

Integrative 3D genomics with multi-omics analysis and functional validation of genetic regulatory mechanisms of abdominal fat deposition in chickens

Received: 23 May 2024

Accepted: 18 October 2024

Published online: 28 October 2024

 Check for updates

Linyong Shen^{1,2,3,4}, Xue Bai^{1,2,3,4}, Liru Zhao^{1,2,3,4}, Jiamei Zhou^{1,2,3}, Cheng Chang^{1,2,3}, Xinquan Li^{1,2,3}, Zhiping Cao^{1,2,3}, Yumao Li^{1,2,3}, Peng Luan^{1,2,3}, Hui Li^{1,2,3}✉ & Hui Zhang^{1,2,3}✉

Chickens are the most abundant agricultural animals globally, with controlling abdominal fat deposition being a key objective in poultry breeding. While GWAS can identify genetic variants associated with abdominal fat deposition, the precise roles and mechanisms of these variants remain largely unclear. Here, we use male chickens from two lines divergently selected for abdominal fat deposition as experimental models. Through the integration of genomic, epigenomic, 3D genomic, and transcriptomic data, we build a comprehensive chromatin 3D regulatory network map to identify the genetic regulatory mechanisms that influence abdominal fat deposition in chickens. Notably, we find that the rs734209466 variant functions as an allele-specific enhancer, remotely enhancing the transcription of *IGFBP2* and *IGFBP5* by the binding transcription factor *IRF4*. This interaction influences the differentiation and proliferation of preadipocytes, which ultimately affects phenotype. This work presents a detailed genetic regulatory map for chicken abdominal fat deposition, offering molecular targets for selective breeding.

Chicken, being the most widely consumed meat globally, is valued for its high-quality protein content and cost-effectiveness¹. However, traits related to abdominal fat deposition (AFD) significantly impact meat quality and production efficiency. Research indicates that excessive fat accumulation in chickens can notably decrease feed conversion efficiency² and carcass lean meat rate³. Moreover, breeder chickens with excessive fat can negatively affect egg production rate, fertilization rate, and hatchability rate, potentially leading to increased mortality during the laying period⁴. Therefore, effectively controlling AFD in chickens is a crucial issue that requires urgent attention in chicken breeding worldwide.

Furthermore, chickens are valuable model organisms in life sciences research, playing a significant role in the study of human obesity and related metabolic diseases^{5,6}. They exhibit characteristics similar to early stages of human type 2 diabetes such as high blood glucose and insulin resistance⁷, making them an ideal animal model for obesity research. For instance, the obesity-inducing virus *SMAM-1*, which leads to obesity in chickens, has been found to have similar effects in humans⁸. Moreover, several quantitative trait loci associated with obesity in chickens contain genes linked to human susceptibility to obesity⁹. Therefore, studying the regulatory mechanisms of AFD traits in chickens also offers valuable insights

¹College of Animal Science and Technology, Northeast Agricultural University, Harbin 150030, PR China. ²Key Laboratory of Chicken Genetics and Breeding, Ministry of Agriculture and Rural Affairs, Harbin 150030, PR China. ³Key Laboratory of Animal Genetics, Breeding and Reproduction, Education Department of Heilongjiang Province, Harbin 150030, PR China. ⁴These authors contributed equally: Linyong Shen, Xue Bai, Liru Zhao. ✉e-mail: lihui@neau.edu.cn; huizhang@neau.edu.cn

into understanding the molecular mechanisms of human obesity and related diseases.

The AFD traits are a complex quantitative trait controlled by multiple genomic variants with intricate genetic regulatory mechanisms. Recent research has made significant progress in identifying genomic variants associated with AFD traits through genome-wide association studies (GWAS). Despite these advancements, the functional mechanisms of these variants, especially those located in non-coding regions, remain poorly understood, leading to a ‘black box’ scenario regarding their specific effects on phenotype. Traditional research methodologies, such as the ‘nearest gene model’ based on locational proximity, simply assign the nearest gene as the regulatory target of a genomic variant¹⁰. However, recent studies revealed that approximately 75–76% of genomic variants affect target genes through long-range interactions¹¹, which challenges the effectiveness of the ‘nearest gene model’.

To overcome these limitations, many researchers have turned to expression quantitative trait locus (eQTL) analysis, which aims to identify correlations between genomic variants and the expression levels of their corresponding target genes. International initiatives, such as the Farm Animal Genotype–Tissue Expression (FarmGTEx) project^{12–14}, have contributed to extensive eQTL datasets on adipose tissues from agricultural animals, enhancing our understanding of the genetic mechanisms underlying genomic variants associated with economic traits. However, eQTL analysis predominantly relies on statistical correlations and fails to directly elucidate the functionality of genomic variants. Emerging epigenetic techniques suggest that variants located in regulatory elements have functional implications¹⁵, but these methods are limited in explicating how functional variants exert their influence through specific gene regulation. In contrast, three-dimensional (3D) genomics approaches based on the frequency of physical chromatin interactions can predict target genes but often neglect the significant influence of epigenetic activity on variants^{11,16,17}. Therefore, the integration of 3D genomics with multiple omics datasets is crucial for comprehensively understanding how noncoding variants influence transcriptional regulation^{18–22}.

In this study, we used two Northeast Agricultural University Broiler Lines (NEAUHLF), which have been selected for divergent abdominal fat content since 1996, as our experimental populations. These two lines have proven to be ideal genetic models for investigating the underlying mechanisms of obesity²³. Building upon this, we constructed an Integrative Multi-omics approach for Variant–Gene Interactions (IMVGI) that considers genomic, epigenomic, 3D genomic, and transcriptomic data and broadly analyzed how genomic variants systematically affect the transcription regulatory mechanisms of AFD traits. The strategy was to use whole-genome sequencing analysis of the NEAUHLF to identify genomic variants in the chicken genome, and selection signature analysis was used to identify variants associated with AFD traits. Subsequently, the functionality of these variants was annotated through epigenetic analysis, and their potential target genes were identified through 3D genomic analysis. By further employing weighted gene coexpression network analysis (WGCNA), we established correlations between these target genes and AFD traits. Finally, experimental validation revealed the role of crucial functional variants in AFD traits and their regulatory mechanisms. In summary, our research unveils a detailed framework that shows how genomic variants influence transcriptional regulation associated with AFD traits, providing insights into the genetic basis of these traits and identifying significant targets for the molecular breeding of chicken AFD traits.

Results

Analysis of genomic selection signatures uncovers candidate variants associated with fat deposition

In this research, we conducted a thorough investigation to identify candidate variants related to abdominal fat deposition (AFD) traits.

Our research focused on a unique population of 330 chickens from the Northeast Agricultural University broiler lines that were divergently selected for abdominal fat content (NEAUHLF) for 19th generations (years). Despite their similar body weights at 7 weeks of age (BW7), chickens from the fat (FL) and lean (LL) lines displayed a notable 7.83-fold difference in abdominal fat weight (AFW) and a 7.51-fold difference in abdominal fat percentage (AFP = AFW/BW7) (Fig. 1a, b and Supplementary Fig. 1). This discrepancy underscores the significant phenotypic divergence between the FL and LL groups in terms of AFD traits.

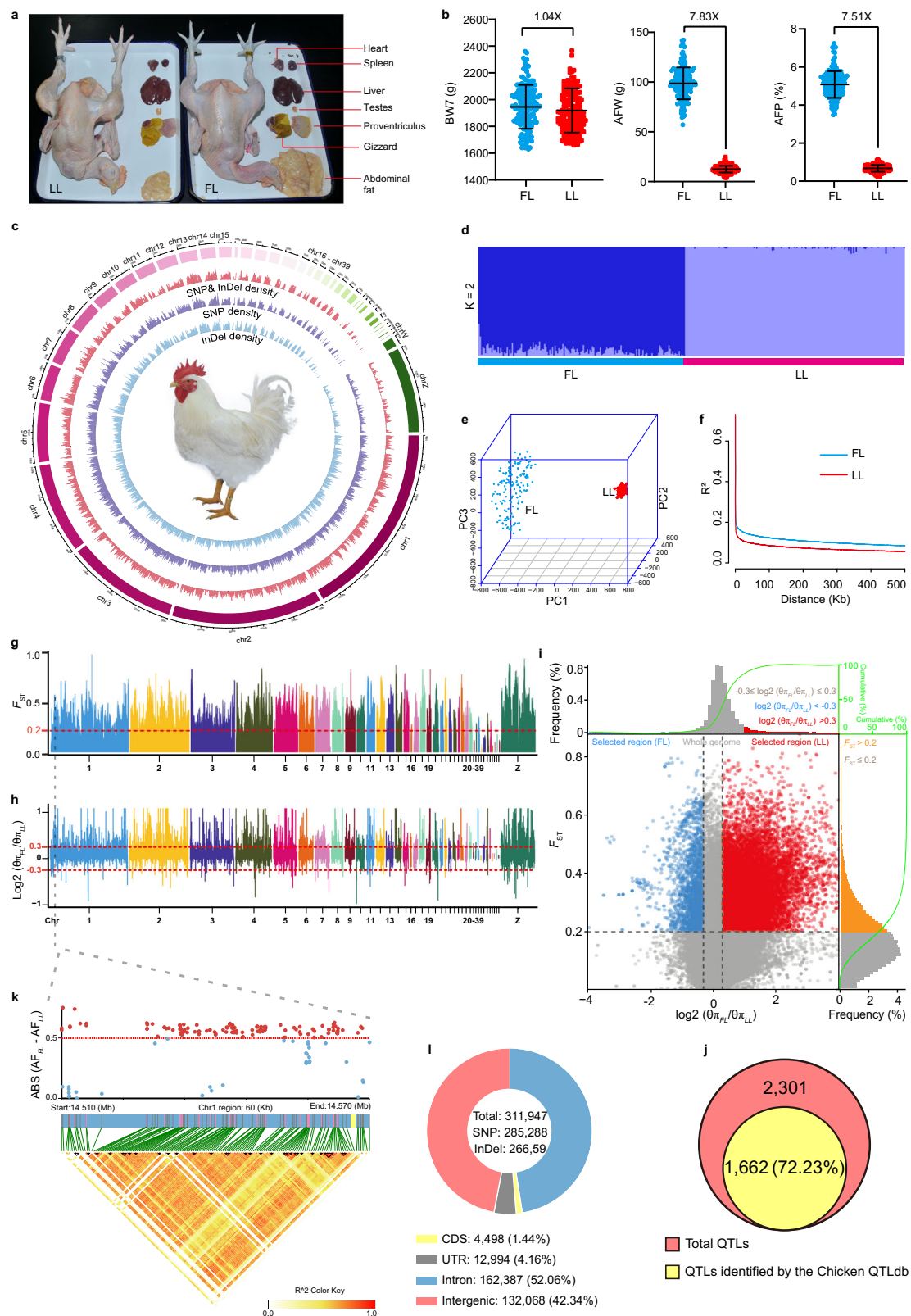
Population structure analysis was subsequently conducted to estimate the genomic relationship between these two groups. Whole-genome sequencing (WGS) of these 330 individuals revealed 4,677,252 high-quality, filtered genomic variants (4,140,284 single nucleotide polymorphisms (SNPs) and 536,968 insertions/deletions (InDels)). Intriguingly, 20.01% of these variants were novel and were not recorded in the dbSNP database (https://www.ncbi.nlm.nih.gov/projects/SNP/snp_summary.cgi, last accessed on August 22, 2021) (Fig. 1c). Admixture analysis partitioned the population into FL and LL subpopulations at $K = 2$ (Fig. 1d), which was supported by principal component analysis (PCA) (Fig. 1e). Linkage disequilibrium (LD) analysis indicated rapid degradation, particularly in the LL subpopulation, suggesting intense selection pressure (Fig. 1f). The significant genetic differences between the FL and LL groups make them ideal subjects for studying traits influenced by artificial selection, such as AFD traits.

To identify candidate variants associated with AFD traits, we employed selection signature analysis. We calculated the fixation indices (F_{ST}) and θ_{π} ratios across the genome, leading to the identification of 2301 regions selected for fat traits (Fig. 1g–i). Remarkably, 72.23% of these regions overlapped with known quantitative trait loci (QTLs) associated with the fat traits from the Chicken QTLdb²⁴ (Fig. 1j), suggesting their potential involvement in AFD traits. By focusing on these regions, we identified variants that showed significant changes in allele frequency ($|\Delta AF| \geq 0.5$) between FL and LL as candidate variants linked to AFD traits. This process allowed us to identify 311,947 candidate variants associated with AFD traits, predominantly located in noncoding intergenic/intronic regions (94.42%), posing challenges in identifying the true functional variants and their target genes (Fig. 1k, l).

Identification of potentially functional variants regulating fat deposition using epigenomic techniques

To identify the potentially functional variants among the 311,947 variants associated with AFD traits, we employed a comprehensive annotation strategy. This methodology entailed the synthesis of data derived from transposase-accessible chromatin sequencing (ATAC-seq) and chromatin immunoprecipitation sequencing (ChIP-seq) of chicken abdominal fat tissues to increase the precision of our identification process.

In our analysis of open chromatin regions (OCRs) using ATAC-seq data, we focused on identifying differentially open chromatin regions (DOCRs) in abdominal fat tissues between the FL and LL groups, as those showing statistically significant changes ($p < 0.05$, $|\log_2(\text{fold change})| > 1$). This approach led us to identify 128,090 OCRs, 4028 of which met the criteria for classification as DOCRs (Fig. 2a, b). These DOCRs exhibited distinct length characteristics and were predominantly enriched in regions flanking transcription start sites (TSSs) (Fig. 2c, d). We further classified these DOCRs into distinct regulatory elements (REs) based on their genomic features using ChIP-seq data. Specifically, we identified 218 promoters (TSS –4 kb/+2 kb, H3K4me3, H3K27ac), 1106 enhancers (H3K4me1, H3K27ac), 312 silencers (H3K27me3), 2337 CTCF-enriched regions (marked by CTCF peaks), and 1264 low-signal regions (regions lacking significant RE signals) (Fig. 2e, f). Notably, some DOCRs may exhibit multiple regulatory functions, leading to their classification as various types of REs



(Supplementary Fig. 2a). These REs showed higher fold changes in activity signals between FL and LL compared to low-signal regions (Fig. 2g), and these low signal regions are highly conserved across all biological replicate samples (Supplementary Fig. 2b), thereby enhancing the credibility of our annotations.

