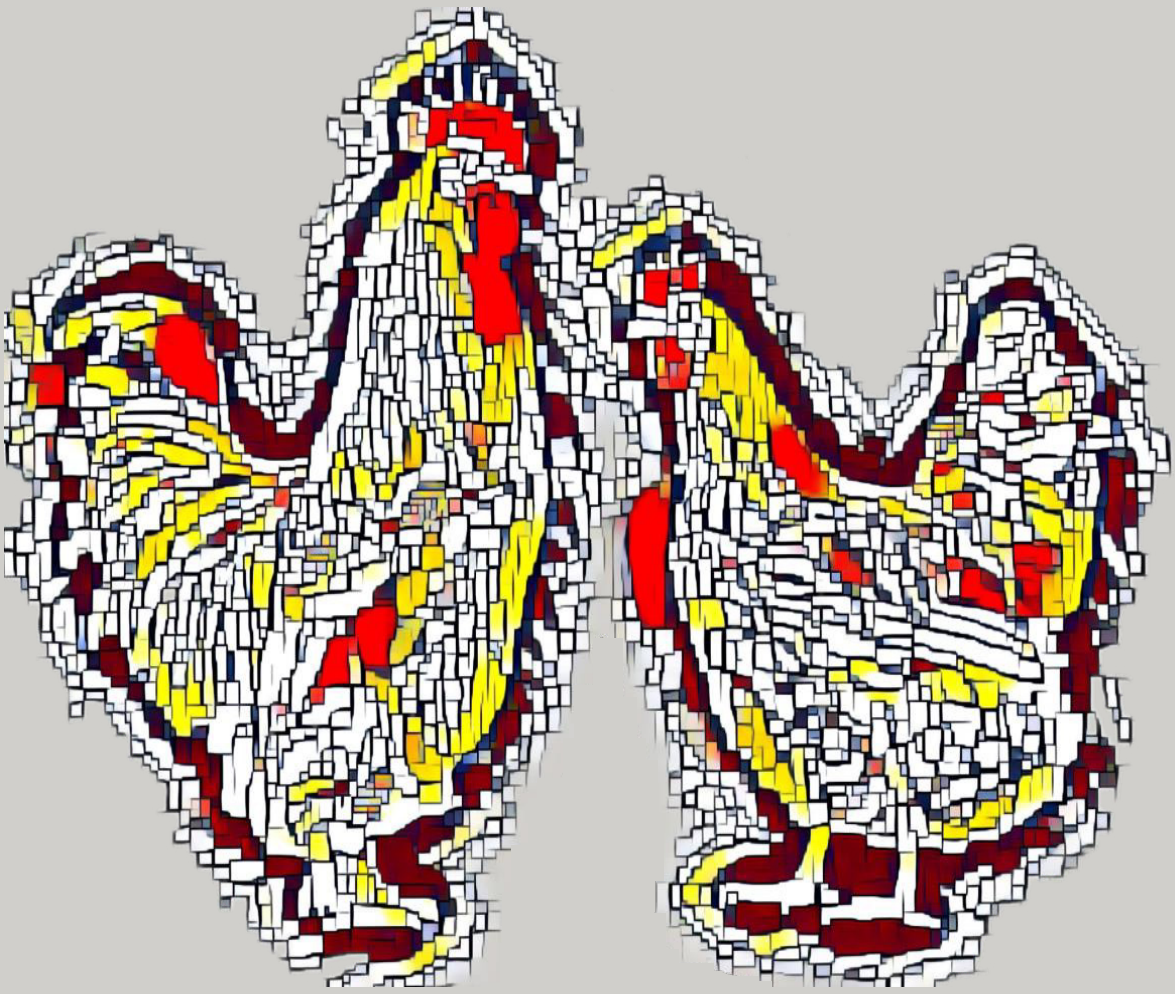
Differential Deposition of Intramuscular and Abdominal Fat in Chicken

邢思远
Siyuan Xing



Propositions

1. In the current chicken industry, increased body weight selection, as a proxy for maximising production, is useless.

   (this thesis)

2. The poor accuracy in measuring the phenotype of intramuscular fat in chicken prevents selection on it.

   (this thesis).

3. Participatory science is key in the advancement of animal breeding.

4. Perfectionism is the best friend for a paper revision but not the writing itself.

5. Insufficient mathematical knowledge is the biggest limiting factor of data analysts.

6. Only considering the first author in scientific papers as a personal academic achievement destroys scientific cooperation.

7. Cyber war has become more potent in disrupting the world than nuclear war.

8. Rat racing is good for employers but not good for the society.

Propositions belonging to the thesis, entitled:

Differential Deposition of Intramuscular and Abdominal Fat in Chicken

Siyuan Xing

Wageningen, 13 October 2021

Differential Deposition of Intramuscular and Abdominal Fat in Chicken

Siyuan Xing

Thesis committee

Thesis supervisor
Prof. Dr M.A.M. Groenen
Professor of Animal Breeding and Genomics
Wageningen University & Research

Thesis co-supervisors
Prof. Dr J. Wen
Professor at the Institute of Animal Sciences
Chines Academy of Agricultural Sciences (CAAS)

Dr R.P.M.A. Crooijmans
Assistant professor, Animal Breeding and Genomics
Wageningen University & Research

Prof. Dr R. Liu
Professor at the Institute of Animal Sciences
Chines Academy of Agricultural Sciences (CAAS)

Other members
Prof. Dr A.H. Kersten, Wageningen University & Research
Prof. Dr M. Siwek, University of Technology and Life Sciences in Bydgoszcz
Dr J.A.M. van Arendonk, Hendrix Genetics
Prof Dr R.P. i Subirà, Universitat de Lleida

This research was conducted under the auspices of the Institute of Animal Sciences (IAS) of Chinese Academy of Agricultural Sciences (CAAS) and Graduate School of Wageningen Institute of Animal Sciences (WIAS).

Differential Deposition of Intramuscular and Abdominal Fat in Chicken

Siyuan Xing

Thesis

submitted in fulfillment of the requirements for the degree of doctor at
Wageningen University
by the authority of the Rector Magnificus,
Prof. Dr A.P.J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Wednesday 13 October 2021
at 11 a.m. in the Aula.

Xing, S.
Differential Deposition of Intramuscular and Abdominal Fat in Chicken

PhD thesis, Wageningen University, the Netherlands (2021)
With references, with summary in English

ISBN: 978-94-6395-878-3
DOI: https://doi.org/10.18174/549548

Abstract

Xing, S. (2021). Differential Deposition of Intramuscular and Abdominal Fat in Chicken. PhD thesis, Wageningen University & Research, the Netherlands

Fat deposition in domesticated chicken has received much attention in modern breeding and poultry production. The aim is for high intramuscular fat (IMF) content for better meat quality and low abdominal fat (AF) for higher feed efficiency. Fatness traits are complex quantitative traits influenced by many variants in the genome some of which are likely to affect gene expression during development. Despite the ever-increasing amount of sequencing data, knowledge of fatness phenotypes and the underlying genes and variants is still limited. In this thesis, a new chicken SNP genotyping array is described based on SNPs segregating in local Chinese breeds and commercial lines. The new array offers potential benefits in breeding programs aimed at increasing both the meat quality and the feed efficiency. Furthermore, I describe the IMF and the AF deposition during development. I explore the transcriptomes of chicken breast muscle, AF, and liver in different stages. More specific, I provide time course transcriptome profiles for breast muscle, AF, and liver in chicken and identify developmentally dynamic genes in the three tissues. Weighted gene co-expression network analysis (WGCNA) results demonstrate that the expression of *L3MBTL1*, *TNIP1*, *HAT1*, and *BEND6* genes correlate to both high breast muscle IMF and low AF weight in breast muscle and AF while *ACSBG2* gene expression in liver is correlated to high AF weight. I finally provide the transcriptome analysis in breast muscle and AF from high-IMF-low-AF and low-IMF-high-AF chickens at marketing time. The expression of *ACSM3* and *CYP2AB1* correlate to both high IMF and low AF weight in breast muscle and AF, respectively. Together, I provide a comprehensive overview of gene expression affecting IMF and AF deposition in chicken. The results described in this thesis provide new insights in chicken fat deposition and allow further SNP array updates and improvements on meat quality in the selection process.

Contents

1

General introduction

1.1 Introduction

Meat has been a crucial protein source of homo sapiens diet for more than 1 million years (Carmody and Wrangham, 2009; Wood, 2017). Over the past 10,000 years, the meat supplied by domesticated animals gradually exceeded that by hunting and fishing. Chicken (Gallus gallus domesticus) is the most abundant domesticated animal. It initially was derived from the red jungle fowl (Fumihito et al., 1994; Wang et al., 2020) around 9500 ± 3300 years ago (West and Zhou, 1988; Rubin et al., 2010; Wang et al., 2020). Up to now, there are 1,825 different recognized chicken breeds/lines/strains across the world (DAD-IS-FAO, 2021). Compared with traditional breeds, the breast meat quality of modern commercial chicken lines has declined after recent decades of intensive selection (Abeni and Bergoglio, 2001; Musa et al., 2006). Intramuscular fat (IMF) content influences chicken meat quality in many aspects, e.g., juiciness, flavour, and tenderness (Chizzolini et al., 1999; Zhao et al., 2007). However, merely increased IMF selection tends to cause high abdominal fat weight (AFW) which will reduce the chicken production efficiency (Zhao et al., 2007; Jiang et al., 2017). The studies aimed at identifying the regulatory genes during tissues development and fat deposition are important to further expand chicken production possibilities and catch the demand of consumers.

Glossary
IMF: Intramuscular fat, deposited inside skeletal muscle fibers.
AF: Abdominal fat, deposited inside the abdominal cavity, packed between the organs.
SNP: Single nucleotide polymorphism.
Genotyping array: A kind of DNA chip, used for massive parallel detection of SNP alleles by hybridization technology.
DDGs: Developmentally dynamic genes, reflecting the gene expression and biological changes across developmental stages.
Hub gene: A driver gene with high correlation and connectivity in a candidate module.
TF: Transcription factor, controlling the rate of transcription of DNA genetic information (gene) to messenger RNA.

1.2 Meat-type chicken industry

1.2.1 Meat-type chicken production

As an important animal protein resource for human beings, global chicken meat production grows slowly but stably (Xin et al., 2021). In 2000, 59.64 million tons of chicken meat were produced worldwide (USDA, 2001), while in 2020, that had increased to 100.59 million tons (USDA, 2021). USA, China, Brazil, and the European Union are the 4 major chicken meat producing countries/regions (Figure 1.1). The fast increase in chicken production both in developed and developing countries/regions is mainly driven by the lower production costs with respect to pork and beef (Trapp, 2019). Nowadays, the USA is the biggest chicken meat producing and second biggest exporting country, while China, Japan, Mexico, Saudi Arabia are the major importing countries.

