. Sel. Evol.* 49:39. doi: 10.1186/s12711-017-0314-5

Liu, Q., Liu, Y., Li, W., Wang, X., Sawaya, R., Lang, F. F., et al. (2015). Genetic, epigenetic, and molecular landscapes of multifocal and multicentric glioblastoma. *Acta Neuropathol.* 130, 587–597. doi: 10.1007/s00401-015-1470-8

Liu, Z., Yang, N., Yan, Y., Li, G., Liu, A., Wu, G., et al. (2019). Genome-wide association analysis of egg production performance in chickens across the whole laying period. *BMC Genet.* 20:67. doi: 10.1186/s12863-019-0771-7

Moreira, G. C. M., Salvian, M., Boschiero, C., Cesar, A. S. M., Reecy, J. M., Godoy, T. F., et al. (2019). Genome-wide association scan for QTL and their positional candidate genes associated with internal organ traits in chickens. *BMC Genomics* 20:669. doi: 10.1186/s12864-019-6040-3

Nakano, T., Imai, S., Koga, T., and Sim, J. S. (1996). Light microscopic histochemical and immunohistochemical localisation of sulphated glycosaminoglycans in the rooster comb and wattle tissues. *J. Anat.* 189 (Pt 3), 643–650.

Navara, K. J., Anderson, E. M., and Edwards, M. L. (2012). Comb size and color relate to sperm quality: a test of the phenotype-linked fertility hypothesis. *Behav. Ecol.* 23, 1036–1041. doi: 10.1093/beheco/ars068

O'Leary, N. A., Wright, M. W., Brister, J. R., Ciufo, S., Haddad, D., McVeigh, R., et al. (2016). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* 44, D733–D745. doi: 10.1093/nar/gkv1189

Pfaff, M. J., Xue, K., Li, L., Horowitz, M. C., Steinbacher, D. M., and Eswarakumar, J. V. P. (2016). FGFR2c-mediated ERK-MAPK activity regulates coronal suture development. *Dev. Biol.* 415, 242–250. doi: 10.1016/j.ydbio.2016.03.026

Pizzari, T., Cornwallis, C. K., Lovlie, H., Jakobsson, S., and Birkhead, T. R. (2003). Sophisticated sperm allocation in male fowl. *Nature* 426, 70–74. doi: 10.1038/nature02004

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* 38, 904–909. doi: 10.1038/ng1847

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. doi: 10.1086/519795

Reardon, W., Winter, R. M., Rutland, P., Pulleyn, L. J., Jones, B. M., and Malcolm, S. (1994). Mutations in the fibroblast growth factor receptor 2 gene cause Crouzon syndrome. *Nat. Genet.* 8, 98–103. doi: 10.1038/ng0994-98

Scheet, P., and Stephens, M. (2006). A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. *Am. J. Hum. Genet.* 78, 629–644. doi: 10.1086/502802

Shen, M., Qu, L., Ma, M., Dou, T., Lu, J., Guo, J., et al. (2016). Genome-wide association studies for comb traits in chickens. *PLoS One* 11:e0159081. doi: 10.1371/journal.pone.0159081

Siegel, P. B., and Dudley, D. S. (1963). Comb type, behavior and body weight in chickens. *Poult. Sci.* 42, 516–522. doi: 10.3382/ps.0420516

Symeon, G. K., Mantis, F., Bizelis, I., Kominakis, A., and Rogdakis, E. (2012). Effects of caponization on growth performance, carcass composition and meat quality of males of a layer line. *Animal* 6, 2023–2030. doi: 10.1017/S175173111200102

Tam, V., Patel, N., Turcotte, M., Bosse, Y., Pare, G., and Meyre, D. (2019). Benefits and limitations of genome-wide association studies. *Nat. Rev. Genet.* 20, 467–484. doi: 10.1038/s41576-019-0127-1

Tan, A., Abecasis, G. R., and Kang, H. M. (2015). Unified representation of genetic variants. *Bioinformatics* 31, 2202–2204. doi: 10.1093/bioinformatics/btv112

van Kampen, M. (1971). Some aspects of thermoregulation in the white leghorn fowl. *Int. J. Biometeorol.* 15, 244–246. doi: 10.1007/BF01803906

Wallis, J. W., Aerts, J., Groenen, M. A., Crooijmans, R. P., Layman, D., Graves, T. A., et al. (2004). A physical map of the chicken genome. *Nature* 432, 761–764. doi: 10.1038/nature03030

Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 38:e164. doi: 10.1093/nar/gkq603

Wang, S. Z., Hu, X. X., Wang, Z. P., Li, X. C., Wang, Q. G., Wang, Y. X., et al. (2012). Quantitative trait loci associated with body weight and abdominal fat traits on chicken chromosomes 3, 5 and 7. *Genet. Mol. Res.* 11, 956–965. doi: 10.4238/2012.April.19.1

Wilke, T. A., Gubbels, S., Schwartz, J., and Richman, J. M. (1997). Expression of fibroblast growth factor receptors (FGFR1, FGFR2, FGFR3) in the developing head and face. *Dev. Dyn.* 210, 41–52. doi: 10.1002/(SICI)1097-0177(199709)210:1<41::AID-AJA5<3.0.CO;2-1

Wright, D., Kerje, S., Brandstrom, H., Schutz, K., Kindmark, A., Andersson, L., et al. (2008). The genetic architecture of a female sexual ornament. *Evolution* 62, 86–98. doi: 10.1111/j.1558-5646.2007.00281.x

Wright, D., Rubin, C., Schutz, K., Kerje, S., Kindmark, A., Brandstrom, H., et al. (2012). Onset of sexual maturity in female chickens is genetically linked to loci associated with fecundity and a sexual ornament. *Reprod. Domest. Anim.* 47(Suppl. 1), 31–36. doi: 10.1111/j.1439-0531.2011.01963.x

Yang, J., Lee, S. H., Goddard, M. E., and Visscher, P. M. (2011). GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* 88, 76–82. doi: 10.1016/j.ajhg.2010.11.011

Yang, Z., Deng, J., Li, D., Sun, T., Xia, L., Xu, W., et al. (2020). Analysis of population structure and differentially selected regions in guangxi native breeds by restriction site associated with DNA sequencing. *G3 (Bethesda)* 10, 379–386. doi: 10.1534/g3.119.400827

Zerbino, D. R., Achuthan, P., Akanni, W., Amode, M. R., Barrell, D., Bhai, J., et al. (2018). Ensembl 2018. *Nucleic Acids Res.* 46, D754–D761. doi: 10.1093/nar/gkx1098

Zhou, X., and Stephens, M. (2012). Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* 44, 821–824. doi: 10.1038/ng.2310

Zuk, M., Thornhill, R., Ligon, J. D., Johnson, K., Austad, S., Ligon, S. H., et al. (1990). The role of male ornaments and courtship behavior in female mate choice in red jungle fowl. *Am. Nat.* 136, 459–473. doi: 10.1086/285107