Intriguingly, in addition to typical REs, we also identified a distinct class of enhancers and silencers, termed superenhancers (SEs) and

supersilencers (SSs), respectively. SEs were defined by applying the ranking of the ROSE pipeline²⁵ to ATAC-seq and H3K4me1 signals, which revealed 280 SEs in DOCRs (Fig. 2h). The authors of a recent study referred to H3K27me3-rich genomic regions as SSs that promote gene repression through chromatin interactions²⁶. Similar to the identification of SEs, we defined SSs by utilizing the ROSE algorithm to rank ATAC-seq signals and H3K27me3 signals. This method allowed

Fig. 1 | Genomic analysis identifying candidate variants for fat deposition in broiler chicken lines. **a** Comparative visualization of fat (FL) and lean (LL) lines, highlighting the marked difference in abdominal fat tissue. **b** Graphical representation of body weight at 7 weeks of age (BW7), abdominal fat weight (AFW), and abdominal fat percentage (AFP = AFW/BW7) in the FL and LL groups. Data were shown as the mean \pm SD, $n = 160$ (FL), 170 (LL), biologically independent animals. Statistical analysis was performed using an unpaired two-tailed Student's t -test. The fold differences observed were as follows: BW7 in FL relative to LL is 1.04-fold ($p < 0.0001$); AFW is 7.83-fold ($p < 0.0001$); and AFP shows a 7.51-fold difference ($p < 0.0001$). **c** Circos plot displaying the distribution characteristics of genomic variants, including SNPs and InDels, across the whole genome. **d** Population structure analysis at $K = 2$, illustrating distinct genetic clusters. **e** Three-dimensional principal component analysis (PCA) plot. **f** Analysis of linkage

disequilibrium (LD) decay distance, quantified by R^2 values. **g** Genome-wide F_{ST} bar plot, with a red dashed line indicating the cutoff threshold ($F_{ST} > 0.2$), highlighting regions of high genetic differentiation. **h** Genome-wide $\theta\pi$ ratio bar plot. The red dashed line indicates the cutoff thresholds for $\log_2(\theta\pi \text{ ratios}) > 0.3$ or < -0.3 , identifying areas under selection or with high genetic variation. **i** Scatter plot showing 2301 selected regions associated with abdominal fat deposition AFD traits identified through joint F_{ST} and $\theta\pi$ ratios. **j** Venn diagram illustrating the overlap between the 2301 adipose-related selected regions and known adipose-related QTLs from the Chicken QTLdb. **k** Manhattan plot highlighting significant changes in allele frequency (AF) between the FL and LL groups, with regions under selection marked. **l** Pie chart categorizing the 311,947 candidate variants linked to AFD traits. Source data are provided as a Source Data file.

for the recognition of 41 SSs in DOCRs (Fig. 2i). These SEs/SSs showed stronger activity signals than regions with typical REs (Fig. 2j, k), enhancing the credibility of our annotations.

After applying the comprehensive annotation strategy using ATAC-seq and ChIP-seq data, we identified 2162 potentially functional variants embedded within annotated REs. The functional relevance of these variants is underscored by their distribution: approximately half were significantly enriched at CTCF-binding sites (50.66%), and the rest (45.24%) were nested within distal regulatory elements such as enhancers, SEs, silencers, and SSs (Fig. 2l). The localization of these 2162 variants within REs has potential functional significance, as they may influence the activity of these elements and consequently modify the transcriptional regulation of target genes.

Identification of variant–gene interactions using three-dimensional genomes

To investigate how 2162 potentially functional variants influence gene regulatory mechanisms, we utilized Hi-C technology to perform high-resolution genome-wide chromatin interaction mapping of the abdominal fat tissues of the FL and LL groups (Supplementary Table 1). Through this chromatin interaction mapping, we constructed chromatin interaction frequency plots, revealing complex chromatin structures (A/B compartments, topologically associating domains [TADs], and loops) in both the FL and LL genomes (Fig. 3a). Furthermore, we crafted individualized three-dimensional (3D) chromatin models for FL and LL, revealing marked structural differences between the two groups. Notably, the LL group exhibited a weaker chromatin conformation, as indicated by increased Von Neumann entropy²⁷ and decreased interaction frequencies (Fig. 3b–d, Supplementary Figs. 3 and 4).

We conducted a comparative analysis of A/B compartments, TADs, and loop structures between the FL and LL groups. Our analysis at a 100 kb resolution revealed that only 6.16% of the total genomic regions showed transitions between the A and B compartments in the FL and LL groups, suggesting high conservation in these compartment regions (Fig. 3e, f). Furthermore, a more refined 10 kb resolution allowed us to identify a total of 3829 TAD boundaries, of which 2898 (75.69%) were shared by the two groups. Additionally, there were 483 (12.61%) FL-specific TAD boundaries and 448 (11.70%) LL-specific TAD boundaries (Fig. 3g, h). Further examination of loop structures at resolutions of 1, 5, 10, and 25 kb yielded a total of 4246 loops, with only 1243 (29.27%) loops being shared by the two groups. Additionally, there were 2100 (49.46%) FL-specific loops and 903 (21.27%) LL-specific loops, which was further confirmed by CTCF enrichment at loop anchors (Fig. 3i, j and Supplementary Fig. 4d). These findings underscored a notable disparity in TADs (24.31%) and loops (70.73%) structures between FL and LL, suggesting their potential regulatory roles.

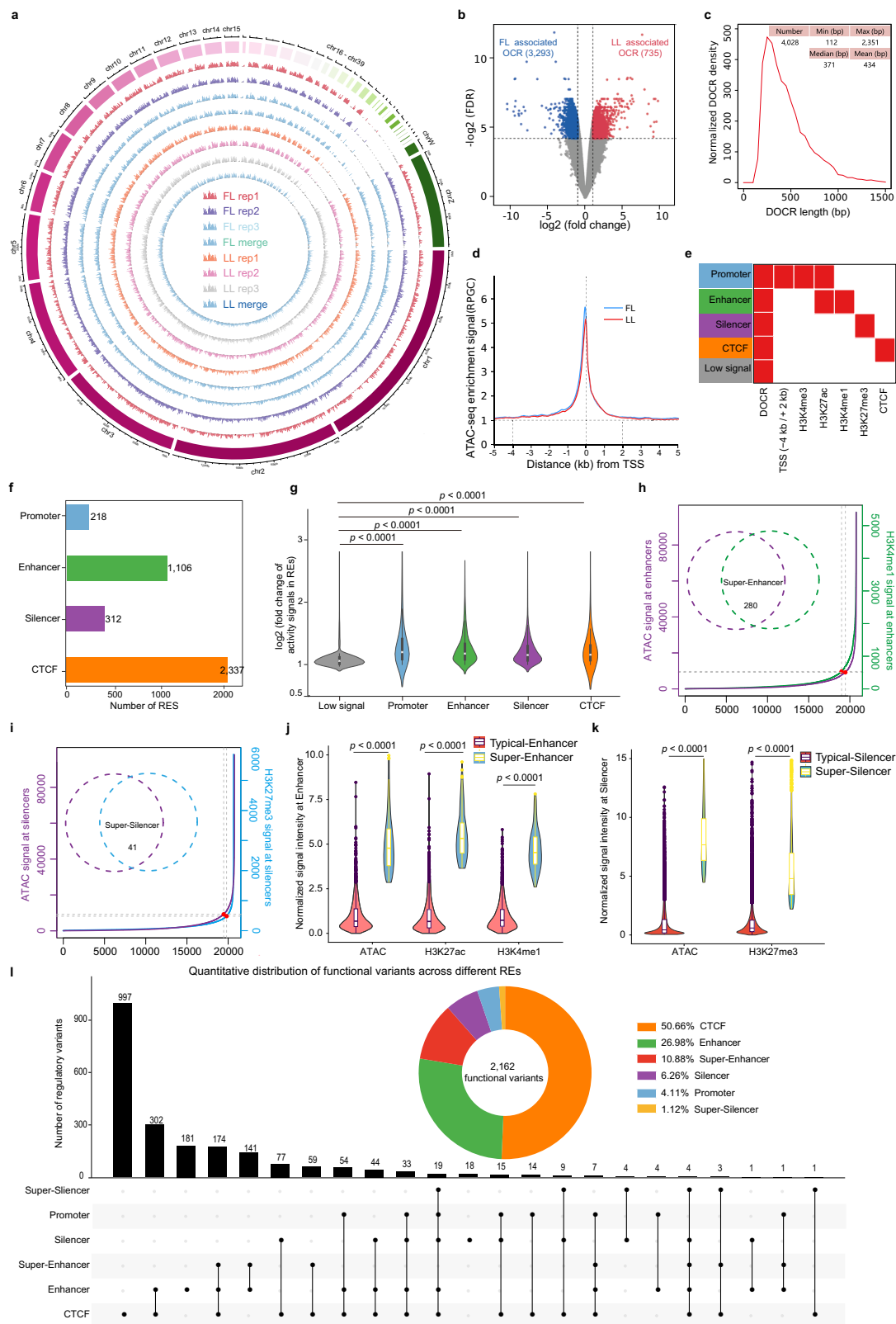
We annotated the 2162 potentially functional variants associated with AFD traits on the high-order structure of the 3D genome, focusing on TAD boundaries and loops, to identify target genes regulated by these variants. This process revealed four primary

transcriptional regulatory mechanisms mediated by functional variants: (1) alterations in proximal RE activity, in which 552 functional variants modulate the activity of 168 proximal REs (promoters, nearby enhancers, and nearby silencers) to regulate the transcription of 154 target genes; (2) modifications in distal RE activity, in which 619 functional variants exert transcriptional control over 261 target genes by altering the activity of 160 REs (enhancers and silencers); (3) reconfiguration of the TAD structure, in which 40 functional variants impact the transcription of 18 target genes by modifying 8 TAD structures, particularly at CTCF binding sites located at TAD boundaries; and (4) loop structure alterations, in which 666 functional variants regulate 267 distant target genes by influencing 163 loop structures, especially at CTCF binding sites within loop anchor points (Fig. 3k, l). In total, our analysis identified a compendium of 1169 functional variants that regulate the expression of 500 target genes by influencing chromatin structures such as proximal and distal regulatory elements, TADs, and loop structures, resulting in a complex network of 2559 variant–gene interactions.

Weighted gene coexpression network analysis reveals key variant–gene interactions related to fat deposition

After identifying the target genes of potentially functional variants, we employed weighted gene coexpression network analysis (WGCNA) to explore the associations between the expression of target genes and fat traits, thereby establishing comprehensive variant–gene networks related to AFD traits. Upon detailed analytical consideration, a power beta value of 11 was chosen for WGCNA to ensure a scale-free network (Fig. 4a, b). Using this power beta value, 439 of the above 500 target genes were successfully grouped into four distinct functional modules (Fig. 4c). Specifically, the *blue module* encompassed 133 genes, the *turquoise module* contained 172 genes, the *brown module* comprised 86 genes, and the *yellow module* included 48 genes. The remaining 61 genes, which did not exhibit a sufficiently strong coexpression relationship with the others, have been allocated to the *gray* group in Fig. 4c.

The association analysis between the gene modules and phenotypes revealed a significant relationship between 439 target genes and AFD traits among the 1134 functional variants. Specifically, 305 target genes of 900 functional variants were associated to fat deposition. This finding is supported by the positive correlations observed in the *blue module*, which exhibited a strong positive correlation with AFW ($r = 0.90$, $p = 4e - 04$) and AFP ($r = 0.87$, $p = 1e - 04$), and in the *turquoise module*, which showed a similarly strong positive correlation with AFW ($r = 0.98$, $p = 8e - 07$) and AFP ($r = 0.99$, $p = 1e - 07$) (Fig. 4c). Conversely, another set of 134 target genes of 518 functional variants appeared to inhibit fat deposition. This observation was supported by the negative correlations in the *brown module* with AFW ($r = -0.64$, $p = 0.01$) and AFP ($r = -0.65$, $p = 0.01$) and in the *yellow module* with AFW ($r = -0.53$, $p = 0.04$) and AFP ($r = -0.54$, $p = 0.04$) (Fig. 4c). Notably, there is an overlap of 284 variants between the two sets, resulting in a total of 1134 unique variants ($900 + 518 - 284$) regulating the 439



genes associated with AFD traits. Additionally, 61 target genes of 35 functional variants, encompassing 304 variant–gene interactions, were identified within the *grey module* and were not associated with AFD traits.

In total, our analysis successfully highlighted 439 significant adipose-related target genes influenced by 1134 functional variants,

encompassing 2255 variant–gene interactions (Fig. 4d). Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses further confirmed that these genes strongly correlated with lipid metabolism-related pathways, primarily signal transduction, metabolic processes, and cell growth and death (Fig. 4e, f and Supplementary Fig. 5d–f).