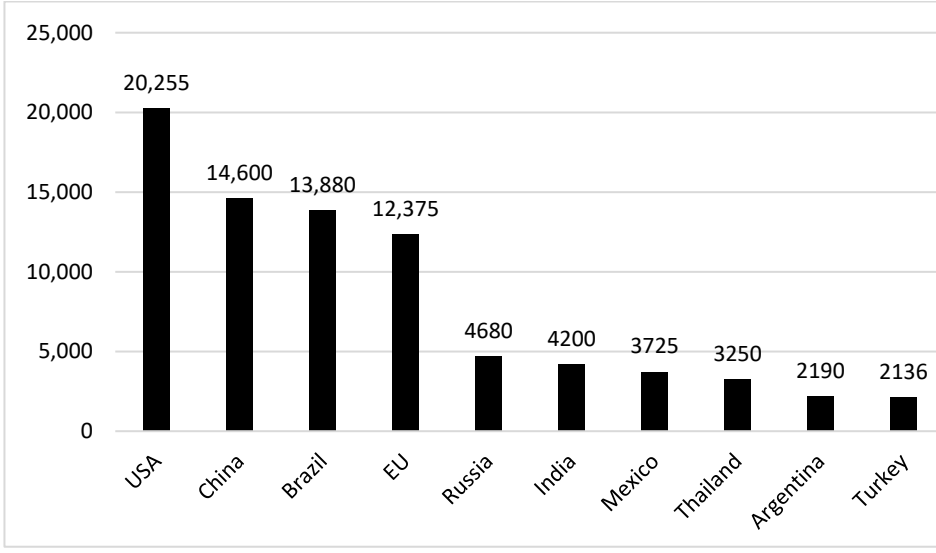


Figure 1.1 Top 10 countries/regions of chicken meat production in 2020
Note: Chicken paws are excluded. Unit in the y-axis is 1000 tons. Data resource: Economics, statistics and market information system of USDA (USDA, 2021).

1.2.2 Chicken meat quality

Meat quality is generally characterized by the compositional quality and the deliciousness factors such as appearance, smell, juiciness, tenderness, and flavour. Chicken meat is a kind of widely accepted white meat (Linseisen et al., 2002). To meet the growing demands of consumers and the increase of the worldwide

population, meat-type chicken has been intensively selected at an industrial scale for the mass market (Schmidt et al., 2009). Modern chicken genetic selection studies mainly focused on body weight (Siegel, 1962), feed efficiency (Pym and Nicholls, 1979), fecundity (Lamoreux et al., 1943), disease resistance (Waters, 1945), and animal behavior/welfare (Komai et al., 1959; Dawkins and Layton, 2012). With the intensive selection on faster growth and a heavier body weight, the incidence of chicken breast muscle abnormalities or myopathies increased e.g., deep pectoralis muscle myopathy or so called green muscle disease (Bianchi et al., 2006), white striping (Mudalal et al., 2014), spaghetti meat (Baldi et al., 2018), wooden breast (Sihvo et al., 2014), and PSE-like meat (pale, soft, and exudative meat) (Barbut, 1998). In the slaughterhouse, such abnormalities or myopathies cause chicken carcass condemnation and economic losses (Zanetti et al., 2018). Recently, the new trend of slow-growing chickens is popular to meet society complains. Consumers generally consider slow-growing chicken to have better meat quality than fast-growing broilers.

By estimating the heritability, genetic and phenotypic correlations between meat yield and quality traits in chicken, previous research concluded that selection for higher body weight may reduce the meat quality but that it is possible to improve meat quality by genetic selection (Zerehdaran et al., 2012).

1.3 Fat deposition in chicken and the molecular pathways involved

1.3.1 Fat origin and functions

Fat is a kind of loose connective tissue, mostly composed of adipocytes (Chizzolini et al., 1999). The origin of adipocytes is complex and studied mainly in mammals. According to the colour, fat can be divided into white adipose tissue, brown adipose tissue, and beige adipose tissue (Cannon and Nedergaard, 2004). White adipocytes have a big droplet that contains triacylglycerols and these cells mainly differentiate from mesenchymal stem cells in the mesoderm (Rosen and MacDougald, 2006). Brown adipocytes, with a high mitochondrial content, originate from specific precursors and are controlled by the transcriptional regulator *PRDM16* (PRD1-BF1-RIZ1 homologous domain containing 16) (Seale et al., 2008). It has been suggested that beige adipocytes are derived from white adipocytes (Reitman, 2017). However, another study reported that the previously identified brown fat in adult humans are composed of beige adipocytes (Wu et al., 2012). The development of adipocytes can be divided into hyperplasia and hypertrophy, which are both correlated with fat deposition (Cartwright, 1991).

Fat tissue is important for energy storage (Harms and Seale, 2013) and adaptive thermogenesis. In the last decades, an endocrine role of adipose was discovered , where adipocytes were shown to secrete adipokines like leptin, TNFα, IL-6, resistin, and adiponectin (Kershaw and Flier, 2004). Chickens were considered to only have white fat tissue. An excess of fat can also lead to type 2 diabetes (Montague and O'rahilly, 2000), insulin resistance (Kern et al., 2001), and inflammatory disease (Marette, 2003) in humans. Understandably, similar studies are still very limited in chickens.

1.3.2 Fat deposition

According to the location, the adipocyte tissues can be classified as follows: 1. Abdominal fat, also known as visceral fat or organ fat, is located within the abdominal wall surrounding organs. 2. Ectopic fat, which is the storage of triglycerides in tissues, that are supposed to contain only small amounts of fat, such as, liver, skeletal muscle, and heart. The fat located inside the skeletal muscle fibers is also called intramuscular fat (IMF). 3. Subcutaneous fat, which is located beneath the skin. 4. Bone marrow adipose tissue, also known as marrow adipose tissue (MAT), is a poorly understood adipose depot that resides in the bone.

Abdominal fat (AF), also called central fat or visceral fat, is the most dominant type of fat in the mature animal body. In the chicken industry, the excess abdominal fat has been one of the main reasons for the decrease in economic efficiency (Fouad and El-Senousey, 2014). AF weight is more variable than total fat and IMF (Jennen, 2004). AF stands for approximately 2-3% of body weight, e.g., in Jingxing-Huang chicken, abdominal fat holds 2.9 ± 1% in live body weight (Xing, 2020, unpublished data).

Intramuscular fat (IMF) is a kind of ectopic fat, which is also deposited in the heart and pancreas (Birbrair et al., 2013). IMF content plays a key role in various quality traits of meat (Chizzolini et al., 1999; Hocquette et al., 2010). Previous research has shown that for meat type animals, adipocyte numbers increased most rapidly in the abdominal wall, while the least rapidly in the intramuscular depot (Allen, 1976). IMF content varies between species, between lines, and between muscle types in the same animal. Moreover, it also varies with age, gender, feed, and even during the season (Hocquette et al., 2010).

Subcutaneous fat (SF) and **bone marrow adipose tissue (MAT)** are also a part of the chicken body fat content. The SF is the major non-visceral fat and is found between the hypodermis and the fascia. In chicken, SF is widely distributed but mainly

grouped together in the axilla and haunch's connective tissue. SF is widely studied in human and can be easily measured by computed tomography. SF is almost impossible to measure in living chicken because of the feathers and therefore was poorly studied in chicken until recently. A recent study described that chicken SF deposition is another important energy supply tissue during late embryonic periods and the first days after hatch (Zhao et al., 2021). MAT accounts for 70% of the bone marrow volume in adult humans (Fazeli et al., 2013). MAT research mainly focused on humans and accelerated marrow adipogenesis has been associated with several chronic diseases and aging (Rosen et al., 2009). Although both tissues are part of the chicken body fat content, they play a minor role and are not considered in our study.

The main components of IMF are triglyceride and phospholipids. The phospholipids approximately represent 20% to 50% of all the lipid, which equals about 0.5% to 1% of the weight of fresh muscle. In AF and SF, the main component is triglyceride (Fu et al., 2013).

1.3.3 Fat deposition influencing chicken meat production and quality

Fats in meat are important nutrients for humans because they are the structural component of cell membranes, essential fatty acids (FA) providers, and fat-soluble vitamins facilitators (Andueza et al., 2017). In animal production, excessive fat is considered to be dietary energy-wasting and because the feed accounts for around 70% of the total costs, excessive fat reduces carcass yield (Emmerson, 1997). While abdominal fat (AF) can be removed from the carcass after slaughtering and subcutaneous fat can be trimmed from the external surface of meat before consuming, IMF cannot be trimmed as in chicken it is almost invisible by the naked eye (Chizzolini et al., 1999). Due to the energy demands of egg laying (Scanes et al., 1987) and other hormonal functions (Cui et al., 2012), hens have more abdominal fat than males (Marx et al., 2016). The excessive fat accumulation in broiler hens negatively affects their reproductive performance both in egg quality (Cahaner et al., 1986) and the egg number from the first egg to 40 weeks (Zhang et al., 2018), which will further reduce broiler production.

When serving to consumers, both the composition of raw meat (Saláková et al., 2009) and the cooking procedures (Park et al., 2020) greatly influence the final flavour. In the raw meat composition field, IMF content plays a major role in broiler meat quality. A decrease of IMF will decrease the juiciness and flavour of the meat (Chizzolini et al., 1999). A previous study reported the pectoral muscle IMF content in partridge to be negatively correlated with the shear force (Wen et al., 2020), which means the IMF content positively affects the tenderness of poultry meat.

1.3.4 Pathways involved in FA biosynthesis

Many hormones and growth factors have been shown to be related to lipid metabolism. The lipid metabolism pathways include glycerolipid, glycerophospholipid, ether lipid, and sphingolipid metabolism. Moreover, glycerolipid is the major component of fat, and consists of glycerol and FA.

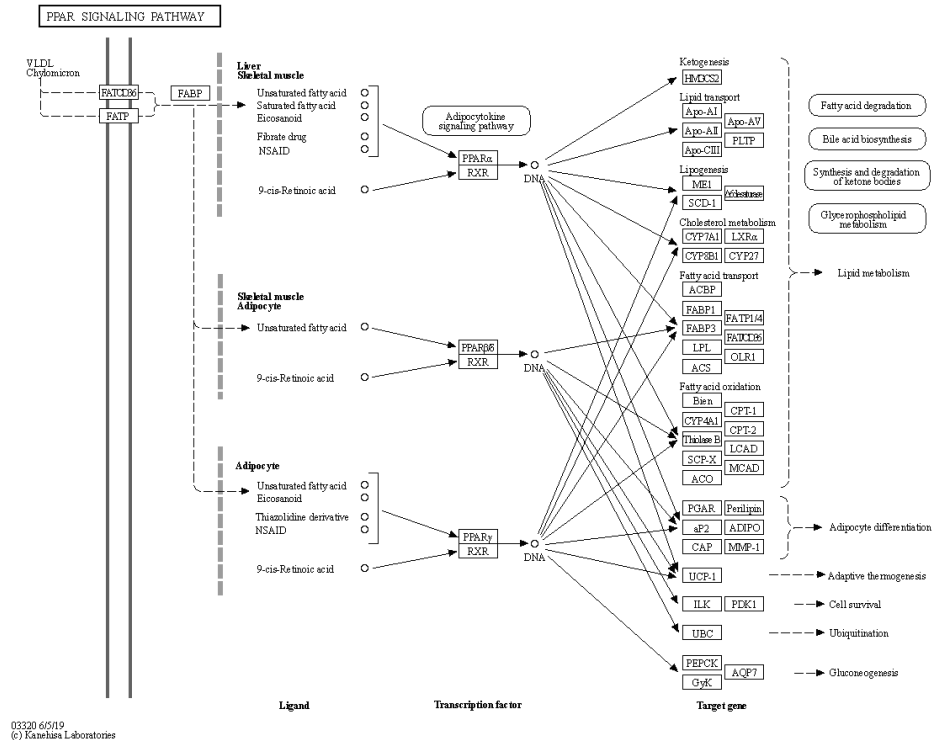


Figure 1.2 Chicken PPAR signaling pathway (https://www.genome.jp/pathway/map03320).