Fig. 2 | Epigenomic screening of functional variants influencing fat deposition. **a** Circos plot showing the distribution and frequency of open chromatin regions (OCRs) across chromosomes identified by ATAC-seq, with each ring representing different sample data. **b** Volcano plot of 4028 differentially expressed OCRs between FL and LL, with significance thresholds. **c** Graphical representation of the size distribution of differential OCRs. **d** Enrichment analysis of ATAC-seq signals near transcription start sites (TSSs). **e** Annotation of DOCRs with regulatory elements (REs) using ChIP-seq data of histone modifications. **f** Number of annotated REs in DOCRs. **g** Comparative analysis of the activity signals of different REs. Violin plots provide a density estimation of the data distribution, and box plots display the median (central line), interquartile range (IQR, bounds), and whiskers representing the minimum and maximum values. Sample sizes are $n = 1264$ (Low signal), 218 (Promoter), 1106 (Enhancer), 312 (Silencer), and 2337 (CTCF), biologically

independent samples. P values were obtained using one-way ANOVA followed by Dunnett's multiple comparisons test. **h** Circular plot depicting the identification of superenhancers within DOCRs. **i** Circular plot showing the identification of super-silencers within DOCRs. **j**, **k** Comparative analysis of chromatin marks in enhancers and silencers, with box plots showing the median, interquartile range, and whiskers representing the range of data, and violin plots providing a density estimation of the data distribution. Data are plotted in box plots showing the median (central line) and IQR (bounds) with whiskers extending to the minimum and maximum values. $n = 1106$ (Typical enhancer), 280 (Super enhancer), 312 (Typical silencer) and 41 (Super silencer), biologically independent samples. P values were obtained using unpaired two-tailed Student's t -test. **l** Bar and pie charts showing the distribution and proportion of functional variants across different REs. Source data are provided as a Source Data file.

Functional characterization of variant–gene interactions in relation to fat deposition

In this study, we constructed an integrated framework, as described above, that combines genomic, epigenomic, 3D genomic, and transcriptomic data to unravel the mechanisms of transcriptional regulation involved in the modulation of AFD traits by functional variants; this framework is termed Integrative Multiomics for Variant–Gene Interactions (IMVGI). This holistic approach allowed a systematic analysis of functional variants and their impact on transcriptional regulatory mechanisms associated with AFD traits. (Fig. 5a). Using IMVGI, we identified 2255 adipose-related variant–gene interactions, including 1134 functional variants (90.48% SNPs and 9.52% InDels) and 439 target genes. These functional variants were predominantly localized in enhancer-associated regions (45.67%) and CTCF-associated regions (40.06%) (Fig. 5b and Supplementary Data 1–3). On average, each functional variant was predicted to regulate 1.94 target genes, whereas each target gene was influenced by 4.23 functional variants. The average genomic distance of these interactions was approximately 201 kb, with about 71.00% of variants skipping over the nearest gene to regulate their target genes through long-range interactions (Fig. 5c–f and Supplementary Fig. 5a–c), highlighting the importance of integrating 3D genomic interactions in this study.

To further characterize the 1134 identified functional variants (IMVGI variants), from three aspects: epigenomic features, transcription factor enrichment, and QTL regions associated with AFD traits, we generated a control set of variants (non-IMVGI variants) using a web tool vSampler with the allele frequencies, number of variants in LD, as well as genomic distribution matched to IMVGI variants. Compared with non-IMVGI variants, IMVGI variants were significantly enriched in active chromatin regions marked by ATAC-seq, H3K27ac, H3K4me3, H3K4me1, and CTCF ($p < 0.0001$, Fig. 5g). Notably, several transcription factors associated with AFD traits, including members of the KLF, TCF, CEBP, and SOX families, showed a preference for binding to IMVGI variants ($p < 0.01$, Fig. 5h). Additionally, 73.02% of IMVGI variants showed significant enrichment in QTL regions related to AFD traits according to the Chicken QTLdb²⁴. In contrast, only 10.05% of non-IMVGI variants displayed similar enrichment in these QTLs (Fig. 5i). Together, these observations suggested that IMVGI variants play a central role in the transcriptional regulatory mechanisms associated with AFD traits.

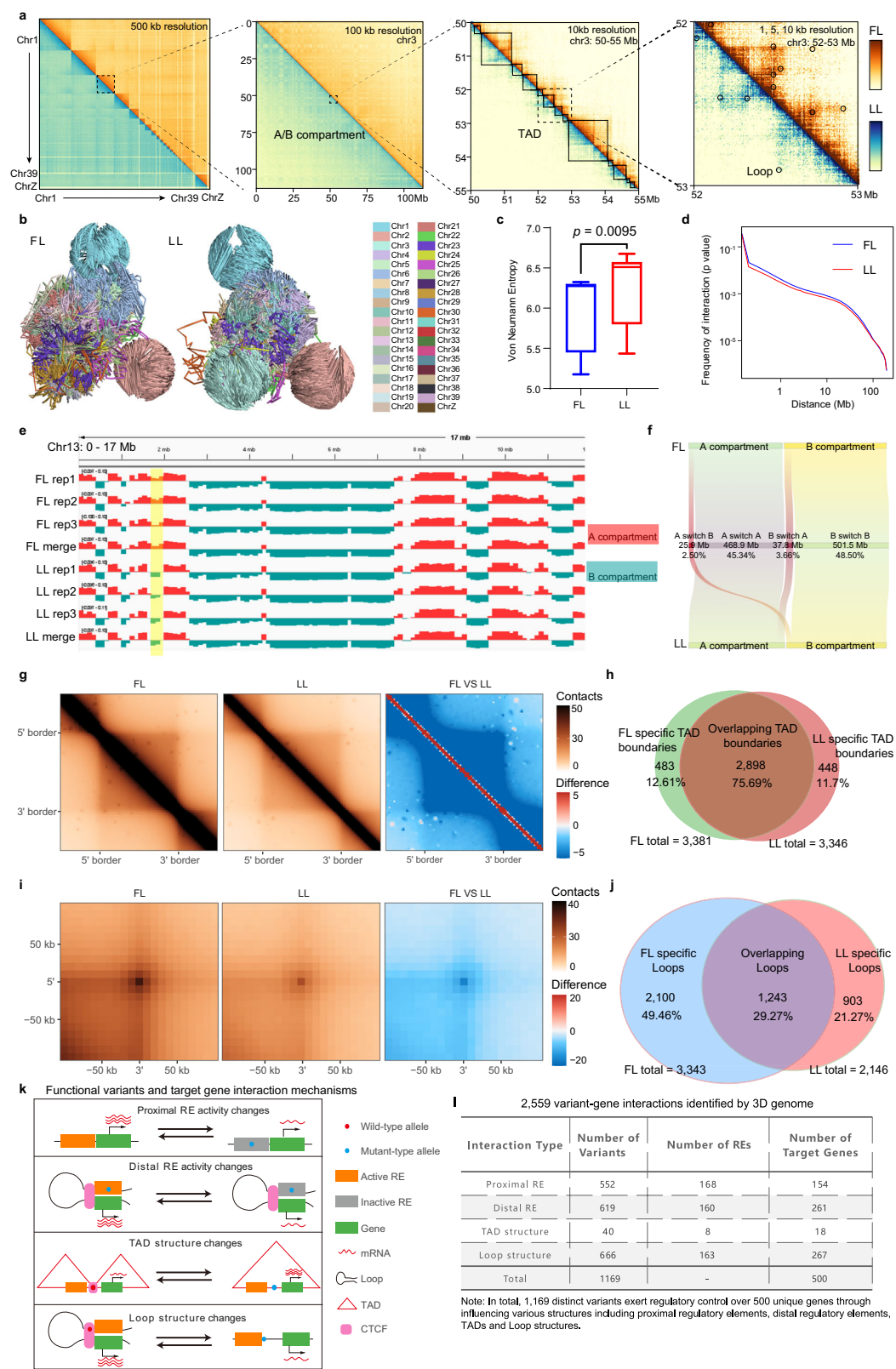
To evaluate the efficacy of the IMVGI model in predicting target genes of functional variants, we compared it with other approaches based on subsets of features such as positional distribution, regulatory element activity, 3D loops and ChickenGTEx eQTL¹³ data. Our results indicated that IMVGI can predict more distal target genes and predicts target genes that show greater mRNA expression differences between the abdominal fat tissues of FL and LL broiler populations, which have significantly different abdominal fat contents (Fig. 5j). These results suggest that IMVGI is effective in integrating diverse genomic features

to accurately link precisely positioned variants with target genes involved in AFD traits.

IMVGI variant rs734209466 as an allele-specific enhancer promoting the transcription of *IGFBP2* and *IGFBP5*

Of the 2255 variant–gene interactions identified by the IMVGI model, we selected a representative interaction for further experimental validation. Our IMVGI model positioned both SNP rs16596562 (G > A) and InDel rs734209466 (a 13-base pair InDel) within a predicted active enhancer region enriched with high levels of H3K27ac and H3K4me1 markers in the FL group (Fig. 6a, Supplementary Fig. 6). Sanger sequencing further confirmed the presence of both rs16596562 and rs734209466 in broilers (Fig. 6b). We then examined the association between the genotypes of these two variants and fat traits (AFW and AFP). The results indicated that individuals with the rs16596562-AA and rs734209466-InIn genotypes had higher AFW and AFP than those with the rs16596562-GG and rs734209466-DelDel genotypes ($p < 0.0001$, Fig. 6c), suggesting that rs16596562-A and rs734209466-In are among the potential functional alleles influencing AFD traits. To detect their functional role, we employed allele-specific luciferase reporter assays. The results showed that both the rs16596562-A and rs734209466-In alleles notably increased enhancer activity compared to the rs16596562-G and rs734209466-Del alleles ($p < 0.0001$, Fig. 6d), thus functioning as allele-specific enhancer.

Our IMVGI model further revealed that rs16596562 and rs734209466 may directly interact with the promoter regions of the distal target genes *IGFBP2* and *IGFBP5* through chromatin loops, while potentially influencing the proximal target genes *CXCR1* and *TNSI* through proximity effects, enhancing their transcriptional activation (Fig. 6a and Supplementary Fig. 7). We then used RT-qPCR analysis to indicate significant differences in the mRNA expression levels of the potential target genes *CXCR1*, *TNSI*, *IGFBP2*, and *IGFBP5* between the adipose tissues of FL and LL ($p < 0.05$, Fig. 6e), suggesting these genes as potential influencers of AFD traits. Luciferase reporter assays further demonstrated that, in comparison to the rs734209466-Del allele, the rs734209466-In allele significantly elevated the promoter transcriptional activities of *IGFBP2* and *IGFBP5* ($p < 0.0001$, Fig. 6f, g), but did not affect the transcriptional activities of *CXCR1* and *TNSI* ($p > 0.05$, Fig. 6f, g). Conversely, the rs16596562-A allele and the rs16596562-G allele did not enhance the transcriptional activity of these target genes ($p > 0.05$, Fig. 6f, g). These results suggested the role of rs734209466-In as an allele-specific enhancer, particularly in the upregulation of *IGFBP2* and *IGFBP5* transcription. Additionally, individuals carrying the rs734209466-In allele displayed elevated mRNA expression levels for *IGFBP2* and *IGFBP5* in adipose tissue, further confirming the allele's role as a specific enhancer in the transcriptional promotion of *IGFBP2* and *IGFBP5* ($p < 0.0001$, Supplementary Fig. 8).



Mediation of differential allelic activity at rs734209466 by the transcription Factor *IRF4*

Given the association of the rs734209466-In allele with increased expression of *IGFBP2* and *IGFBP5*, we hypothesized that rs734209466-In is more prone to attracting enhancer-associated transcription factors than is rs734209466-Del. To test this hypothesis, electrophoretic mobility shift assay (EMSA)

experiments were conducted. The results revealed that probes containing the biotin-labeled rs734209466-In allele could successfully form complexes with nuclear extracts, in contrast to the biotin-labeled rs734209466-Del allele (Fig. 7a). These findings suggested that sequences containing rs734209466-In serve as regulatory elements by recruiting specific transcription factors.

Fig. 3 | 3D genome mapping of variant–gene interactions in fat and lean broiler chicken lines. **a** Hi–C contact matrices at varying resolutions for chromosomes in the FL and LL genomes, with darker shades indicating higher chromatin interaction frequencies. **b** Three-dimensional average models of FL and LL chromatin structures, color-coded to represent different chromosomes at a 500 kb resolution, illustrating average inter-chromosomal spatial relationships rather than specific diploid configurations. **c** Box plots of Von Neumann entropy comparing chromatin conformational looseness in the FL and LL genomes. Boxplot center line, median; box limits, upper and lower quartiles; whiskers, 1.5× interquartile range and data beyond that threshold indicated as outliers. *P* values were obtained using unpaired two-tailed Student's *t*-test, *n* = 40 biologically independent samples. **d** Log–log plot of contact frequency versus distance, highlighting structural differences between the FL and LL genomic architectures. **e** Chromatin A/B compartment distribution

maps for specific chromosomal regions in the FL and LL genomes at 100 kb resolution. **f** Frequency charts of A/B compartment transitions, showing the proportion of transitions between A/B compartments in the FL and LL genomes. **g** Aggregate peak analysis (APA) of all topologically associating domains (TADs) within the FL and LL genomes at 10 kb resolution. **h** Statistical summary of specific and shared TAD boundaries, including counts and percentages. **i** APA for a comparative analysis of all chromatin loops within the FL and LL genomes at resolutions of 1, 5, 10, and 25 kb. **j** Distribution charts summarizing the number and percentage of specific and shared chromatin loops. **k** Schematic diagrams illustrating how functional variants influence gene interactions through changes in proximal and distal RE activity and in TAD and loop structures. **l** Table detailing 2559 variant–gene interactions identified, categorized by interaction type, variant number, REs, and target genes involved. Source data are provided as a Source Data file.

Furthermore, our in-depth investigation into the transcription factor binding sites at the rs734209466 locus using Homer²⁸ and AnimalTFDB²⁹ revealed a specific *IRF4* binding site on the rs734209466–In allele (Fig. 7b). Illuminating this interaction, luciferase reporter gene analysis revealed that overexpression of *IRF4* significantly increased the luciferase activity of the enhancer carrying the rs734209466–In allele compared to the control group (cells transfected with the pCMV–HA empty vector) ($p < 0.0001$, Fig. 7c, d). However, there was no significant change in activity of the enhancer containing the rs734209466–Del allele ($p > 0.05$, Fig. 7c, d). To elucidate whether *IRF4* directly interacts with the rs734209466–In allele, we performed a supershift EMSA. The results showed that the biotin-labeled rs734209466–In probe formed complexes with *IRF4*–HA nuclear extracts, and a supershift band appeared upon the addition of anti–HA (Fig. 7e). In conclusion, our results confirmed that *IRF4* can directly bind to and regulate enhancer activity associated with the rs734209466–In allele.