Fatty acid synthesis, transportation, and storage is a complex process. In Figure 1.2 an example is shown of one of the most important fat deposition pathways, the PPAR (peroxisome proliferator-activated receptor) signaling pathway. In chicken, the liver is the central organ for lipid metabolism where around 90% of the FA is *de novo* synthesized (O'Hea and Leveille, 1969). The hepatic fatty acid synthase protein synthesizes lipids that are secreted in very low-density lipoprotein (VLDL). VLDL and chylomicrons transport endogenous and exogenous (dietary) lipids to cells, respectively. The fatty acid translocase (FAT), also called cluster of differentiation 36 (CD36) and fatty acid transport proteins (FATP), both located in the cell membrane,

facilitate uptake of FA into cells by fatty acid binding protein (FABP) (Jensen-Urstad and Semenkovich, 2012). The FAs can subsequently activate the PPARs and retinoid X receptor (RXR) complex. Then the PPARs, as TFs, activate the downstream pathways e.g., lipid metabolism, adipocyte differentiation, and adaptive thermogenesis. Different PPAR isoforms are expressed in different tissues, e.g., PPARα is expressed in liver and skeletal muscle, whereas PPARδ is expressed in skeletal muscle adipocytes, and PPARγ is expressed in adipocytes (Berger and Moller, 2002).

1.4 Jingxing Huang Chicken as the study animal

1.4.1 Origin of Jingxing Huang Chicken

Jingxing Huang breed is a dwarf yellow-feathered meat-type chicken, which was initially developed from a cross of Beijing-You chicken (China), Beijing-Huang chicken (China), and White Bro chicken (France). Among the three base ancestors, Beijing-You chicken has commonly been considered a high-quality meat chicken breed (Zheng, 1988; Zhang, 2011).

Figure 1.3 Male and female Jingxing-Huang chickens.

1.4.2 Rearing, phenotyping, and selection methods

The IMF-up selection line was originally developed in the animal experiment station of the Institute of Animal Sciences, Chinese Academy of Agricultural Sciences in the year 2000. Founder animals consisted of 40 sires and 400 dams, which were

randomly selected from the base population of Jingxing-100 chicken. The selection maintained 1 generation per year (Zhao et al., 2006).

In each generation, two batches of fertilized eggs were hatched. Chickens had *ad libitum* access to water and feed, which was based on a corn-soybean type diet throughout generations (Table 1.1).

Table 1.1 Feed energy levels and crude protein rate in different rearing stages.

Diet	Period (weeks)	Energy (MJ/Kg)	Crude Protein (%)
Starter diet	0 – 8	12.1	20
Grower diet	9 – 18*	11.7	17
Pullet diet	19 – 24	11.5	16
Adult diet	25 and afterwards	11.5	16

Note *: Slaughtering test was performed in this period (13 to 14 weeks).

The mean family IMF value was used as individuals' phenotype. All the selected sire breeders had a higher mean family IMF content than the population IMF content. At maximum, 3 full-sib male breeders in any family were selected. For each mate, there is no inbreeding within 3 generations for the sire and dams.

The balanced selection line (IMF-up and abdominal fat percentage-down selection line) was randomly chosen from the base population of the IMF-up selection line in the year 2008. Then the two lines were independently selected. In each generation, 25 to 30 sire families were selected for breeding the next generation. The sires in the balanced selection line have higher family IMF values and lower abdominal fat percentage (AFP) values than the population average value (Jiang et al., 2017).

The birds used in this thesis for whole genome sequencing (Liu et al., 2019) and developmental transcriptome studies (Xing et al., 2020) were derived from the IMF-up selection line (Zhao et al., 2006; Zhao et al., 2007).

1.5 Aims and outline of this thesis

In this thesis, I aim to provide new insight on the possibilities to improve the IMF content while at the same time reducing the AF content in commercial meat-type chickens. First, I joined a new mid-density chicken SNP genotyping array development, which has already been used for commercial genomic selection, GWAS for feed efficient trait (Li et al., 2021), and to study diversity of different chicken breeds (Liu et al., 2020). Second, I used transcriptome data in breast muscle, AF, and liver (three important tissues relevant for fat deposition) to study the gene regulation in different tissues and the corresponding correlation with IMF and AF deposition.

In **Chapter 2**, SNPs are selected and used in the design of a new SNP genotyping array derived from five different resources i) whole genome sequencing data of representative local breeds and commercial broiler lines, ii) significant results of GWAS on interesting traits, iii) candidate genes of economic traits, iv) residual feed intake related, and v) previous genotyping data. The newly developed SNP genotyping array was validated by genotyping 313 animals derived from 13 different chicken breeds/lines. In **Chapters 3 and 4**, gene expression in breast muscle, AF, and liver during tissue development in the IMF-up selected chicken line are addressed. In this study I used 9 representative developmental stages. By profiling gene expression, I provided a useful resource for transcriptome data for these 3 tissues. After quantifying gene expression between adjacency stages, differentially expressed genes and enriched pathways are identified. Using clustering analysis, developmentally dynamic genes and transcription factors are identified at different time points. In **Chapter 5**, I describe the correlations between gene expression and breast muscle IMF content and AFW in Jingxing-Huang chickens. Applying a weighted correlation network analysis, new hub genes correlated to breast muscle IMF content and AFW are identified in marketing time. Finally, in **Chapter 6** the knowledge contributions of chicken genotyping and gene regulation with the overall findings in this thesis are discussed. Moreover, I discuss future utilization and new possibilities of fat related traits selection in chickens.

References

Abeni, F., and Bergoglio, G. (2001). Characterization of different strains of broiler chicken by carcass measurements, chemical and physical parameters and NIRS on breast muscle. *Meat Science* 57(2), 133-137.

Allen, C. (1976). Cellularity of adipose tissue in meat animals. *Federation proceedings* 35(11), 2302-2307.

Andueza, D., Mourot, B.-P., Hocquette, J.-F., and Mourot, J. (2017). Phenotyping of Animals and Their Meat: Applications of Low-Power Ultrasounds, Near-Infrared Spectroscopy, Raman Spectroscopy, and Hyperspectral Imaging. *Lawrie´ s Meat Science*, 501-519.

Baldi, G., Soglia, F., Mazzoni, M., Sirri, F., Canonico, L., Babini, E., et al. (2018). Implications of white striping and spaghetti meat abnormalities on meat quality and histological features in broilers. *Animal* 12(1), 164-173.

Barbut, S. (1998). Estimating the magnitude of the PSE problem in poultry. *Journal of Muscle Foods* 9(1), 35-49.

Berger, J., and Moller, D.E. (2002). The mechanisms of action of PPARs. *Annual review of medicine* 53(1), 409-435.

Bianchi, M., Petracci, M., Franchini, A., and Cavani, C. (2006). The occurrence of deep pectoral myopathy in roaster chickens. *Poultry science* 85(10), 1843-1846.

Birbrair, A., Zhang, T., Wang, Z.-M., Messi, M.L., Enikolopov, G.N., Mintz, A., et al. (2013). Role of pericytes in skeletal muscle regeneration and fat accumulation. *Stem cells and development* 22(16), 2298-2314.

Cahaner, A., NITSAN, Z., and NIR, I. (1986). Reproductive performance of broiler lines divergently selected on abdominal fat. *Poultry Science* 65(7), 1236-1243.

Cannon, B., and Nedergaard, J. (2004). Brown adipose tissue: function and physiological significance. *Physiological reviews* 84(1), 277-359.

Carmody, R.N., and Wrangham, R.W. (2009). The energetic significance of cooking. *Journal of human evolution* 57(4), 379-391.

Cartwright, A.L. (1991). Adipose cellularity in Gallus domesticus: investigations to control body composition in growing chickens. *The Journal of nutrition* 121(9), 1486-1497.

Chizzolini, R., Zanardi, E., Dorigoni, V., and Ghidini, S. (1999). Calorific value and cholesterol content of normal and low-fat meat and meat products. *Trends in Food Science & Technology* 10(4-5), 119-128.

Cui, H., Zhao, G., Liu, R., Zheng, M., Chen, J., and Wen, J. (2012). FSH stimulates lipid biosynthesis in chicken adipose tissue by upregulating the expression of its receptor FSHR. *Journal of lipid research* 53(5), 909-917.

DAD-IS-FAO (2021). *Domestic Animal Diversity Information System* [Online]. Available: http://www.fao.org/dad-is/dataexport/en/ [Accessed June 10, 2021].

Dawkins, M., and Layton, R. (2012). Breeding for better welfare: genetic goals for broiler chickens and their parents. *Animal Welfare-The UFAW Journal* 21(2), 147.

Emmerson, D. (1997). Commercial approaches to genetic selection for growth and feed conversion in domestic poultry. *Poultry Science* 76(8)**,** 1121-1125.

Fazeli, P.K., Horowitz, M.C., MacDougald, O.A., Scheller, E.L., Rodeheffer, M.S., Rosen, C.J., et al. (2013). Marrow fat and bone—new perspectives. *The Journal of Clinical Endocrinology & Metabolism* 98(3)**,** 935-945.

Fouad, A., and El-Senousey, H. (2014). Nutritional factors affecting abdominal fat deposition in poultry: a review. *Asian-Australasian journal of animal sciences* 27(7)**,** 1057.

Fu, R., Zhao, G., Liu, R., Zheng, M., Chen, J., and Wen, J. (2013). Research on body fat tistribution and deposition pattern of Beijing-You Chickens. *Chinese Journal of Animal Nutrition* 25(7)**,** 1465-1472.

Fumihito, A., Miyake, T., Sumi, S.-I., Takada, M., Ohno, S., and Kondo, N. (1994). One subspecies of the red junglefowl (Gallus gallus gallus) suffices as the matriarchic ancestor of all domestic breeds. *Proceedings of the National Academy of Sciences* 91(26)**,** 12505-12509.

Harms, M., and Seale, P. (2013). Brown and beige fat: development, function and therapeutic potential. *Nature medicine* 19(10)**,** 1252-1263.

Hocquette, J.-F., Gondret, F., Baeza, E., Medale, F., Jurie, C., and Pethick, D. (2010). Intramuscular fat content in meat-producing animals: development, genetic and nutritional control, and identification of putative markers. *Animal* 4(2)**,** 303-319.

Jennen, D.G.J. (2004). *Chicken fatness: From QTL to candidate gene.*

Jensen-Urstad, A.P., and Semenkovich, C.F. (2012). Fatty acid synthase and liver triglyceride metabolism: housekeeper or messenger? *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids* 1821(5)**,** 747-753.

Jiang, M., Fan, W.L., Xing, S.Y., Wang, J., Li, P., Liu, R.R., et al. (2017). Effects of balanced selection for intramuscular fat and abdominal fat percentage and estimates of genetic parameters. *Poult Sci* 96(2)**,** 282-287.

Kern, P.A., Ranganathan, S., Li, C., Wood, L., and Ranganathan, G. (2001). Adipose tissue tumor necrosis factor and interleukin-6 expression in human obesity and insulin resistance. *American Journal of Physiology-Endocrinology And Metabolism* 280(5)**,** E745-E751.

Kershaw, E.E., and Flier, J.S. (2004). Adipose tissue as an endocrine organ. *The Journal of Clinical Endocrinology & Metabolism* 89(6)**,** 2548-2556.

Komai, T., Craig, J., and Wearden, S. (1959). Heritability and repeatability of social aggressiveness in the domestic chicken. *Poultry Science* 38(2)**,** 356-359.

Lamoreux, W., Hutt, F., and Hall, G. (1943). Breeding for low fecundity in the fowl with the aid of the progeny test. *Poultry Science* 22(2)**,** 161-169.

Li, W., Zheng, M., Zhao, G., Wang, J., Liu, J., Wang, S., et al. (2021). Identification of QTL regions and candidate genes for growth and feed efficiency in broilers. *Genetics Selection Evolution* 53(1)**,** 1-17.

Linseisen, J., Kesse, E., Slimani, N., Bueno-De-Mesquita, H., Ocké, M., Skeie, G., et al. (2002). Meat consumption in the European Prospective Investigation into

Cancer and Nutrition (EPIC) cohorts: results from 24-hour dietary recalls. *Public Health Nutrition* 5(6b), 1243-1258.

Liu, L., Cui, H., Xing, S., Zhao, G., and Wen, J. (2020). Effect of divergent selection for intramuscular fat content on muscle lipid metabolism in chickens. *Animals* 10(1), 4.

Liu, R., Xing, S., Wang, J., Zheng, M., Cui, H., Crooijmans, R.P., et al. (2019). A new chicken 55K SNP genotyping array. *BMC genomics* 20(1), 1-12.

Marette, A. (2003). Molecular mechanisms of inflammation in obesity-linked insulin resistance. *International Journal of Obesity* 27(S3), S46.

Marx, S., Soares, J., Prestes, R., Schnitzler, E., Oliveira, C., Demiate, I., et al. (2016). Influence of sex on the physical-chemical characteristics of abdominal chicken fat. *Brazilian Journal of Poultry Science* 18(2), 269-276.

Montague, C.T., and O'rahilly, S. (2000). The perils of portliness: causes and consequences of visceral adiposity. *Diabetes* 49(6), 883-888.

Mudalal, S., Babini, E., Cavani, C., and Petracci, M. (2014). Quantity and functionality of protein fractions in chicken breast fillets affected by white striping. *Poultry Science* 93(8), 2108-2116.

Musa, H., Chen, G., Cheng, J., Shuiep, E., and Bao, W. (2006). Breed and sex effect on meat quality of chicken. *International Journal of Poultry Science* 5(6), 566-568.

O'Hea, E.K., and Leveille, G.A. (1969). Lipid biosynthesis and transport in the domestic chick (Gallus domesticus). *Comp Biochem Physiol* 30(1), 149-159.

Park, C., Lee, B., Oh, E., Kim, Y., and Choi, Y. (2020). Combined effects of sous-vide cooking conditions on meat and sensory quality characteristics of chicken breast meat. *Poultry Science* 99(6), 3286-3291.

Pym, R., and Nicholls, P. (1979). Selection for food conversion in broilers: Direct and correlated responses to selection for body‐weight gain, food consumption and food conversion ratio. *British Poultry Science* 20(1), 73-86.

Reitman, M.L. (2017). How does fat transition from white to beige? *Cell metabolism* 26(1), 14-16.

Rosen, C.J., Ackert-Bicknell, C., Rodriguez, J.P., and Pino, A.M. (2009). Marrow fat and the bone microenvironment: developmental, functional, and pathological implications. *Critical Reviews™ in Eukaryotic Gene Expression* 19(2), 109-124.

Rosen, E.D., and MacDougald, O.A. (2006). Adipocyte differentiation from the inside out. *Nature reviews Molecular cell biology* 7(12), 885-896.

Rubin, C.-J., Zody, M.C., Eriksson, J., Meadows, J.R., Sherwood, E., Webster, M.T., et al. (2010). Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature* 464(7288), 587-591.

Saláková, A., Straková, E., Válková, V., Buchtová, H., and Steinhauserová, I. (2009). Quality indicators of chicken broiler raw and cooked meat depending on their sex. *Acta Veterinaria Brno* 78(3), 497-504.

Scanes, C.G., Campbell, R., and Griminger, P. (1987). Control of energy balance during egg production in the laying hen. *The Journal of nutrition* 117(3)**,** 605-611.

Schmidt, C.J., Persia, M., Feierstein, E., Kingham, B., and Saylor, W. (2009). Comparison of a modern broiler line and a heritage line unselected since the 1950s. *Poultry science* 88(12)**,** 2610-2619.

Seale, P., Bjork, B., Yang, W., Kajimura, S., Chin, S., Kuang, S., et al. (2008). PRDM16 controls a brown fat/skeletal muscle switch. *Nature* 454(7207)**,** 961-967.

Siegel, P.B. (1962). Selection for body weight at eight weeks of age: 1. Short term response and heritabilities. *Poultry science* 41(3)**,** 954-962.

Sihvo, H.-K., Immonen, K., and Puolanne, E. (2014). Myodegeneration with fibrosis and regeneration in the pectoralis major muscle of broilers. *Veterinary pathology* 51(3)**,** 619-623.

Trapp, J.N. (2019). "Production Cost Comparisons Among Beef, Pork, and Chicken," in *Beef Cattle Science Handbook*. CRC Press), 767-774.

USDA (2001). *Livestock and Poultry: World Markets and Trade* [Online]. Available: https://downloads.usda.library.cornell.edu/usda-esmis/files/73666448x/pr76f381t/gq67jr501/livestock-poultry-ma-10-01-2001.pdf [Accessed June 9th, 2021].

USDA (2021). *Livestock and Poultry: World Markets and Trade* [Online]. Available: https://downloads.usda.library.cornell.edu/usda-esmis/files/73666448x/zs25z463c/np194471v/livestock_poultry.pdf [Accessed January 12, 2021].

Wang, M.-S., Thakur, M., Peng, M.-S., Jiang, Y., Frantz, L.A.F., Li, M., et al. (2020). 863 genomes reveal the origin and domestication of chicken. *Cell research* 30(8)**,** 693-701.

Waters, N.F. (1945). Breeding for resistance and susceptibility to avian lymphomatosis. *Poultry Science* 24(3)**,** 259-269.

Wen, Y., Liu, H., Liu, K., Cao, H., Mao, H., Dong, X., et al. (2020). Analysis of the physical meat quality in partridge (Alectoris chukar) and its relationship with intramuscular fat. *Poultry science* 99(2)**,** 1225-1231.

West, B., and Zhou, B.-X. (1988). Did chickens go north? New evidence for domestication. *Journal of archaeological science* 15(5)**,** 515-533.

Wood, J.D. (2017). "Meat composition and nutritional value," in *Lawrie s Meat Science*. Elsevier), 635-659.

Wu, J., Boström, P., Sparks, L.M., Ye, L., Choi, J.H., Giang, A.-H., et al. (2012). Beige adipocytes are a distinct type of thermogenic fat cell in mouse and human. *Cell* 150(2)**,** 366-376.

Xin, X., Zheng, M., Wen, J., and Wang, J. (2021). The situation analysis, future prospect and countermeasures of meat type chicken industry in China in 2020. *Chinese Journal of Animal Science* 57(3)**,** 217-222.

Xing, S., Liu, R., Zhao, G., Liu, L., Groenen, M.A., Madsen, O., et al. (2020). RNA-Seq analysis reveals hub genes involved in chicken intramuscular fat and abdominal fat deposition during development. *Frontiers in genetics* 11.

Zanetti, M.A., Tedesco, D.C., Schneider, T., Teixeira, S.T.F., Daroit, L., Pilotto, F., et al. (2018). Economic losses associated with Wooden Breast and White Striping in broilers. *Semina: Ciências Agrárias* 39(2)**,** 887-892.

Zerehdaran, S., Lotfi, E., and Rasouli, Z. (2012). Genetic evaluation of meat quality traits and their correlation with growth and carcase composition in Japanese quail. *British poultry science* 53(6)**,** 756-762.

Zhang, B. (2011). China National Commission of Animal Genetic Resources. *Animal genetic resources in China: Horses, Donkeys, Camels. China Agricultural Press, China*.

Zhang, X., Wu, M., Wang, S., Zhang, H., Du, Z., Li, Y., et al. (2018). Genetic selection on abdominal fat content alters the reproductive performance of broilers. *Animal* 12(6)**,** 1232-1241.

Zhao, G., Chen, J., Zheng, M., Wen, J., and Zhang, Y. (2007). Correlated responses to selection for increased intramuscular fat in a Chinese quality chicken line. *Poultry Science* 86(11)**,** 2309-2314.

Zhao, G., Wen, J., Chen, J., Zheng, M., and Li, X. (2006). Selection response and estimation of the genetic parameters for intramuscular fat in a quality chicken line. *J Acta veterinaria et zootechnica Sinica* 37(9)**,** 870-873.

Zhao, H., Wu, M., Tang, X., Li, Q., Yi, X., Wang, S., et al. (2021). Function of chick subcutaneous adipose tissue during the embryonic and posthatch period. *Frontiers in Physiology* 12**,** 831.

Zheng, P. (1988). Breeds of domesticated animal and poultry in China. *Shanghai Scientific and Technical Publishers, Shanghai, China*.

2

A new chicken 55K SNP genotyping array

Ranran Liu[1,#], Siyuan Xing[1,2,#], Jie Wang[1], Maiqing Zheng[1], Huanxian Cui[1], Richard P. M. A. Crooijmans[2], Qinghe Li[1], Guiping Zhao[1] and Jie Wen[1]

[1] Institute of Animal Sciences, Chinese Academy of Agricultural Sciences; State Key Laboratory of Animal Nutrition; Key Laboratory of Animal (Poultry) Genetics Breeding and Reproduction, Ministry of Agriculture, Beijing 100193, P. R. China.