Since rs734209466 can directly affect the transcription of the target genes *IGFBP2* and *IGFBP5*, we speculate that if a transcription factor is essential for rs734209466-related enhancer binding, its upregulation will promote the expression of the target genes *IGFBP2* and *IGFBP5*. The RT–qPCR results showed that compared with those in the control group (pCMV–HA), the mRNA expression levels of the target genes *IGFBP2* and *IGFBP5* in the *IRF4*–overexpressing group (pCMV–HA–*IRF4*) were significantly greater ($p < 0.0001$, Fig. 7f).

In summary, these findings suggested that the transcription factor *IRF4* likely plays a pivotal role in controlling enhancer activity in regions centered on rs734209466. The rs734209466–In allele promotes increased expression of the target genes *IGFBP2* and *IGFBP5* by *IRF4* binding.

***IGFBP2* and *IGFBP5* influence fat deposition by promoting preadipocyte proliferation and differentiation**

We established that the rs734209466–In allele, which is associated with an increase in AFD traits, can increase enhancer activity by the binding of *IRF4* transcription factor, thereby promoting the expression of the target genes *IGFBP2* and *IGFBP5*. Consequently, we were interested in determining whether alterations in the expression levels of the target genes *IGFBP2* and *IGFBP5* could influence cell phenotypes related to AFD traits. Initially, we analyzed the effects of silencing or overexpressing *IGFBP2* and *IGFBP5* on the proliferation of chicken preadipocytes. The increased proliferation assay results showed that *IGFBP2* and *IGFBP5* increased cell viability ($p < 0.0001$, Fig. 8a). RT–qPCR revealed that *IGFBP2* and *IGFBP5* increased the expression of the preadipocyte proliferation marker genes *PCNA* and *Ki67* ($p < 0.0001$, Fig. 8b). EdU staining results demonstrated that *IGFBP2* and *IGFBP5* promoted DNA synthesis in cells ($p \leq 0.0004$, Fig. 8c, d). Flow cytometry revealed that *IGFBP2* and *IGFBP5* significantly reduced the proportion of cells in the G1 phase and significantly increased the proportion of cells in the S and G2 phases ($p < 0.0001$, Fig. 8e, f and

Supplementary Fig. 9). These results suggested that *IGFBP2* and *IGFBP5* promote the proliferation of chicken preadipocytes.

Subsequently, we analyzed the effects of silencing or overexpressing *IGFBP2* and *IGFBP5* on the differentiation of chicken preadipocytes. Oil Red O staining and extraction colorimetry results showed that *IGFBP2* and *IGFBP5* promote lipid droplet deposition in preadipocytes ($p < 0.0001$, Fig. 8g, h). The RT–qPCR results showed that *IGFBP2* and *IGFBP5* promote the expression of the preadipocyte differentiation marker genes *PPARG*, *CEBPA*, and *AP2* ($p \leq 0.0002$, Fig. 8i). These results suggested that *IGFBP2* and *IGFBP5* promote the differentiation of chicken preadipocytes.

Discussion

Our research employs integrated 3D genomics with multi-omics analysis to systematically explore the key genetic variations and regulatory mechanisms influencing AFD traits in chickens. It comprehensively maps functional variations, regulatory elements, transcription factors, and the 3D genome structure involved in chicken fat traits, establishing complex genetic network structures (Fig. 9). Using rs734209466 as a case study, we demonstrate how non-coding variants impact the transcription of target genes by altering the 3D conformation of chromatin, ultimately influencing AFD traits in chickens (Fig. 9). These findings not only provide a perspective on understanding the molecular mechanisms underlying AFD traits but also offer crucial molecular markers for chicken genetic breeding.

The identification of functional variants and their target genes is a major challenge for noncoding variants, as they do not directly encode proteins³⁰. Our IMVGI model, which integrates multiomics datasets from genomics (WGS), epigenetics (ATAC–seq, ChIP–seq), 3D genomics (Hi–C), and transcriptomics (RNA–seq), established a variant–gene interaction network. This network reveals complex genetic regulatory patterns pervasive throughout the genome. Traditionally, studies have concentrated on gene regulation involving single–gene variants (one–to–one)³¹. However, emerging evidence suggests that a single variant can regulate multiple target genes (one–to–many)³², and a gene may be influenced by several variants (many–to–one)³³. The variant–gene interactions examined in this study suggested that each variant potentially affects an average of 1.94 genes, and an average of 4.23 functional variants likely regulate each gene. Furthermore, our functional elucidation of the IMVGI variant rs734209466 revealed its location in a specific enhancer region, enabling long–range regulation of *IGFBP2* and *IGFBP5* expression through loop structures without affecting the nearby genes *CXCR1* and *TNSI*, thus influencing AFD traits. For instance, the enhancer–bearing variant rs4810856 has been found to regulate the transcription of *PREX1*, *CSE1L*, and *STAU1* in colorectal cancer studies³². These results support the presence of complex genetic regulatory patterns, indicating that the phenomena of one–to–many and many–to–one interactions are widespread throughout the genome.

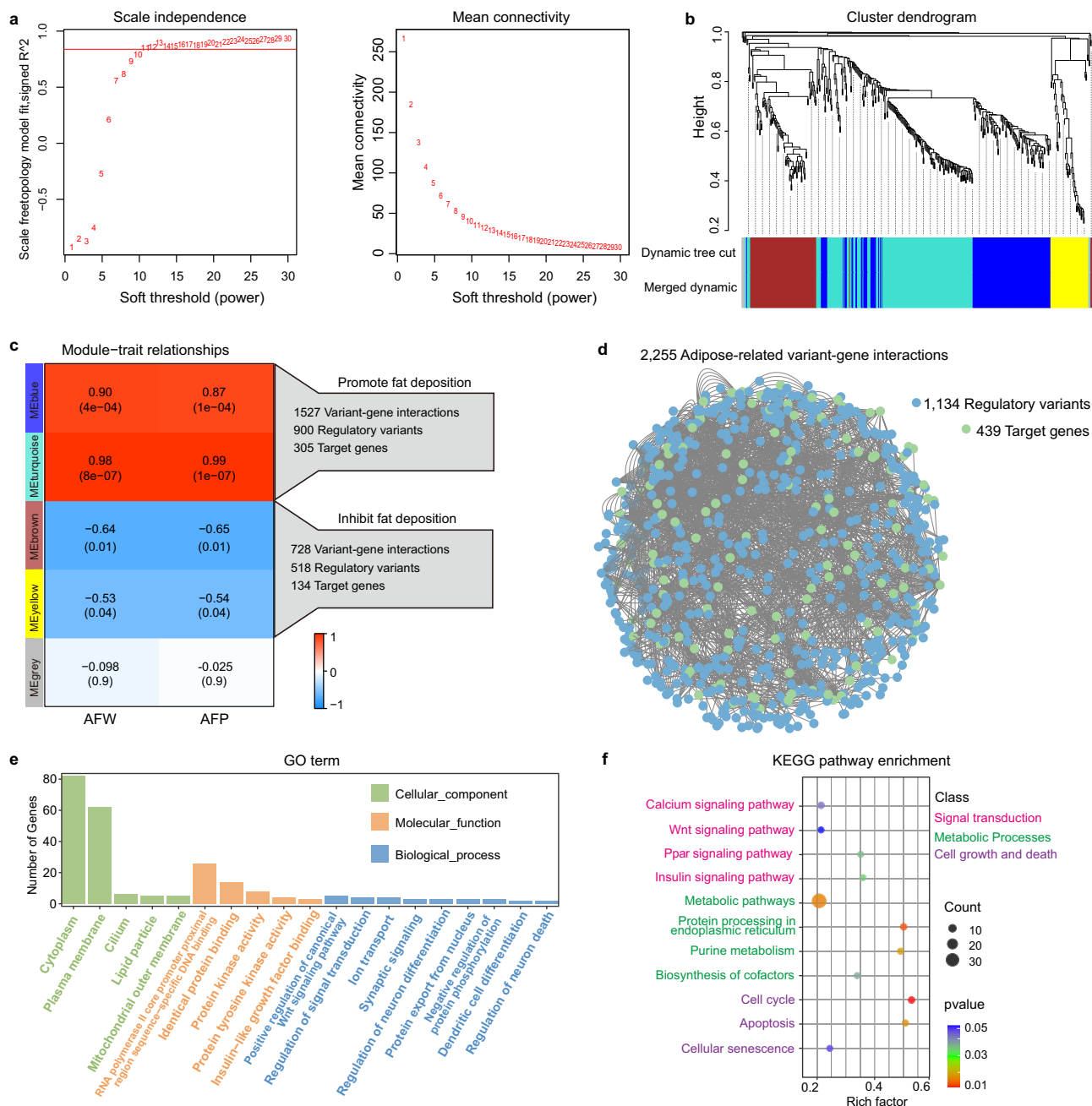
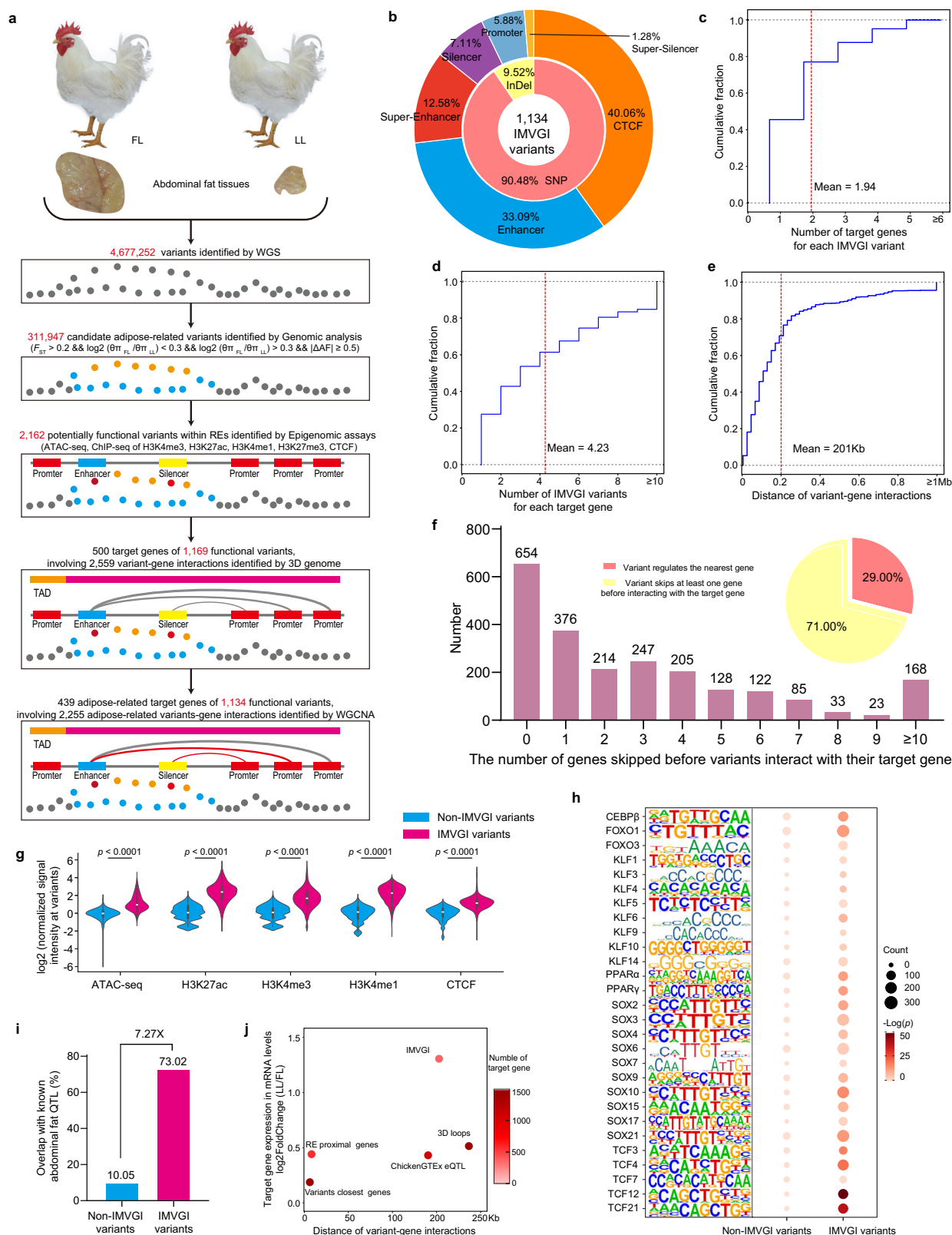


Fig. 4 | Weighted gene coexpression network analysis (WGCNA) and functional enrichment analysis. **a** Scale independence (left) and mean connectivity (right) analysis for choosing the soft threshold in establishing the WGCNA network. **b** Gene clustering and identification of gene modules using WGCNA. **c** Heatmap illustrating the relationships between gene modules and traits. The module-trait relationships were assessed by calculating the Pearson correlation coefficients (r) between the module eigengenes and trait data (AFW and AFP). P values were calculated based on the Pearson correlation test, reflecting the statistical significance

of the correlations. **d** Network diagram of adipose-related variant-gene interactions. **e** Gene Ontology (GO) term enrichment analysis for target genes. **f** Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of target genes, providing insight into biological pathways. KEGG pathway enrichment analysis was performed using a two-sided Fisher's exact test. Multiple comparisons were adjusted using the Benjamini-Hochberg (BH) method to control the false discovery rate (FDR). Source data are provided as a Source Data file.