[2] Animal Breeding and Genomics, Wageningen University & Research, 6708PB Wageningen, The Netherlands.

[#] Ranran Liu and Siyuan Xing contributed equally to this work

BMC Genomics (2019) 20:410

Abstract

Background: China has the richest local chicken breeding resources in the world and is the world's second largest producer of meat-type chickens. Development of a moderate-density SNP array for genetic analysis of chickens and breeding of meat-type chickens taking utility of those resources is urgently needed for conventional farms, breeding industry and research areas.

Results: Eight representative local breeds or commercial broiler lines with 3 pools of 48 individuals within each breed/line were sequenced and supplied the major SNPs resource. There were 7.09 million - 9.41 million SNPs detected in each breed/line. After filtering using multiple criteria such as preferred incorporation of trait-related SNPs and uniformity of distribution across the genome, 52,184 SNPs were selected in the final array. It consists of: (i) 19.2 K SNPs from the genomes of yellow-feathered, cyan-shank partridge and white-feathered chickens; (ii) 6.0 K SNPs related to economic traits from the Illumina 60 K SNP Bead Chip, which were found as significant associated SNPs with 15 traits in a Beijing-You crossed Cobb F2 resource population by genome-wide association study analysis; (iii) 7.6 K SNPs from 861 candidate genes of economic traits; (iv) the 1.0 K SNPs related to residual feed intake; and (v) 18.4 K from chicken SNPdb. The polymorphisms of 9 extra local breeds and 3 commercial lines were examined with this array, and 40 K - 47 K SNPs were polymorphic (with minor allele frequency >0.05) in those breeds. The MDS result showed that those breeds can be clearly distinguished by this newly developed genotyping array.

Conclusions: We successfully developed a 55K genotyping array by using SNPs segregating in typical local breeds and commercial lines. Compared to the existing Affy 600K and Illumina 60K arrays, there are 21.4 K new SNPs included on our Affy 55K array. The results of the 55K genotyping data can therefore be imputed to high-density SNP genotyping data. The array offers a wide range of potential applications such as genomic selection breeding, GWAS of interested traits, and investigation of diversity of different chicken breeds.

Key words: Chicken, commercial line, Genotyping array, SNP

2.1 Introduction

With a total of 107 chicken breeds, China has one of the richest local breed resources (Zhang, 2011). This diverse chicken genetic resource is a vital part of the diversity of biological genetic resources around the world and provides excellent material for breeding new varieties or to genetically improve breed.

China is the second-largest broiler producer and consumer all over the world, which accounts for approximately 40% of the chicken production across the globe (FAOSTAT, 2018). In China, chicken is the second largest meat product after pork, making up to 17% of the total meat production. Chicken meat is mainly obtained from the introduced white feather broilers and domestic, yellow-feathered meat-type chickens (meat-type local chicken breed, meat-type bred variety and a relevant strain containing the consanguinity of Chinese indigenous chicken), each accounting for half of the consumption. However, the current challenge is how to effectively protect and maintain the existing local varieties. On the other hand, improving breeding efficiency will accelerate breeding of new chicken lines. The genome-wide SNP chips, also known as SNP array, arranges up to 25 million of DNA marker flanks on glass or special silicon chip to form the SNP probe array. It's function is based on base pairing between the chip fixed DNA marker flanks with the target genome, to accurately identify the genetic information.

Genotyping arrays have been developed for pig (Ramos et al., 2009), cow (Matukumalli et al., 2009), dairy cattle (Dash et al., 2017), sheep (Anderson et al., 2014), salmon (Houston et al., 2014), and buffalo (Iamartino et al., 2017). In chicken, the first 3 K genotyping array was developed in 2005 with 3,072 SNPs (Muir et al., 2008). After that, in 2008, Groenen et al. (2011) developed a 60K bead chip for chicken which evenly covered the whole genome. To date, the only available commercial array for chicken is the Chicken Affy 600K SNP Array (Axiom Genome-Wide Chicken Genotyping Array), which was developed by Kranis et al. (2013) The other arrays are privately owned by commercial companies. The array provides an important tool for the genetic diversity analysis, breeds relationship analysis, GWAS, quantitative character positioning analysis of QTL, selective evolution investigation, and Genomic Selection (Derks et al., 2018). Up till now the most efficient ways for SNP genotyping, biodiversity measuring, QTL mapping and genomic selection is using SNP arrays. These applications provide improved technical support for the conservation of indigenous breeds and development of new genetic lines/breeds.

One pitfall of all current chicken SNP arrays is the bias towards western commercial lines. The current chicken arrays, however, lack the genomic variation information of Chinese indigenous breeds. Therefore, it is imperative to develop a new type of genome-wide SNP chip with moderate flux in the chicken breeding industry that also contains the genetic variation information specific to Chinese indigenous breeds. Overlap with the current arrays of the different platforms (Axiom and Illumina) is essential to link the commercial SNP arrays.

Through whole genome re-sequencing of a variety of Chinese native breeds and commercial chicken lines, integrating SNPs associated with economic traits detected in cross breeds (either indigenous or commercial), a new publicly available moderate density (55K) chicken array (IASCHICK) has been developed.

2.2 Results

The SNPs selection was performed in four groups. The roadmap is shown in Figure 2.1, and the establishment of the four groups is indicated in the following paragraphs.

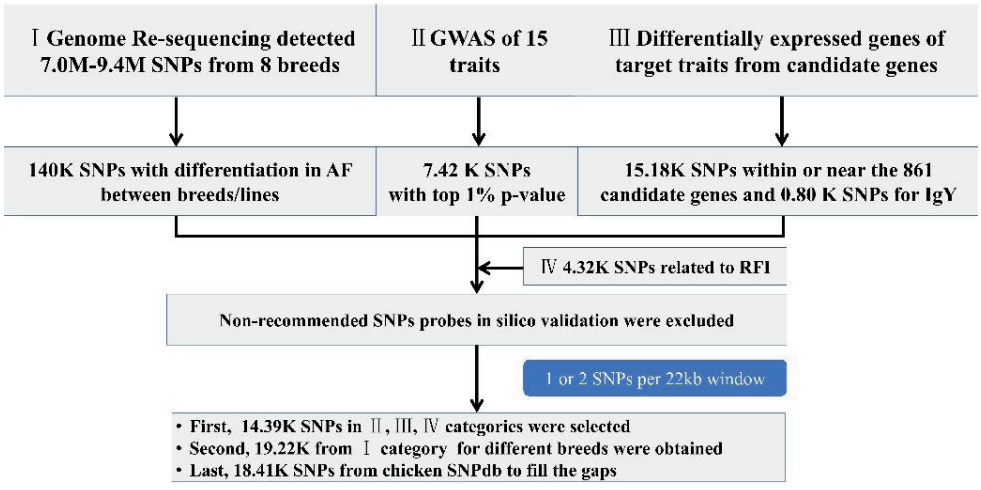


Figure 2.1. The roadmap for the design of the new chicken 55K SNP array.

Genome re-sequencing of chickens supplying the first SNP group

Eight Chinese local chicken breeds or inbred lines were selected for whole genome sequencing. Each breed/line holds 3 pools of 16 individuals per library without individual barcodes (Table 2.1). The data summary of each library is provided in the

Supplemental table 2.1. The number of SNPs per breed/line varied from 7.09 M to 9.41 M SNPs. The average number of detected SNPs was 8.61 M in the local lines, and 7.73 M in the commercial broilers. The total number of SNPs detected overall 8 breeds/lines was 15.2 M. The SNPs with minor allele frequency (MAF) < 0.05 and with low ΔF were excluded for further steps. The 140.0 K SNPs, which allelic frequencies distinct to the control breeds, were subsequently used as the first group of candidate SNPs.

Table 2.1 Sequenced chickens and the number of SNPs detected from different breeds.

Type	Breeds	Individuals in each pool*	No. of pools	Number of detected SNPs (with QC ≥ 20) **
Local Yellow-feathered chicken	Beijing-You	16	3	8,505,214
	Jingxing-Huang	16	3	8,349,627
	Sanhuang	16	3	9,405,319
Cyan shank partridge	Cyan shank partridge (fast growth rate)	16	3	8,954,795
	Cyan shank partridge (mediate growth rate)	16	3	8,884,232
Commercial white-feathered	Cobb maternal line	16	3	7,093,225
	Cobb paternal line	16	3	8,372,769
	Recessive White	16	3	7,556,464
Total				15,312,402

*Each pool contained 8 males and 8 females.

**Based on Gallus_gallus-4.0.

Selection of the second group of candidate SNPs based on the GWAS of 15 traits

The 7.4 K SNPs were demonstrated to have the top 1% genome-wide significance in 15 traits and were selected as the second group of SNPs. The details are shown in Supplemental file 2.2.

Selection of the third group of candidate SNPs based on the genes associated with economic traits

SNPs in the regions of 861 candidate genes related to economic traits were used according to previous studies of gene/protein expression profiles. A total of 66.4K SNPs in 383 genes for breast muscle and intramuscular fat development in

embryonic and post-hatching periods (Liu et al., 2017), 24.7 K SNPs in 286 genes for body fat metabolism (Huang et al., 2018), 32.6 K SNPs in 146 genes for disease resistance (Li et al., 2018), and 7.2 K SNPs in 46 genes that exhibited possible influence on other chicken economic traits (Fan et al., 2015) were selected (Supplemental table 2.3). Flanking SNPs located within 5 Kb up- and down-stream of the selected SNPs were also considered.

According to the SNPs detected by genome resequencing of the previously mentioned 8 breeds, 15.2 K candidate SNPs were selected from 118,470 K SNPs on all those genes, which had priority with mutations in exons, splicing regions, promotors, and the 3' and 5' untranslated regions (UTRs).

In addition, a batch of 0.8 K SNPs from an unpublished capture sequencing of chicken chromosomes 11, 16, and 19 were included in the third candidate group (Supplemental table 2.4). The SNPs were significantly related to high IgY levels in Beijing-You and White Leghorn chickens.

There were 16.0 K candidate SNPs that were selected for the design of the final genotyping array.

The fourth group of candidate SNPs are derived from whole genome sequences of low- and high-RFI chickens

Whole genome sequencing of low- and high-RFI chickens were performed to locate the genomic variants for RFI based on differences in allelic frequency between high- and low-RFI chickens as described in our previous study (Liu et al., 2018). The selected 4.3 K SNPs (3.7 K RFI related SNP in Beijing-You chickens and 0.6 K RFI related SNPs in Cobb chickens) were used as the candidate SNPs for the design of the final genotyping array in the next step.