Furthermore, the IMVGI model was used to explore the biological mechanisms underlying these functional variants. Our results indicate that key transcription factors involved in AFD traits, such as the KLF family³⁴, TCF family³⁵, CEBP family³⁶, and SOX family³⁷, are significantly enriched at IMVGI variant sites. This result suggests that these transcription factors regulate the expression of target genes by directly binding to IMVGI variant sites, thereby altering the activity of regulatory elements. For instance, studies have shown that transcription factors can modulate gene expression by binding to enhancer regions,

either facilitating or impeding the recruitment of RNA polymerase³⁸. Our experimental investigations using the IMVGI variant rs734209466 as a model further validated this mechanism. Motif analysis indicated that the rs734209466-In allele possesses a binding site for *IRF4*, whereas the rs734209466-Del allele lacks this binding site. *IRF4* is a transcription factor that has been reported to play a crucial role in adipose deposition³⁹. Our experimental results demonstrated that the *IRF4* transcription factor, by binding to the rs734209466-In allele, increases the activity of the specific enhancer region, impacting the



expression of the target genes *IGFBP2* and *IGFBP5*, which in turn influence adipocyte proliferation and differentiation, thereby promoting fat deposition. In summary, our results provide a clear understanding of the specific molecular mechanisms through which transcription factor-mediated variant-gene interactions influence the expression of specific genes.

This study utilized the innovative IMVGI method to construct a comprehensive genetic regulatory network map of functional variant sites that regulate AFD traits. These regulatory maps highlighted the significance of non-coding region variants in AFD traits and their key roles in trait-related genes. Experimental validation confirmed the functionality of the variant site rs734209466 as a specific enhancer,

Fig. 5 | Functional analysis of variant–gene interactions in relation to AFD traits. **a** Flowchart model for Integrative Multi-omics for Variant–Gene Interactions (IMVGI) identification of functional variants and target genes related to AFD traits. **b** Functional annotations for IMVGI variants. **c** Cumulative fractions of target genes per IMVGI variant (mean = 1.94). **d** Cumulative fractions of IMVGI variants per target gene (mean = 4.23). **e** Cumulative fractions of genomic distances for identified variant–gene connections (mean = 201 kb). **f** The bar chart indicates the number of genes skipped by variants before interacting with their target genes. The x-axis represents the number of genes skipped, and the y-axis represents the number of variants. **g** Comparative analysis of the enrichment of non-IMVGI and IMVGI variants in active chromatin regions, including ATAC-seq, H3K27ac, H3K4me3, H3K4me1, and CTCF markers. Violin plots provide a density estimation of the data distribution, and box plots display the median (central line), inter-quartile range (IQR, bounds), and whiskers representing the minimum and

maximum values. Sample sizes are $n = 1134$ biologically independent samples (unpaired two-tailed Student's t -test). **h** Heatmaps showing transcription factor binding preferences for non-IMVGI vs. IMVGI variants, with a greater binding preference of IMVGI variants for transcription factors involved in AFD traits, such as members of the KLF, TCF, CEBP, and SOX families. $n = 1134$ biologically independent samples. P values were calculated using hypergeometric tests and adjusted for multiple comparisons with the Benjamini-Hochberg method. **i** Bar charts comparing the enrichment of non-IMVGI vs. IMVGI variants in adipose-related QTLs with data from the Chicken QTLdb. **j** Scatter plot illustrating the relationship between genomic distance and the impact of variant–gene interactions on gene expression, highlighting the predictive power of the IMVGI model for genes with differential mRNA expression in abdominal fat tissue. Source data are provided as a Source Data file.

demonstrating its ability to interact with the transcription factor *IRF4* and regulate the transcription of *IGFBP2* and *IGFBP5* over long distances through a loop structure. *IGFBP2* and *IGFBP5* were found to stimulate the differentiation and proliferation of preadipocytes, thereby influencing AFD traits. These findings not only offer valuable targets for molecular breeding of AFD traits in chickens, but also present a research framework for understanding the genetic regulation mechanisms of important economic traits in agricultural animals.

Methods

Ethics

The Animal Ethics Committee of Northeast Agricultural University granted approval for the experimental designs and procedures (approval number: [NEAUEC20220228]). Rigorous measures were undertaken to assure the well-being and humane handling of the broiler chickens engaged in this study. The gathering of specimens and data was executed in conformity with the pertinent regulations and directives on animal welfare and protection.

Experimental population and phenotypic measurements

The study involved 330 7-week-old roosters (FL = 160, LL = 170) from the 19th generation of the NEAUHLF population. The chickens were raised under identical environmental conditions, with the temperature in the chicken coop maintained at 18–25 °C and humidity at 60–65%. Free access to food and water was provided throughout the study. The feed was a commercial corn–soybean–based diet meeting the National Research Council⁴⁰ recommendations for broiler nutritional requirements. From hatching to 3 weeks of age, the chicks were fed a starter diet (metabolizable energy [ME] 3000 kcal/kg; crude protein [CP] 210 g/kg), and from 4 weeks of age to slaughter, they were fed a grower diet (ME 3100 kcal/kg; CP 190 g/kg). At 7 weeks of age, just before slaughter, the BW7 of the chickens was measured. All chickens were slaughtered by cervical dislocation and jugular vein bleeding. The AFW was manually separated and weighed, and the AFP concentration was calculated.

Genomic analysis

Whole-genome sequencing and genomic variant calling.

Whole-genome sequencing was carried out on 330 roosters (FL = 160, LL = 170) from the NEAUHLF breed by Novogene (Beijing, China). This process generated 150 bp paired-end reads with an average sequencing depth of 6.62× coverage per sample (Supplementary Data 4). We assessed the quality of the base sequences using FastQC software (version 0.11.9)⁴¹. Next, we processed and removed low-quality bases and artifacts using Trim Galore software (version 0.6.7). High-quality 150 bp paired-end reads were aligned to the reference GRCh7b genome (https://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/016/699/485/GCF_016699485.2_bGalGal1.mat.broiler.GRCh7b/, last accessed May 21, 2021) using BWA-MEM software (version 0.7.17)⁴² with default

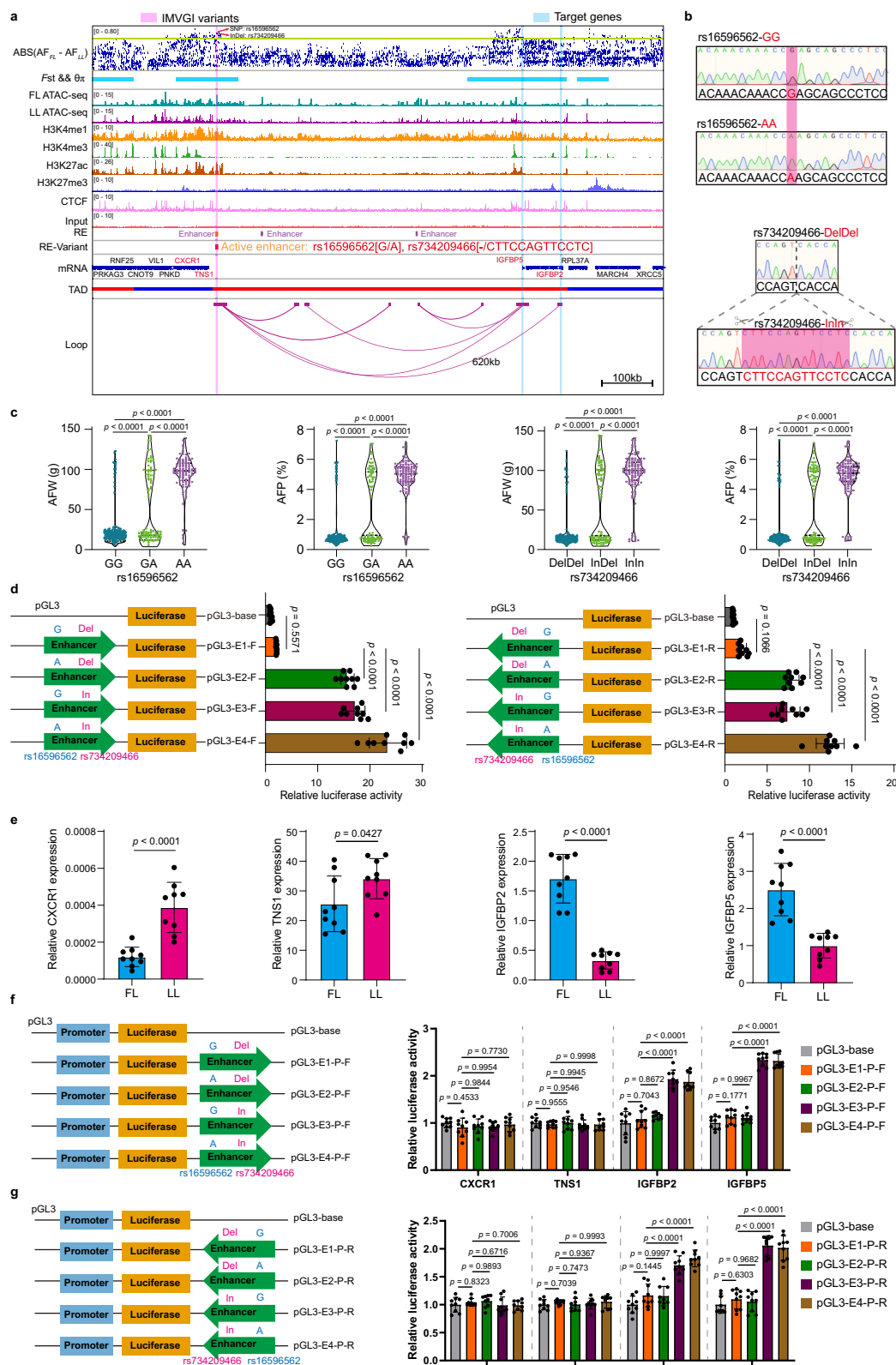
parameters (Supplementary Data 5). The mapped reads were converted to bam files and sorted using SAMtools software (version 1.12)⁴³. Duplicates were removed using the MarkDuplicates module in GATK software (version 4.2.6.1)⁴⁴. Genomic variants (SNPs and InDels) were identified using the GATK HaplotypeCaller module, adhering to GATK Best Practices⁴⁵. Finally, ANNOVAR software (version 2020Jun7)⁴⁶ was used for annotation of genomic variant positions. Genomic variants were categorized into noncoding and coding region variants based on their positions in the reference genome and gene position information for the reference genome.

Population structure analysis. To investigate the genetic relationship between the FL and LL populations, PCA was conducted using PLINK software (version 1.90)⁴⁷. Additionally, the population structure of these chicken populations was predicted using ADMIXTURE software (version 1.20)⁴⁸ as described by Shringarpure⁴⁹. LD decay analysis was performed using PopLDdecay software (version 3.41)⁵⁰, setting a reasonable extraction interval for candidate variants.

Selection signature analysis. The genome-wide distributions of the F_{ST} values and $\theta\pi$ ratios for FL and LL were calculated using VCFtools software (version 0.1.16)⁵¹, with a window size of 50 kb sliding in 10 kb steps. The $\theta\pi$ ratios were log₂-transformed. Windows with $F_{ST} > 0.2$ and $|\log_2(\theta\pi \text{ ratio})| > 0.3$ were identified as candidate adipose-related selection regions^{11,52,53}. The allele frequencies of genomic variants in these regions were calculated separately for FL and LL using VCFtools software. Genomic variants with a $\Delta AF \geq 0.5$ were identified as candidate variants associated with AFD traits^{11,52,53}.

Epigenomic analysis

ATAC-seq sequencing. Abdominal fat tissue samples from six NEAUHLF individuals (three from each line) were subjected to ATAC-seq by Frasergen (Wuhan, China). The ATAC-seq data analysis followed the pipeline from the ENCODE project (https://github.com/kundajelab/atac_dnase_pipelines)⁵⁴. Approximately 28.89 Gb of raw sequence data per sample were obtained, with base sequence quality assessed using FastQC software. Low-quality bases and artifacts were removed using Trim Galore software. High-quality 150 bp paired-end reads were aligned to the GRCh7b reference genome using BWA-MEM software. PCR duplicates and mitochondrial data were removed using SAMtools software. BAM files from three biological replicates were independently used for downstream analysis. Filtered BAM files were converted to bedpe format and then to tagAlign files using BEDTools software (version 2.30.0)⁵⁵. Peaks were called using MACS3 software (version 3.0.0b1)⁵⁶ with specified parameters, and differential analysis of peaks between groups was conducted using the edgeR package (version 4.0.2)⁵⁷ to identify DOCRs related to AFD traits. The detailed ATAC-seq data can be found in Supplementary Fig. 10 and 11 and Supplementary Tables 2 and 3.



ChIP-seq data analysis. Raw ChIP-seq data⁵⁸ for histone modifications (H3K27ac, H3K27me3, H3K4me1, H3K4me3) and CTCF, along with input controls, from chicken adipose tissue were obtained from the NCBI Gene Expression Omnibus (GEO; <https://www.ncbi.nlm.nih.gov/geo/>, accession number GSE158430). The raw data were remapped to the GRCg7b genome using BWA-MEM software. BAM file processing was consistent with the ATAC-seq sequencing

method described above. We merged BAM files from two biological replicates for downstream analysis. MACS3 software was used for peak calling, with parameters set for both broad and narrow peaks, following the ENCODE project pipeline (<https://github.com/ENCODE-DCC/chip-seq-pipeline>)⁵⁹. The detailed ChIP-seq data can be found in Supplementary Fig. 12.

Fig. 6 | Role of the IMVGI variant rs734209466 in enhancing *IGFBP2* and *IGFBP5* transcription. **a** Integrative Genomics Viewer (IGV) plot illustrating the regulatory role of candidate functional variants in AFD traits, as identified by IMVGI analysis. **b** Sanger sequencing validation results displaying the sequencing curves for SNP rs16596562 and InDel rs734209466, confirming the presence of these variants in the selected broiler chicken population. **c** Violin plot depicting phenotypic differences in AFD traits (AFW and AFP) among 330 NEAHLF individuals with various genotypes ($n = 136$ for rs16596562-GG, $n = 86$ for rs16596562-GA, $n = 108$ for rs16596562-AA, $n = 131$ for rs734209466-Del/Del, $n = 87$ for rs734209466-In/Del and $n = 112$ for rs734209466-In/In, biologically independent animals). P values were obtained using one-way ANOVA followed by Tukey's multiple comparisons test.

d Luciferase-based enhancer assay results demonstrating the enhancer activity of different alleles in the chicken preadipocyte cell line (ICP2). **e** RT-qPCR assessment of the mRNA expression levels of candidate target genes of IMVGI variants in the FL and LL of abdominal adipose tissue. P values were obtained using unpaired two-tailed Student's t -test. **f, g** Luciferase-based promoter assays were used to evaluate the transcriptional effects of enhancer alleles of the two IMVGI variants on the promoter regions of their target genes. **d, f, g** P values were obtained using one-way ANOVA followed by Dunnett's multiple comparisons test. **d, e, f, g** Data were presented as mean \pm SD, $n = 9$ biologically independent samples. Source data are provided as a Source Data file.

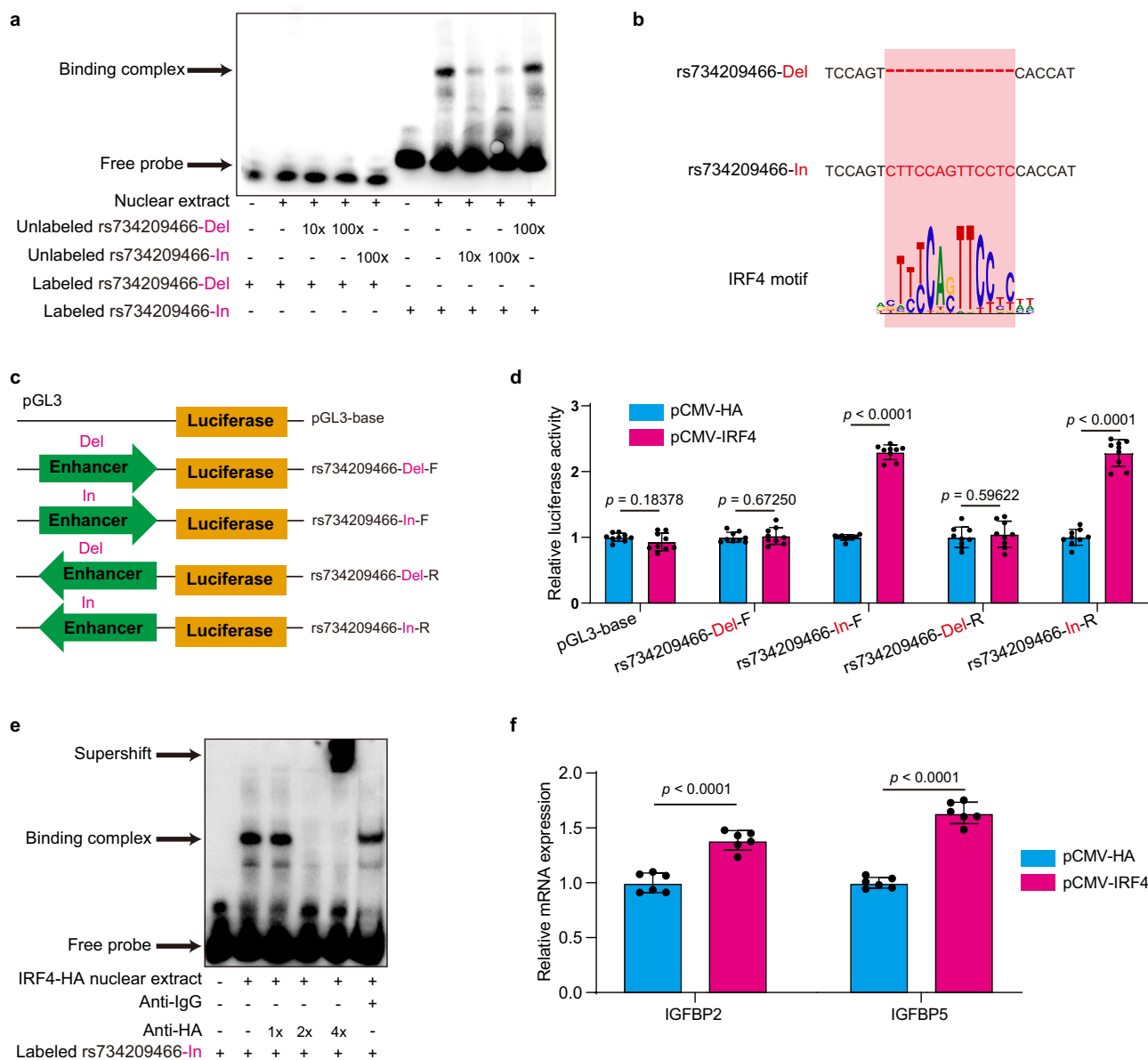
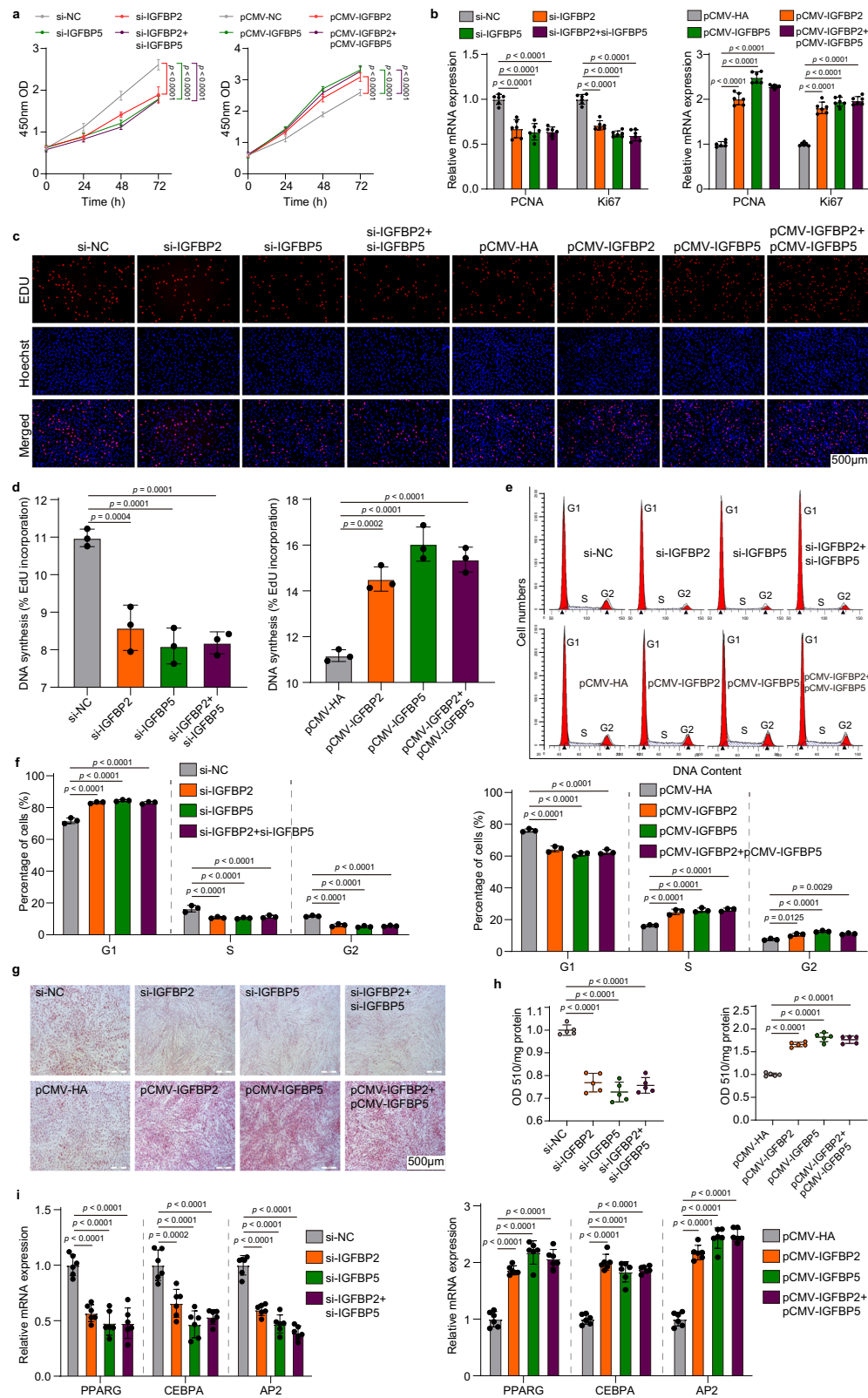


Fig. 7 | The transcription factor *IRF4* mediates differential allelic activity of rs734209466. **a** The results of the competitive EMSA of ICP2. The binding affinity of the protein to DNA oligos was demonstrated by adding unlabeled rs734209466-Del or unlabeled rs734209466-In to the reaction. **b** rs734209466-In resides within the *IRF4* binding motif predicted by both the Homer and AnimalTFDB4 databases. **c, d** Luciferase-based enhancer assays were performed to detect changes in enhancer activity before and after the overexpression of *IRF4* in

ICP2. Data were shown as the mean \pm SD, $n = 9$ biologically independent samples (unpaired two-tailed Student's t -test). **e** *IRF4* supershift EMSA of ICP2. 1 \times , 2 \times and 4 \times represent 1 μ g, 2 μ g and 4 μ g of HA antibody, respectively. **f** RT-qPCR analysis was utilized to detect the expression of *IGFBP2* and *IGFBP5* rs734209466-In/In genotyped preadipocytes, with and without the overexpression of *IRF4*. Data were shown as the mean \pm SD, $n = 6$ biologically independent samples (unpaired two-tailed Student's t -test). Source data are provided as a Source Data file.



Annotation of DOCR regions. DOCRs were classified and annotated using BEDTools software. Integrating ChIP-seq data for H3K27ac, H3K27me3, H3K4me1, H3K4me3, and CTCF, DOCRs were categorized into different regulatory elements: promoters (TSS -4 kb/+2 kb, H3K4me3, H3K27ac), enhancers (H3K4me1, H3K27ac), silencers (H3K27me3), and CTCF sites. We employed the ROSE algorithm²⁵ to annotate SEs and SSs within DOCRs. SEs were identified by stitching

enhancer-like regions with high ATAC-seq and H3K27ac enrichment⁶⁰, and SSs were annotated similarly, focusing on dense ATAC-seq and H3K27me3 modification regions²⁶.

Three-dimensional genome

Hi-C. We used abdominal fat tissues from the same six NEAHLF individuals (three from each line) selected for ATAC-seq analysis.

Fig. 8 | Effects of silencing and overexpression of *IGFBP2* and *IGFBP5* on fat deposition: preadipocyte proliferation and differentiation. **a** Cell viability was analyzed using the CCK-8 assay. *P* values were obtained using two-way ANOVA followed by Dunnett's multiple comparisons test, *n* = 6 biologically independent samples. **b** The mRNA levels of the preadipocyte proliferation markers *PCNA* and *Ki67* were assessed by RT-qPCR (*n* = 6 biologically independent samples). **c, d** Assessment of DNA synthesis in cells using EdU staining (scale bars: 500 μ m, *n* = 3 biologically independent samples). **e, f** The distribution of the cell cycle was

assessed by flow cytometry (*n* = 3 biologically independent samples).

g, h Measurement of preadipocyte lipid droplet accumulation by Oil Red O staining and spectrophotometric quantification (scale bars: 500 μ m, *n* = 5 biologically independent samples). **i** The mRNA levels of the preadipocyte differentiation markers *PPARG*, *CEBPA*, and *AP2* were assessed by RT-qPCR (*n* = 6 biologically independent samples). **b, d, f, h, i** Data were shown as the mean \pm SD, *p* values were obtained using one-way ANOVA followed by Dunnett's multiple comparisons test. Source data are provided as a Source Data file.

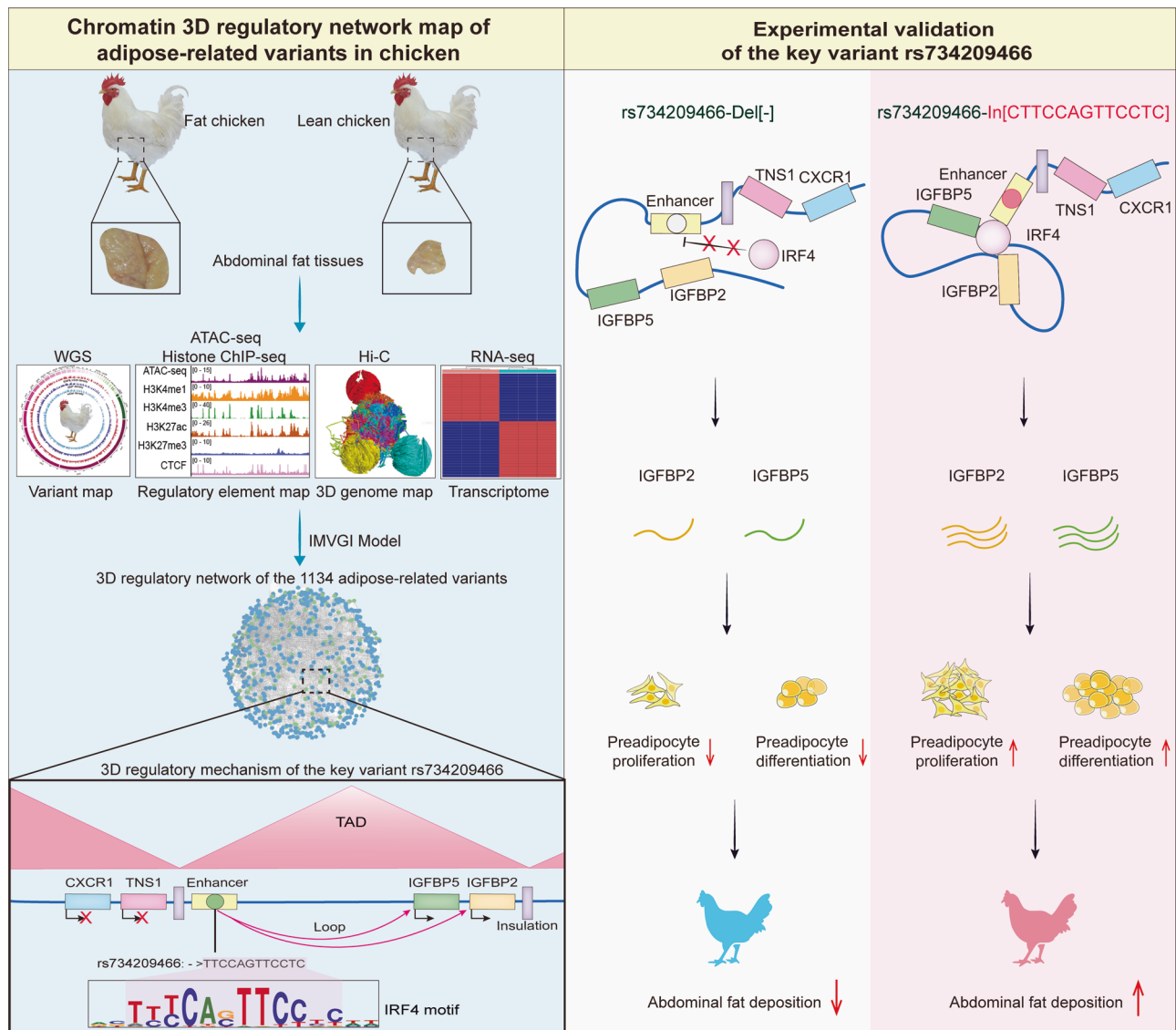


Fig. 9 | Schematic overview of the study. This study successfully established the comprehensive chromatin 3D regulatory network for AFD traits in chickens, integrating WGS, ATAC-seq, histone ChIP-seq, Hi-C, and RNA-seq. The network elaborates on the interactions among regulatory elements, super-regulatory elements, transcription factors, and gene transcription maps. Further, experimental validation of the key variant rs734209466 confirmed that it enhances