Designing the Affy 55K genotyping array

Based on the above four groups of candidate SNPs, a custom-made algorithm was used to fix the final array. Finally, 52,184 SNPs were selected for the final array. The mean physical distance of SNPs in each involved chromosome is show in Table 2. The priority 1 SNPs (the SNPs in group 2, 3 and 4) and 25 insertions and deletions were first placed on the final SNP panel. The next step was addition of the priority 2 SNPs (the SNPs in group 1). The remaining 18.4 K SNPs were selected for the blank windows in the whole chicken genome which the SNPs in the four groups do not cover.

Table 2.2 The number of SNPs of the 55K array on each chromosome and their distance*.

Chromosome	Number of SNPs	Mean Distance (K bp)
1	10,228	19.21
2	7,077	21.14
3	5,196	21.44
4	4,589	19.91
5	2,705	22.10
6	1,750	20.36
7	1,684	21.63
8	1,314	22.81
9	1,236	18.98
10	1,399	14.59
11	1,373	14.67
12	1,389	14.34
13	1,041	17.67
14	1,118	14.43
15	761	16.71
16	81	7.37
17	724	14.31
18	736	14.94
19	725	13.74
20	867	16.04
21	503	13.53
22	153	30.00
23	321	17.76
24	390	15.98
25	106	26.67
26	339	15.36
27	277	20.22
28	317	15.91
Z	3785	21.67
Summary	52,184	

* The distance between SNPs based on Gallus_gallus-5.0.

The SNPs positions of 55K array are given in Supplemental table 2.5. The selected SNPs were derived from the following five groups (Table 2.3): (i) 19.2 K SNPs from whole genome sequencing of the eight chicken breeds/lines; (ii) 7.4 K trait-related SNPs from the Illumina 60K SNP Bead Chip, which were found as SNPs significantly associated with 15 economic traits; (iii) 16.0K SNPs from 861 candidate genes of target traits and high IgY level related region; (iv) 4.3 K SNPs related to chicken RFI;

and (v) 18.4 K from chicken SNPdb. In the final genotyping array, 99.85% of the SNPs could be annotated (Table 2.4). The distribution of SNPs on the chromosomes is shown in Figure 2.2.

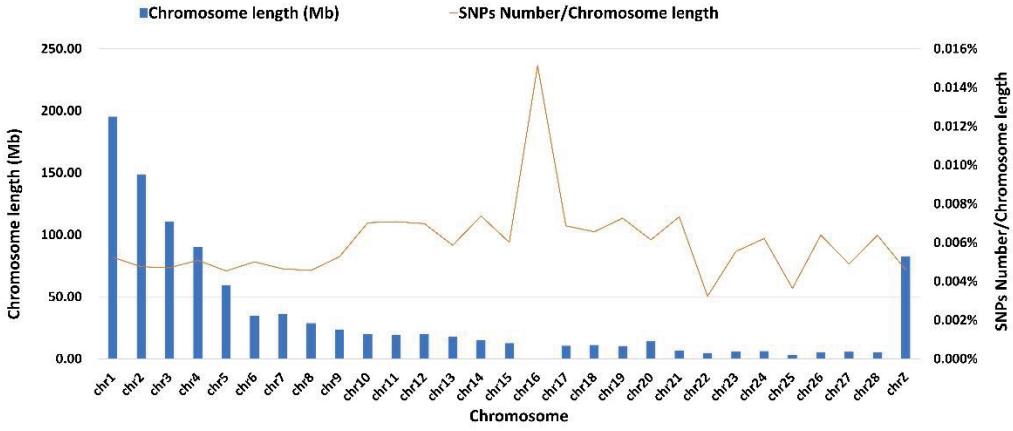


Figure 2.2 The chromosome-wise SNP density of the 55K SNP array. Chromosome length shows in left axis (based on galGal-5) and SNP density shows in right axis.

Table 2.3 The number of SNPs from five candidate groups in the final 55K array.

Resource Category	Number of SNPs in 55K array
I. Genome Re-sequencing of eight breeds	
White-feathered	12,555
Yellow-feathered	3,940
Cyan-shank Partridge Chicken	2,724
II. SNPs based on GWAS of 15 traits	5,980
III. SNPs on the candidate genes	7,630
IV. SNPs related to RFI	943
V. SNPs from chicken SNPdb	18,412
Total	52,184

Table 2.4 Summary of the SNPs effect prediction in 55K array.

Item	Count	Percent (%)
Total number of SNPs in the panel	52,184	
Annotation possible	52,108	99.85
intergenic variant	16,106	30.86
intron variant	25,275	48.43
intron variant & noncoding transcript variant	3,981	7.63
missense variant	590	1.13
missense variant & splice region variant	13	0.02
synonymous variant	1,601	3.07
Splicing	187	0.36
start/stop gained/lost/retained	12	0.02
3 prime UTR variant	1,358	2.60
5 prime UTR variant	229	0.44
upstream gene variant (1kb)	871	1.67
downstream gene variant (1kb)	1014	1.94
noncoding transcript exon variant	871	1.67

The comparisons of the Affy 55K array with the existing chicken arrays (Affy 600K array, and Illumina 60K)

All the SNPs of the current 55K array, Affy 600K array (Kranis et al., 2013), and Illumina 60K array (Groenen et al., 2011) were mapped to the latest chicken genome (GRCg6a). The overlap of the 3 arrays is shown in Figure 3. There are 6,740 SNPs (13%) which overlap between the Affy 55K array and the Illumina 60K array. When comparing to the Affy 600K array, there are 24,227 SNPs that overlap between the 55K array which accounts for 46%. There are 21,412 new SNPs included in 55K array compared to the existing arrays.

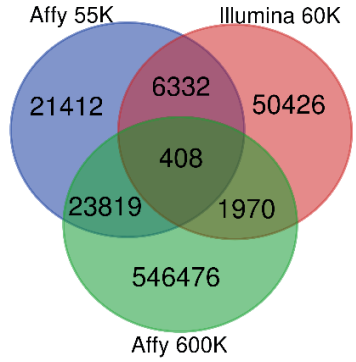


Figure 2.3 The comparison of the overlap of the SNP positions among the Affy 55K array, Affy 600K array and Illumina 60K array.

Validation of the 55K array in 13 chicken breeds/lines

All samples from 10 Chinese local breeds (Chahua, Dagu, Liyang, Luhua, Qingyuan, Silkie, Wenchang, Bai'er, Xianju, and Jingxing-Huang) and 3 commercial lines (Hubbard, Cobb, and White Leghorn) were genotyped with the new 55K array.

The average call rate for each breed ranged from 97.0% (Qingyuan) to 98.7% (Cobb). Across all populations, 76.7% to 88.0% of the 52,184 SNPs were polymorphic, with MAF ≥ 0.05. The average MAF ranged from 0.22 (Bai'er chicken) to 0.27 (Wenchang chicken) (Table 2.5).

Table 2.5 Number of polymorphic loci in local breeds and introduced lines.

Breeds	Average Call Rate (%)	Polymorphic loci*	Mean MAF**
Chahua	97.56	42.3K	0.242
Qingyuan	97.00	45.2K	0.267
Wenchang	97.10	46.5K	0.277
Luhua	97.21	44.4K	0.261
Liyang	97.23	40.4K	0.229
Dagu	97.46	43.7K	0.253
Bai'er	97.36	40.0K	0.222
Xianju	97.13	40.1K	0.235
Silkie	97.03	45.0K	0.258
Hubbard	97.88	46.1K	0.269
Cobb	98.70	45.2K	0.237
White Leghorn	97.99	43.5K	0.249

*MAF > 0.05,

**Across all 52.2 K loci.

An MDS analysis was performed using the genotyped data to investigate the ability of the 55K panel to detect population stratification in the validated samples. Figure 4 shows the relative coordinates of individuals when plotted using the two largest principal components. Individuals originating from the commercial broilers, Hubbard and Cobb tightly clustered. The Chinese indigenous meat-type breeds clustered together. The two Chinese indigenous egg-type breeds, Xianju and Bai'er, clustered together. The remaining local breeds (mainly characteristic of meat-types) were located relatively close to each other compared to egg-type breeds and commercial broilers. The commercial layer White Leghorn chickens were placed relative far away from to the Chinese local breeds and commercial broilers in Figure 2.4.

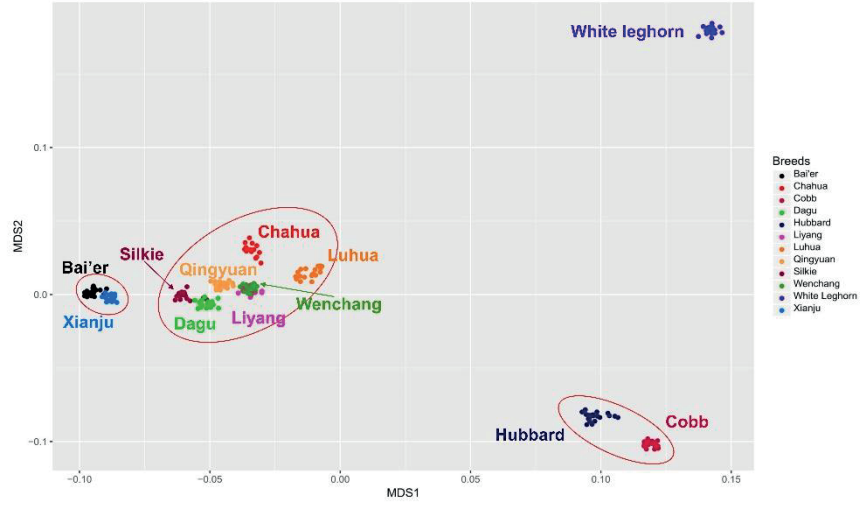


Figure 2.4 Results of multidimensional scaling analysis of 12 breeds/lines. The scatters show the individuals' position in the MDS plot, different colors represent different breeds/lines.

The linkage disequilibrium (LD) in Jingxing-Huang chicken and Cobb paternal line chicken respectively, were calculated. Figure 2.5 a and b shows the LD decay of the Jingxing-Huang and Cobb paternal lines for chromosome 1 and 2, respectively. The average levels of LD between adjacent SNPs in Jingxing-Huang breeds of chromosome 1 is 0.61 and for chromosome 2 is 0.58, whereas in the Cobb line these LD levels are 0.56 and 0.46, respectively. The mean LD level decay is around 0.22 in 40 Kb. The r2 of LD in the Jingxing-Huang chickens is larger than that in the Cobb paternal line. Supplemental figure 2.1 and 2.2 shows the r2 of LD decreasing with the increased SNPs distance in the two populations in whole genome level.