enhancer activity by interacting with *IRF4* and orchestrates the long-range regulation of the *IGFBP2* and *IGFBP5* genes through loop structures. This modulation promotes adipocyte proliferation and differentiation, thereby intensifying AFD traits. Importantly, this variant does not affect the expression of adjacent genes *CXCR1* and *TNS1*, demonstrating the specificity and precision of the regulatory network.

This sequencing was conducted by FraserGen (Wuhan, China), yielding an average sequencing depth of 231.99 \times per sample. To construct a more comprehensive Hi-C interaction matrix, we amalgamated the Hi-C data from three biological replicates of each line. The data were processed using the Juicer pipeline (<https://github.com/aidenlab/juicer/>)⁶¹. Contact matrices at varying resolutions (1, 5, 10, 25, 100, and 500 kb) were normalized using the KR

algorithm⁶¹. The detailed Hi-C sequencing data can be found in Supplementary Fig. 13, Supplementary Tables 1 and 4 and Supplementary Data 6.

Chromatin 3D modeling. The 3D genome structure was reconstructed from interchromosomal contacts (at 500 kb resolution) using the Python package miniMDS software (version 2018-09-27)⁶². PyMOL

software (version 2.5.0)⁶³ facilitated the visualization of these structures.

A/B compartment identification. We employed a PCA-based approach⁶⁴ for delineating the A and B compartments. Initially, observed/expected matrices were computed for each chromosome at 100 kb resolution using normalized and ICE-corrected interaction matrices. Subsequent steps included the computation of Pearson correlation and covariance matrices based on these matrices. PCA eigenvectors were then derived from the covariance matrices, using the first principal component (PC1) to assign the A and B compartments. This assignment was based on the direction of the eigenvalues, manually adjusted for gene and GC density, using hicPCA in HiCEXplorer software (version 3.6.0)^{65,66}.

TAD identification. TADs and boundaries were identified at 10 kb resolution using hicFindTADs in HiCEXplorer software^{65,66}. To ensure the conservation of TAD boundaries between the FL and LL groups, we considered boundaries overlapping or within a distance less than one bin size (10 kb) to be conserved⁶⁷.

Chromatin loop identification. Chromatin loops were identified using hicDetectLoops in HiCEXplorer software^{65,66}. Given the extensive loops observable in Hi-C contact maps, often spanning several megabases, we called loops at various resolutions (1, 5, 10, 25 kb). These loops, identified across all resolutions, were then consolidated within a 25 kb range using hicMergeLoops in HiCEXplorer software. For loop conservation between the FL and LL groups, loops with anchor genomic intervals overlapping or within less than one bin size (25 kb) were considered conserved⁶⁷.

Transcriptome

In this part of the study, abdominal fat tissues were collected from ten selected chickens (five from each line, FL and LL) from the initial 330-chicken cohort. The selection was random, ensuring representativeness. These samples were subjected to RNA-seq by Novogene (Beijing, China). Alignment to the GRCg7b chicken genome was accomplished using HISAT2 software (version 2.0.1)⁶⁸. BAM files from five biological replicates were independently used for downstream analysis. Read counts were obtained using FeatureCounts software (version 2.0.1)⁶⁹, and gene expression was assessed based on TPM values. WGCNA was conducted using the WGCNA R package (version 1.69)⁷⁰ to identify gene networks and pathways pertinent to adiposity. The detailed RNA-seq data can be found in Supplementary Fig. 14 and Supplementary Tables 5 and 6.

Identification of target genes associated with functional variants

To identify target genes regulated by functional variants, we conducted a comprehensive analysis integrating epigenomic genomic annotations and 3D genome structure. The methodology involved the following steps:

Mapping variants to promoter regions and proximal REs. We used BEDTools software to intersect functional variants with the genomic coordinates of promoters and nearby regulatory elements. Variants located within promoter regions (TSS -4 kb/+2 kb) or within 10 kb of other REs (enhancers and silencers) were associated with their proximal genes based on genomic proximity.

Linking variants in REs to distal target genes via chromatin interactions. For variants within REs that might affect distal genes, we utilized chromatin interaction data to capture long-range regulatory relationships. Using PGLTOOLS software (version 3.0.0)⁷¹, we mapped these REs to Hi-C interaction anchor points, identifying regulatory elements overlapping with these anchors. We then identified target

genes by determining whether the interacting anchor points mapped to promoter regions of protein-coding genes in the GRCg7b genome assembly (TSS -4 kb/+2 kb).

Identifying variants affecting TAD structures at CTCF binding sites.

We investigated functional variants located at CTCF binding sites coinciding with TAD boundaries. Using PGLTOOLS software, we identified CTCF binding sites overlapping with TAD boundaries defined from our Hi-C data. Variants within these CTCF sites were considered potential modifiers of TAD structures. To identify target genes affected by changes in TAD structures, we analyzed genes located within the TADs altered by these variants, as modifications in TAD boundaries can influence gene regulation within these domains.

Assessing variants influencing chromatin loop structures at CTCF binding sites.

We also analyzed functional variants at CTCF binding sites located at loop anchor points, using PGLTOOLS software. Variants at these sites may disrupt or alter chromatin loop formation, impacting the regulation of distal genes. We mapped these variants to loop anchor points and identified target genes by determining whether the interacting anchor points overlapped with gene promoters.

Quantifying variant-gene interactions. For each variant, we calculated the number of gene promoter regions interacting with that variant through chromatin loops or altered chromatin structures. Conversely, for each gene promoter region, we quantified the number of independent interactions involving at least one functional variant. This analysis enabled us to construct a comprehensive network of potential regulatory interactions mediated by chromatin architecture modifications.

Functional enrichment analysis

The GO term and KEGG pathway analyses were performed using the DAVID database (<http://david.abcc.ncifcrf.gov/>, last accessed on June 5, 2022)⁷². Statistical significance was set at a *p* value < 0.05. GSEA was conducted via the omicstudio online platform (<https://www.omicstudio.cn/tool/123>, last accessed on June 5, 2022)⁷³.

Adipocyte culture

The immortalized chicken preadipocyte (ICP2) cell line was previously developed in our laboratory⁷⁴. All cell lines were authenticated through nucleic acid sequence analysis to confirm their identities. They were tested negative for Mycoplasma using the Myco-Blue Mycoplasma Detector Kit (Vazyme, D101-01). ICP2 cells were cultured in DMEM/F12 medium containing 10% serum and 1% penicillin-streptomycin in an incubator at 37 °C with 5% CO₂. When the cells reached approximately 90% confluence, they were washed with PBS buffer and passaged using trypsin (Sigma, MO, USA).

For the induction of differentiation into mature adipocytes, cells were plated in 12-well culture plates. Once the cells reached approximately 50% confluence, the basal medium was replaced with an induction medium containing oleic acid (300 μM). The oleic acid induction medium was refreshed daily during the induction process.

Synthesis of plasmids and siRNA

The eukaryotic gene expression vector pCMV-HA was preserved in our laboratory. The expression plasmids for *IGFBP2*, *IGFBP5*, and *IRF4* were constructed by inserting the respective cDNA fragments into the pCMV-HA vector. The primer sequences used for constructing the overexpression vectors were synthesized by RiboBio (Guangzhou, China), with detailed information provided in Supplementary Data 7.

The luciferase reporter gene vector pGL3-Promoter was preserved in our laboratory. Enhancer sequences surrounding the rs16596562 and rs734209466 sites were downloaded from the NCBI

database. Promoter sequences of *CXCR1*, *TNSI*, *IGFBP2*, and *IGFBP5* were similarly downloaded from the NCBI database. These sequences were inserted upstream of the SV40 promoter in the pGL3-Promoter vector. Mutant constructs were generated by site-specific mutagenesis at the rs16596562 (G > A) and the rs734209466 (Del > In), and were cloned using the same strategy as for the wild-type sequences. The promoter sequences of *CXCR1*, *TNSI*, *IGFBP2*, and *IGFBP5* were inserted downstream of the SV40 promoter. The primer sequences for the dual-luciferase reporter gene assay were synthesized by RiboBio (Guangzhou, China), with detailed information provided in Supplementary Data 8.

The sequences of small interfering RNA (siRNA) oligonucleotides targeting *IGFBP2*, *IGFBP5*, and the negative control siRNA are listed by RiboBio (Guangzhou, China), with detailed information provided in Supplementary Table 7.

Cell transfection

Transfection of cells was performed according to the manufacturer's protocol for Lipofectamine 3000 reagent (Invitrogen, CA, USA).

Oil red O staining and quantification

The cells were washed three times with PBS and fixed in 4% formaldehyde for 30 minutes. Subsequently, the cells were stained with 60% oil red O solution (Sigma, MO, USA) for 15 minutes. After staining, the cells were washed, and 60% isopropanol was added for lipid extraction. The lipid droplets were then visualized under a microscope. For quantification, oil red O was dissolved in 100% isopropanol, and the absorbance was measured at 510 nm. The accumulation of lipid droplets was normalized to the protein content, which was measured using a BCA Protein Quantitative Analysis Kit (Beyotime, Nanjing, China).

Dual luciferase reporter assay

For functional validation, we utilized a dual luciferase reporter assay system (Promega, WI, USA) following the manufacturer's protocol. Constructs with different allelic sequences were cloned and inserted into the pGL3-promoter vector (Promega, WI, USA). Cells (seeded in 24-well plates) were transfected with Lipofectamine 3000 (Invitrogen, CA, USA), after which firefly and Renilla luciferase activity was measured after 24 hours.

Electrophoretic mobility shift assay (EMSA)

Single-stranded complementary oligonucleotide sequences containing variant rs734209466 alleles were synthesized by RiboBio (Guangzhou, China) and labeled with biotin at the 5' end (Supplementary Table 8). Nuclear extracts from ICP2 cells were prepared using the Nuclear and Cytoplasmic Protein Extraction Kit (Thermo Fisher Scientific, USA). Electrophoretic mobility shift assays (EMSA) were conducted with the EMSA/Gel-Shift Kit (Thermo Fisher Scientific, USA) according to the manufacturer's instructions. For the competitive binding assay, unlabeled probes were added to the reaction mixtures at 10-fold or 100-fold excess relative to the labeled probes and incubated for 20 minutes prior to the addition of the labeled probes. For super-shift EMSA, 1 µg, 2 µg, and 4 µg of anti-HA antibody (Abcam; Cat: ab9110) were added to the reaction mixtures and incubated for 20 minutes before the addition of labeled probes. As a control for non-specific binding, 2 µg of normal rabbit IgG (Cell Signaling Technology; Cat: 2729S) was added to parallel reaction mixtures under the same conditions. The binding products were detected using a streptavidin-horseradish peroxidase conjugate in accordance with the SuperSignal West Femto Trial Kit (Thermo Fisher Scientific, USA).

Real-time quantitative PCR (RT-qPCR)

RNA was extracted with TRIzol reagent (Takara, Dalian, China). Reverse transcription was performed with the Evo M-MLV Mix Kit with gDNA

Clean for qPCR (Accurate Biology, Hunan, China). RT-qPCR was conducted on a QuantStudio 5 system (Applied Biosystems) using a SYBR Green Premix Pro Taq HS qPCR Kit (Accurate Biology, Hunan, China). *TBP* was used as an internal control. The primers used are listed in Supplementary Table 9.

Cell proliferation assays

Cell proliferation was assessed using a Cell Counting Kit-8 (CCK-8, Dojindo, Kumamoto, Japan) and an EdU incorporation assay (RiboBio, Guangzhou, China). For the CCK-8 assay, cells were seeded into 96-well plates at a density of 1×10^4 cells per well in 100 µL of culture medium, with three replicates for each transfection group. At designated time points post-transfection, 10 µL of CCK-8 reagent was added to each well and incubated at 37 °C for 2 hours. Absorbance at 450 nm was measured using a microplate reader (Molecular Devices, Sunnyvale, CA, USA). For the EdU incorporation assay, cells were plated in 24-well plates at a density of 5×10^4 cells per well in 500 µL of culture medium, with three replicates for each transfection group. Cell staining was conducted using the EdU Cell Proliferation Detection Kit following the manufacturer's protocol. Stained cells were observed under a fluorescence microscope (Leica, Wetzlar, Germany). Three fields were randomly selected from each well, and the percentage of EdU-positive cells was calculated using ImageJ 1.46 R software (NIH, Bethesda, MD, USA). The percentage of EdU-positive cells was determined by the ratio of red fluorescent cells to blue fluorescent cells, multiplied by 100%.

Cell cycle analysis

Cell cycle analysis was performed using a Cell Cycle Staining Kit (Multisciences, Hangzhou, China). Cells cultured in 12-well plates were harvested and centrifuged at 800 g for 5 minutes. After discarding the supernatant, the cells were washed once with cold phosphate-buffered saline (PBS). The cells were then resuspended in 1 mL of reagent A from the kit and 10 µL of reagent B, followed by vortexing for 10 seconds and incubation at room temperature for 30 minutes. The cell suspension was subsequently subjected to flow cytometry analysis using a FACS Canto™ II system (BD BioSciences, San Jose, CA, USA).

Statistical analysis

Statistical analysis was performed using R 4.0.3 and GraphPad 8.0.2. Data normality was assessed using the Anderson-Darling test. For normally distributed data, significant differences were determined using Student's *t*-test for pairwise comparisons, or by analysis of variance (ANOVA) for three or more groups, followed by Tukey or Dunnett's multiple comparisons test for further analysis. For non-normally distributed data, the Mann-Whitney U test was used for pairwise comparisons, while the Kruskal-Wallis test was employed for three or more groups, followed by Dunn's post hoc test for multiple comparisons. $P < 0.05$ was considered statistically significant.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The ATAC-seq and Hi-C data generated in this study have been deposited in the GEO database under accession numbers [GSE255760](#) and [GSE260559](#), respectively. Additionally, the WGS and RNA-Seq data used in this study are available in the NCBI BioProject database under accession numbers [PRJNA353057](#) and [PRJNA354990](#). The ChIP-seq data utilized in this study were obtained from the GEO database under accession number [GSE158430](#). Source data are provided with this paper.

Code availability

All the code used for data analysis is available at <https://github.com/shenlinyong/IMVGI>. All codes have been deposited in Zenodo at: <https://doi.org/10.5281/zenodo.13902538>⁷⁵.

References

- Uzundumlu, A. S. & Dilli, M. Estimating chicken meat productions of leader countries for 2019–2025 years. *Ciência Rural* **53**, e20210477 (2022).
- Chen, C. et al. Estimation of the genetic parameters of traits relevant to feed efficiency: result from broiler lines divergent for high or low abdominal fat content. *Poultry Science* **100**, 461–466 (2021).
- Baéza, E., Guillier, L. & Petracci, M. Production factors affecting poultry carcass and meat quality attributes. *Animal* **16**, 100331 (2022).
- Zhang, X. et al. Genetic selection on abdominal fat content alters the reproductive performance of broilers. *Animal* **12**, 1232–1241 (2018).
- Mellouk, N. et al. Chicken Is a Useful Model to Investigate the Role of Adipokines in Metabolic and Reproductive Diseases. *International Journal of Endocrinology* **2018**, 4579734 (2018).
- Ji, B. et al. Dynamic regulation of adipose tissue metabolism in the domestic broiler chicken – an alternative model for studies of human obesity. *BMC Proceedings* **6**, P67 (2012).
- Dupont, J., Tesserand, S. & Simon, J. Insulin signaling in chicken liver and muscle. *General and Comparative Endocrinology* **163**, 52–57 (2009).
- Dhurandhar, N. V., Kulkarni, P. R., Ajinkya, S. M., Sherikar, A. A. & Atkinson, R. L. Association of adenovirus infection with human obesity. *Obesity Research* **5**, 464–469 (1997).
- Nadaf, J. et al. QTL for several metabolic traits map to loci controlling growth and body composition in an F2 intercross between high- and low-growth chicken lines. *Physiological Genomics* **38**, 241–249 (2009).
- Chua, E. H. Z., Yasar, S. & Harmston, N. The importance of considering regulatory domains in genome-wide analyses—the nearest gene is often wrong! *Biol. Open* **11**, bio059091 (2022).
- Li, J. et al. Enhancer-promoter interaction maps provide insights into skeletal muscle-related traits in pig genome. *BMC Biol* **20**, 136 (2022).
- Teng, J. et al. A compendium of genetic regulatory effects across pig tissues. *Nat. Genet.* **56**, 112–123 (2024).
- Guan, D. et al. The ChickenGTEx pilot analysis: a reference of regulatory variants across 28 chicken tissues. *bioRxiv* 2023.06.27.546670 (2023).
- Liu, S. et al. A multi-tissue atlas of regulatory variants in cattle. *Nat. Genet.* **54**, 1438–1447 (2022).
- Stikker, B. S., Hendriks, R. W. & Stadhouders, R. Decoding the genetic and epigenetic basis of asthma. *Allergy* **78**, 940–956 (2023).
- Jin, L. et al. Dynamic chromatin architecture of the porcine adipose tissues with weight gain and loss. *Nat. Commun.* **14**, 3457 (2023).
- Li, D. et al. Dynamic transcriptome and chromatin architecture in granulosa cells during chicken folliculogenesis. *Nat. Commun.* **13**, 131 (2022).
- Aneas, I. et al. Asthma-associated genetic variants induce IL33 differential expression through an enhancer-blocking regulatory region. *Nature Communications* **12**, 6115 (2021).
- Joslin, A. C. et al. A functional genomics pipeline identifies pleiotropy and cross-tissue effects within obesity-associated GWAS loci. *Nature Communications* **12**, 5253 (2021).
- Sobreira, D. R. et al. Extensive pleiotropism and allelic heterogeneity mediate metabolic effects of IRX3 and IRX5. *Science* **372**, 1085–1091 (2021).
- Liu, Y. et al. Integration of multi-omics data reveals cis-regulatory variants that are associated with phenotypic differentiation of eastern from western pigs. *Genetics Selection Evolution* **54**, 62 (2022).
- Dareng, E. O. et al. Integrative multi-omics analyses to identify the genetic and functional mechanisms underlying ovarian cancer risk regions. *Am J Hum Genet* **111**, 1061–1083 (2024).
- Guo, L. et al. Comparison of adipose tissue cellularity in chicken lines divergently selected for fatness. *Poult. Sci.* **90**, 2024–2034 (2011).
- Hu, Z. L., Park, C. A. & Reecy, J. M. Bringing the animal QTLdb and CorrDB into the future: meeting new challenges and providing updated services. *Nucleic Acids Res* **50**, D956–D961 (2022).
- Whyte, W. A. et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319 (2013).
- Zhang, Y. et al. Super-silencer perturbation by EZH2 and REST inhibition leads to large loss of chromatin interactions and reduction in cancer growth. *bioRxiv* 2023.08.29.555291 (2023).
- Lindsly, S. et al. 4DNvestigator: time series genomic data analysis toolbox. *Nucleus* **12**, 58–64 (2021).
- Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
- Shen, W. K. et al. AnimalTFDB 4.0: a comprehensive animal transcription factor database updated with variation and expression annotations. *Nucleic Acids Res* **51**, D39–D45 (2023).
- Wu, K. et al. Exploring noncoding variants in genetic diseases: from detection to functional insights. *J. Genet. Genom.* **51**, 111–132 (2024).
- Zhang, Z. et al. The rs1421085 variant within FTO promotes brown fat thermogenesis. *Nat. Metab.* **5**, 1337–1351 (2023).
- Ying, P. et al. Genome-wide enhancer-gene regulatory maps link causal variants to target genes underlying human cancer risk. *Nat. Commun.* **14**, 5958 (2023).
- Feng, Y. et al. Integrative functional genomic analyses identify genetic variants influencing skin pigmentation in Africans. *Nat. Genet.* **56**, 258–272 (2024).
- García-Niño, W. R. & Zazueta, C. New insights of Krüppel-like transcription factors in adipogenesis and the role of their regulatory neighbors. *Life Sci* **265**, 118763 (2021).
- Fehrenschild, D. et al. TCF/Lef1-mediated control of lipid metabolism regulates skin barrier function. *J. Invest. Dermatol.* **132**, 337–345 (2012).
- Chatterjee, R. et al. Suppression of the C/EBP family of transcription factors in adipose tissue causes lipodystrophy. *J. Mol. Endocrinol.* **46**, 175 (2011).
- Shen, H. et al. SOX4 promotes beige adipocyte-mediated adaptive thermogenesis by facilitating PRDM16-PPAR γ complex. *Theranostics* **12**, 7699 (2022).
- Suter, D. M. Transcription factors and DNA play hide and seek. *Trends Cell Biol* **30**, 491–500 (2020).
- Guo, S. et al. Metabolic crosstalk between skeletal muscle cells and liver through IRF4-FSTL1 in nonalcoholic steatohepatitis. *Nat. Commun.* **14**, 6047 (2023).
- National Research Council. *Nutrient Requirements of Poultry: 1994*, (National Academies Press, 1994).
- Brown, J., Pirrung, M. & McCue, L. A. FQC dashboard: integrates FastQC results into a web-based, interactive, and extensible FASTQ quality control tool. *Bioinformatics* **33**, 3137–3139 (2017).

42. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
43. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *Giga-science* **10**, giab008 (2021).
44. DePristo, M. A. et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491–498 (2011).
45. Van Der Auwera, G. A. et al. From FastQ data to high confidence variant calls: the genome analysis toolkit best practices pipeline. *Curr. Protoc. Bioinform.* **43**, 11.10.1–11.10.33 (2013).
46. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* **38**, e164 (2010).
47. Chang, C. C. et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, S13742–015–0047–8 (2015).
48. Alexander, D. H. & Lange, K. Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinform* **12**, 246 (2011).
49. Shringarpure, S. S., Bustamante, C. D., Lange, K. & Alexander, D. H. Efficient analysis of large datasets and sex bias with ADMIXTURE. *BMC Bioinform* **17**, 218 (2016).
50. Zhang, C., Dong, S. S., Xu, J.-Y., He, W. M. & Yang, T. L. PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics* **35**, 1786–1788 (2019).
51. Danecek, P. et al. The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
52. Zhou, Z. et al. An intercross population study reveals genes associated with body size and plumage color in ducks. *Nat. Commun.* **9**, 2648 (2018).
53. Zhao, P. et al. Evidence of evolutionary history and selective sweeps in the genome of Meishan pig reveals its genetic and phenotypic characterization. *Gigascience* **7**, giy058 (2018).
54. Kim, D. S. ATAC-seq data processing. *Chromatin Access. Methods Protoc.* **2611**, 305–323 (2023).
55. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
56. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137 (2008).
57. Chen, Y., Lun, A. T. & Smyth, G. K. From reads to genes to pathways: differential expression analysis of RNA-Seq experiments using Rsubread and the edgeR quasi-likelihood pipeline. *F1000Res* **5**, 1438 (2016).
58. Kern, C. et al. Functional annotations of three domestic animal genomes provide vital resources for comparative and agricultural research. *Nat. Commun.* **12**, 1821 (2021).
59. Landt, S. G. et al. ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Res* **22**, 1813–1831 (2012).
60. Zhao, S. et al. Genome-wide chromatin interaction profiling reveals vital role of super-enhancers and rearrangements in host enhancer contacts during BmNPV infection. *Genome Res* **33**, 1958–1974 (2023).
61. Durand, N. C. et al. Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Syst* **3**, 95–98 (2016).
62. Rieber, L. & Mahony, S. miniMDS: 3D structural inference from high-resolution Hi-C data. *Bioinformatics* **33**, i261–i266 (2017).
63. DeLano, W. L. Pymol: an open-source molecular graphics tool. *CCP4 Newsl. Protein Crystallogr.* **40**, 82–92 (2002).
64. Dixon, J. R. et al. Chromatin architecture reorganization during stem cell differentiation. *Nature* **518**, 331–336 (2015).
65. Wolff, J. et al. Galaxy HiCExplorer: a web server for reproducible Hi-C data analysis, quality control and visualization. *Nucleic Acids Res* **46**, W11–W16 (2018).
66. Wolff, J., Backofen, R. & Grüning, B. Loop detection using Hi-C data with HiCExplorer. *Gigascience* **11**, giac061 (2022).
67. Liao, Y. et al. The 3D architecture of the pepper genome and its relationship to function and evolution. *Nat. Commun.* **13**, 3479 (2022).
68. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nature Methods* **12**, 357–360 (2015).
69. Liao, Y., Smyth, G. K. & Shi, W. FeatureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
70. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform* **9**, 559 (2008).
71. Greenwald, W. W. et al. Pgltools: a genomic arithmetic tool suite for manipulation of Hi-C peak and other chromatin interaction data. *BMC Bioinformatics* **18**, 207 (2017).
72. Huang, D. W. et al. DAVID bioinformatics resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Res* **35**, W169–W175 (2007).
73. Lyu, F. et al. OmicStudio: A composable bioinformatics cloud platform with real-time feedback that can generate high-quality graphs for publication. *iMeta* **2**, e85 (2023).
74. Wang, W. et al. Immortalization of chicken preadipocytes by retroviral transduction of chicken TERT and TR. *PLoS One* **12**, e0177348 (2017).
75. Shen, L. et al. Replication code: Integrative 3D genomics with multi-omics analysis and functional validation of genetic regulatory mechanisms of fat deposition in chickens. *Zenodo* <https://doi.org/10.5281/zenodo.13902538> (2024).

Acknowledgements

This research was funded by the National Key R&D Program of China (Grant No. 2021YFD1300100 to H.Z.), the National Natural Science Foundation of China (Grant No. 32372868 to H.Z.), the STI2030–Major Projects (Grant No. 2023ZD0406404 to H.Z.), the earmarked fund for CARS–41 (Grant No. CARS–41 to H.Z.), and the Heilongjiang Postdoctoral Financial Assistance program (Grant No. LBH–Q20006 to H.Z.). The funding body did not have any influence on the design, analysis or reporting of the study.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-53692-6>.

Correspondence and requests for materials should be addressed to Hui Li or Hui Zhang.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

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

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024