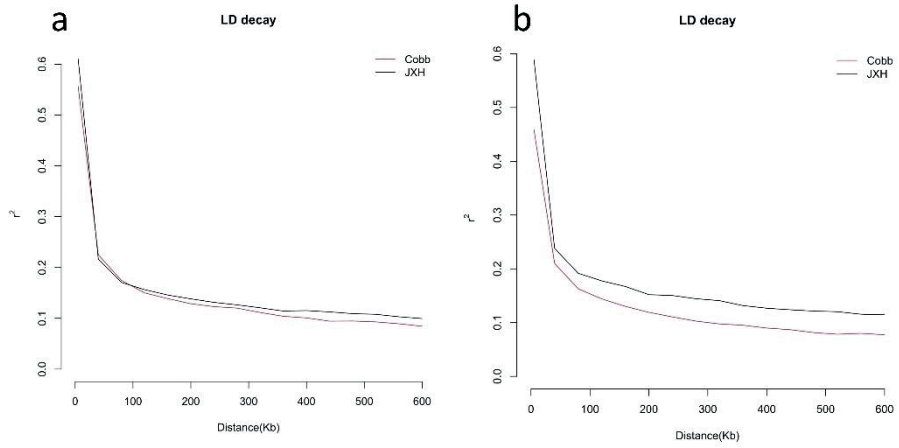


Figure 2.5 The LD decay plots. (a). from the Cobb and Jingxing-Huang (JXH) chickens in chromosome 1; (b). from Cobb and JXH chickens in chromosome 2.

2.3 Discussion

The selection of the 52,184 SNPs was performed in four groups and NCBI SNPdb using several criteria. The first group includes 140 K SNPs screened in genome sequences of eight Chinese local and commercial chickens. The Beijing-You chicken and Guangxi Sanhuang chicken are representative of Chinese, yellow-feathered chickens, which possess excellent meat quality and flavor (Qi et al., 2010). Two Cyan-shank partridge lines possess meat flavor and appearance that are usually chosen by consumers. The Jingxing-Huang line is widely used in local breeding programs because of its dwarfism, feed-saving and space-saving characteristics (Merat, 1984). The Cobb paternal line is a type of fast-growing line. The Recessive White chicken is a fast-growing line, which is popular in Chinese breeding programs because it can improve the growth rate of commercial generations without changing the appearance of offspring, when crossed with local colorful breeds. The SNPs with high ΔF between the local breeds and the commercial lines were used to determine the polymorphisms that have a larger difference in allelic frequency between breeds/lines. The main aim of whole-genome sequencing of different chicken breeds is to detect SNPs, although the pooled sequencing might generate potential bias.

The second, third and fourth groups are those potentially associated with economic traits, including 7.4 K SNPs associated with weight, carcass, immune and meat quality traits and 16.0 K SNPs for breast muscle, body fat and reproduction traits. Improving feed efficiency is an important goal in poultry industry to reduce costs.

RFI was considered independent of body weight and weight gain, selection for RFI would improve the feed efficiency without changing the economic traits (Aggrey et al., 2010). For special interests, 4.3 K SNPs were selected from a whole genomic sequencing research of low- and high- RFI Cobb and Beijing-You chickens. The strategy for the first selection of SNPs in candidate genes for the array is that these SNPs have a higher potential to be in linkage disequilibrium (LD) to the causative SNPs for the target traits. Finally, the last 18.4 K SNPs were selected from chicken SNPs database to make all SNPs cover the whole genome evenly. The average distance among SNPs is 22 kb generally (Figure 2). Due to specific selection on immune genes, a high SNP density on chromosome 16 is observed. Based on the limited information and substandard assembly (galGal-5.0) of the micro-chromosomes, we could not obtain SNPs of sufficient quality on all the micro-chromosomes. In summary, the setting of the algorithm was to select SNPs with relevant function and even distribution across the genome in terms of physical distance and obtain a representation of SNPs fron local or commercial breeds.

When comparing to the Affy 600K array, there are 24,227 SNPs that do overlap with the 55K array which accounts for 46%. The reason for this high percentage of overlap is that the 18.4 K SNPs for filling the gaps in the whole chicken genome were selected from chicken SNPdb, and the probe validated SNPs hold a high priority. This result showed that there were 21.4 K new SNPs included on the Affy 55K array compared to the two existing arrays. The results indicate that imputation of the 55K genotyping data to the high-density SNPs genotyping data is possible. In the new 55K genotyping array, 69% of SNPs are within genes (non-intergenic variant), the proportion is higher than the proportion in the Affy 600K array (54%), and lower than the proportion in Illumina 60K array (86%).

To investigate the ability of our 55K panel to detect polymorphisms and population structure in local or commercial breeds/lines. Nine Chinese local breeds (Chahua, Dagu, Liyang, Luhua, Qingyuan, Silkie, Wenchang, Bai'er, and Xianju) and 3 commercial lines (Hubbard, Cobb, and White Leghorn) were tested. The average call rate for each breed ranged from 97.0% to 98.7%. Across all populations, 76.7% to 88.0% SNPs were polymorphic (Table 2.5), which indicates that the 55K genotyping array can be used to determine genetic variation both in various local Chinese breeds and in commercial meat-type and egg-type breeds.

According to the results of MDS analysis (Figure 2.4), individuals originating from the commercial broilers, Hubbard and Cobb clustered together tightly and the two Chinese indigenous egg-type breeds, Xianju and Bai'er, clustered together. It might

be due to the fact that the two breeds were selected in the same direction (Resources, 2011). This result was supported by a previous study on the genetic diversity of Chinese domestic fowls based on a mtDNA analysis (BAO Wen-bin, 2008). The inter-population net nucleotide divergence (Da) between Xianju and Bai'er was 1.006, which was lower than the Da (1.115) between Xianju and Dagu chickens. The remaining local breeds (mainly characteristic of meat-types) were located relatively close to each other compared to egg-type breeds and commercial broilers. The commercial layer White Leghorn chickens were placed relative far away from the Chinese local breeds and commercial broilers. The relative proximity of Chinese local meat-type chicken and Chinese egg-type chicken in the MDS plot might be due to their shared region and ancestry. Thus, the MDS results are also in agreement with the existing knowledge of the lines/breeds and are also in agreement with the previous studies, which showed phylogenetic relationships among different chicken breeds.

The linkage disequilibrium (LD) in the Jingxing-Huang breeds and Cobb paternal line were calculated and compared. The mean LD level decay to around 0.22 in 40 Kb. This result is similar to the previous result (Fu et al., 2015). The r^2 of LD in the Jingxing-Huang breed is larger than that in Cobb paternal line. The Jingxing-Huang is an inbred line that has a relatively small effective population size whereas that of the Cobb paternal line is three to four times larger.

In China, yellow-feathered indigenous chickens are highly diverse (more than 100 local breeds and 70 crosses). The major obstacle in applying genomic selection for improvement of local breeds is the cost of genotyping array. The 55K array has a medium SNPs density, is cost-efficient, and optimal for Chinese local breeds compared with the existing 600K commercial array. Furthermore, the 55K genotyping array incorporated known SNPs loci that possess a high potential for association with economic traits and traits that are expensive and difficult to measure, which will be interesting for both GWAS and genomic selection (GS) projects.

With the rapid development of next-generation sequencing technologies and reduction of the costs, genotyping by sequencing will be the focus of future research. In the current phase, however, the genotyping by sequencing system is more complex and not as solid as the SNP array. The array genotyped data can be easily analyzed and standardized according to constant array SNP positions. The batch effect can be excluded by different laboratories and companies.

2.4 Conclusions

In conclusion, we developed an Affy 55K genotyping array that was designed to use SNPs that are segregating in Chinese local chicken breeds and commercial lines/breeds, and with large number of SNPs that are associated with economic traits. Compared to the existing Affy 600K and Illumina 60K arrays, 21.4 K new SNPs were included on the 55K SNP array. The results from our Affy 55K genotyping array can be imputed to the high-density SNP genotyping data. This array offers a wide range of potential applications, such as the evaluation of germplasm resources of chicken breeds, investigation of diversity of different chicken breeds, implementation of genome-wide association studies and genomic selection.

2.5 Methods

Animals

For whole genome sequencing, 384 chickens were sampled from eight local breeds or inbred lines (Table 2.1). Chickens were supplied by Institute of Animal Sciences of CAAS (local breed Beijing-You, inbred Jingxing-Huang line), Jiangsu Lihua Co. Ltd. (Cyan-shank Partridge lines with fast and mediate growth rates, respectively), Institute of Poultry Sciences of CAAS (Sanhuang chicken and Recessive White chicken), Xinguang Nongmu Co. Ltd. (paternal and maternal line of Cobb in parental generation). In addition, 15 to 21 chickens in each breed/line were used for SNP array evaluation, which were sampled from 9 local breeds and 3 commercial lines. Chickens were supplied by the Institute of Poultry Sciences of CAAS (Bai'er chicken, Chahua chicken, Dagu chicken, Liyang chicken, Qingyuan chicken, Silkie, Wenchang chicken, Luhua chicken and Xianju chicken), Xinguang Nongmu Co. Ltd. (paternal lines in parent generation from Cobb and Hubbard), the Institute of Animal Sciences of CAAS (White Leghorn). Two groups with 87 and 100 chickens from Jingxing-Huang and Cobb, respectively were also used for SNP array evaluation. The blood samples used in this study were all collected from chickens under the veterinary supervision and the Guidelines for Experimental Animals established by the Ministry of Science and Technology (Beijing, China), and with the approval of Animal Ethics Committee of the Institute of Animal Sciences. No anesthesia or euthanasia methods were used. There was no evidence at health examination that any of the involved chickens had clinical diseases caused by the sampling.

Whole genome re-sequencing

Genomic DNA was isolated from blood samples by the phenol-chloroform method. DNA quality was evaluated by gel electrophoresis and Nanophotometer. The

individual DNA samples (48 from each breed/line) were pooled to construct three libraries, with each library containing 8 males and 8 females. The libraries were constructed using the Nextera DNA Library Preparation Kit (Illumina Inc., San Diego, CA) according to the manufacturer's standard protocol. All libraries were sequenced on the Illumina Hiseq2500 (2 × 125 bp).

Genome sequence alignment and detection of the first Group of candidate SNPs

Reads were filtered for low quality (> 10 consecutive nucleotides with Phred scores < 10), adaptor sequences, and sequences without a quality control-passed paired read using NGSQC toolkit (v2.3.3) (Patel and Jain, 2012). Sequence coverage for each trimmed pool is shown in Table S5. Filtered sequenced reads were mapped to the reference genome (Gallus_gallus_4.0) by BWA software (v0.7.10) (Li et al., 2009). PCR duplicates were removed with *rmdup* function in Samtools (version 0.1.1.18) (Li et al., 2009). SNPs were identified and genotyped for each data set with *mpileup* function in Samtools, then called by VarScan (Koboldt et al., 2009). Only those highly confident variants supported by both methods were kept for downstream analyses. SNP calling parameters were described by Liu *et al.,* (2018). SNPs with MAF < 0.05 and INDELs in each breed/line were filtered by vcftools (Petr et al., 2011). To calculate the allele frequency between Chinese indigenous breeds and commercial lines (ΔF). We used Beijing-you chicken, Jingxing-Huang chicken, Sanhuang chicken, and the two lines of cyan-shank partridges minus the MAFs of Cobb paternal line, as well as the MAFs of Recessive White chicken, and the paternal and maternal generation of Cobb minus the MAFs of Beijing-You chicken, respectively. The SNPs with low ΔF were excluded. The value of ΔF was adjusted for 140K SNPs reserved in local breeds and commercial lines to generate the first group of candidate SNPs. The threshold of ΔF in local breeds and commercial lines are 0.609 and 0.731, respectively. SNPs acquired through genome re-sequencing of eight breeds/lines supplied the major data for the first group of SNPs in the array. SNPs specific for chromosome W were removed and were not considered in the current designing. There are also 25 INDELs of special interest, which were defined as Priory 1.

Selection of the second group of candidate SNPs based on GWAS analysis of 15 traits

The second group of candidate SNPs was selected according to a GWAS analysis of 15 traits. Phenotype and genotype data were generated from the CAAS chicken F2 resource population as previously described (Sun et al., 2013). Briefly, the population was derived from a cross between local Beijing-You chickens and

commercial Cobb broilers (Cobb-Vantress, Inc.). Weight, carcass, immune and meat quality traits were measured from 367 F2 chickens. The 15 traits were as follows, (a.) body weight at day 28 and day 42, (b.) carcass traits including total weight percentage after slaughtering, breast muscle weight percentage, leg muscle weight percentage, abdominal fat percentage, (c.) meat quality traits including the breast muscle intramuscular fat ratio, ultimate pH (24 h), meat lightness, redness value and yellowness value of breast muscle, (d.) immune traits including IgY level to sheep red blood cell, the heterophil and lymphocyte ratio, IgY level in serum, and the average red blood cell backlog.

SNPs were genotyped by using Illumina 60K SNP Bead chip for chicken (Groenen et al., 2011). All phenotypes have been described by Sun *et al.* (2013). To maximize the polymorphism resources for SNP array, the GLM procedures were used for the GWAS which was performed using the PLINK software (version 1.07) (Purcell et al., 2007) with 42,585 SNPs passed quality control. The SNPs with top 1% lowest p-values were used in the following procedures.

Selection of the third group of candidate SNP based on the associated genes for target traits

Known candidate genes for economic traits were collected and used for the SNP array design. All genes were identified through previous research by our group (Cui et al., 2012; Liu et al., 2016; Liu et al., 2017; Huang et al., 2018). We retrieved total 861 genes related to skeletal muscle and intramuscular fat development, chicken fat metabolism, and salmonella enteritidis resistance (Supplemental table 2.2). The SNPs were annotated by the Ensembl tool VEP (Mclaren et al., 2016). Mutations and the SNPs in exons, splicing regions, and UTRs were selected out. A maximum of 5 candidate SNPs was selected out for each gene.

In addition, the SNPs in this group also included 0.8 SNPs detected from a set of capture sequencing of Chr. 11, Chr. 16, and Chr. 19 of White Leghorns and Beijing-You chickens with low or high serum IgY (Liu *et al.*, unpublished, Supplement Table S3).

Selection of the fourth group of candidate SNPs for RFI

The fourth group candidate SNPs were selected from a whole genomic re-sequencing research of low- and high- RFI Cobb and Beijing-You chickens. SNPs calling results showed that 8.51 M and 8.48 M SNPs were detected in low- and high-RFI Beijing-You chickens, respectively; 8.35 M and 8.37 M SNPs were detected in low- and high-RFI Cobb chickens, respectively. The SNPs with Fst value <5% in

each breed were excluded followed by SNPs with mean ΔF < 0.35 between low- and high-RFI chickens. Through the above filtering processes, 3.7 K SNPs assigned to 1,137 candidate genes in Beijing-You chickens and 0.6 K SNPs (448 genes) in Cobb chickens were reserved (Liu et al., 2018).

Selection of the SNPs from chicken SNPs database

The first four groups cannot cover the whole genome evenly. In the fifth group, SNPs were selected from chicken SNPs database from NCBI (ftp://ftp.ncbi.nih.gov/snp/organisms/archive/chicken_9031/).

SNP screening according to the scoring of probes

All the SNPs' positions were transformed from WASHUC2.1 (Illumina 60K), and Gallus_gallus-4.0 (Affy 600K) to Gallus_gallus-5.0 (Affy 55K) by the LiftOver tool on UCSC Genome Browser. The five candidate groups above, in silico validation, was performed using the AxiomGTv1 algorithm of APT, which generated an output score file containing p-convert values, signifying the SNP array quality and list of recommended and non-recommended SNP probes. For a high-quality SNP array design, non-recommended SNP probes were all excluded in the following procedure.

SNPs selection procedure for the final 55K array

The final SNPs selection was done in multiple steps using several criteria. The roadmap is shown in Figure 1.

A custom-made algorithm was applied as described below. According to the Gallus_gallus-5.0, the chicken genome length is about 1.2 Gb. To ensure the probe position evenly distributed in the chicken genome, the whole genome was distributed by windows with 22 Kb length. The backward window started from the probe position of the forward probe position. The selection of the final array was performed on each chromosome separately. The first four groups of SNPs were divided as 2 priorities. The SNPs in group 2, group 3, group 4, and the INDELs in group 1 were defined as priority 1, and the SNPs in group 1 were defined as priority 2.

1. a) The selection of the SNPs in priority 1. If there is no SNP in a 22kb window, the window will be reserved. b) If there are one or two SNPs in the window, the SNP(s) was reserved. c) If there are 3 or more SNPs in a window, only 2 SNPs in this window will be reserved, which can make the SNPs even distributed in this window according to the following formula. $SD^2=\frac{(S-\bar{x})^2+(N_i-\bar{x})^2+(N_j-\bar{x})^2+(E-\bar{x})^2}{4}$. In the

formula above, the S and E are the start position and the end position of the window respectively; and N_i and N_j are the target SNPs position in the window. The SNPs N_i and N_j which can minimum the SD^2, will be reserved.

2. The selection of priority 2 SNPs. Each window reserved 1 or 2 SNPs will be skipped. The windows without SNP will be filled by one SNP of priority 2 according to the formula described above.

3. The windows without any SNP will be filled by 1 SNP from the NCBI SNPdb of chicken, while the validated SNPs will have a priority for filling.

The final array contains 55 K probes for 52 K SNPs, which were manufactured by Affymetrix® using photolithography. The redundant probes are used for interrogating each SNPs (Gunderson et al., 2005; Syvänen, 2005). The final 52K SNPs were annotated by the online tool Ensembl VEP (Zerbino et al., 2018).

The comparisons of the 55K Affy array with the existing arrays (Affy 600K array, and Illumina 60K)

All SNP positions were transformed from WASHUC2.1 (Illumina 60K), Gallus_gallus-4.0 (Affy 600K) and Gallus_gallus-5.0 (Affy 55K) to GRCg6a by the LiftOver tool on UCSC Genome Browser. All SNP positions of the three genotyping arrays were compared. SNPs on the 600K and 60K array were also analysed by Ensembl VEP (Mclaren et al., 2016). Overlapping Venn plot was performed by the Calculate and draw custom Venn diagrams (http://bioinformatics.psb.ugent.be/webtools/Venn/).

Validation of the 55K array in 13 chicken breeds/lines

The genomic DNA from 12 breeds/lines (Chahua, Dagu, Liyang, Luhua, Qingyuan, Silkie, Wenchang, Bai'er, and Xianju, Hubbard, Cobb, and White Leghorn) and two lines with larger populations (Jingxing-Huang and Cobb) were isolated as mentioned above. The genotyping was done on Axiom® arrays using the Affymetrix® GeneTitan® system according to the procedure described by Affymetrix (http://media.affymetrix.com/support/downloads/manuals/axiom_genotyping_solution_analysis_guide.pdf) in the Beijing Compass Biotechnology Co., Ltd (Beijing, China).

Basic genotype statistics for each marker, including call rate, MAF, Hardy-Weinberg Equilibrium (HWE), allele and genotype counts were calculated using the Quality Assurance Module from the SNP Variation Suite version 7 (SVS; Golden Helix Inc., Bozeman, Montana: www.goldenhelix.com). The following quality control criteria (filtering) were used to remove SNPs with less than 95% call rate for further

analysis. The SNPs with less than 0.05 MAF. SNPs were tested for HWE (P <0.001) to identify possible typing error. Samples with more than 10% missing genotypes were removed from the study.

The MDS was performed using the genotype data of the SNPs from the 55K panel on all breed samples (n = 226) to assess the utility of the panel in detecting population structure. Population structure between 12 breeds was carried out using PLINK software (version 1.90b3) (Purcell et al., 2007) with the MDS method, and the plot was performed by ggplot2 (Wickham, 2009). The linkage disequilibrium in 2 populations was performed by GAPIT (Alexander E et al., 2012). The LD decay plot performed by PopLDdecay software is presented as whole genome levels and as chromosome levels with the parameter of min break point size of 5 Kb and small break point size of 40 Kb (Zhang et al., 2019).

2.6 Declarations

Abbreviations

SNP: Single Nucleotide Polymorphism; GWAS: Genome-Wide Association Study; QTL: Quantitative Traits Loci; RFI: Residual Feed Intake; MAF: Minor Allele Frequency; UTRs: Untranslated Regions; LD: Linkage Disequilibrium; MDS: Multidimensional Scaling; CAAS: Chinese Academy of Agricultural Sciences; GLM: general liner model; Chr: Chromosome; HWE: Hardy-Weinberg Equilibrium; Da: inter-population net nucleotide divergence; VEP: Variant Effect Predictor; INDEL: insert/deletion.

Ethics approval and consent to participate

All experimental procedures with chickens were performed according to the Guidelines for Experimental Animals established by the Ministry of Science and Technology (Beijing, China). Ethical approval on animal survival was given by the animal ethics committee of the Institute of Animal Sciences (IAS), Chinese Academy of Agricultural Sciences (CAAS, Beijing, China) with the following reference number: IASCAAS-AE-03.

Availability of data and material

The whole genome sequencing clean data reported in this paper have been deposited in the Genome Sequence Archive (Wang et al., 2017) in BIG Data Center (Members, 2018) under accession number CRA001289 which can be publicly accessed at http://bigd.big.ac.cn/gsa.

Funding

The design of the study and collection of data were supported by the National Key Technology R&D Program (2015BAD03B03); the collection, analysis and interpretation of data were supported by the earmarked fund for the modern agro-industry technology research system (CARS-41) and Agricultural Science and Technology Innovation Program (ASTIP-IAS04; ASTIP-IAS-TS-15); the interpretation of data and writing the manuscript were supported by the National Nonprofit Institute Research Grant (2017ywf-zd-2).

Authors' contributions

RRL contributed to the design and performing of the study, the analysis and interpretation of data and writing of the manuscript. SX contributed to the data analysis, SNPs selection, and manuscript writing. JW contributed to the data analysis. RC contributed to the manuscript writing. MQZ, HXC, and QHL contributed to the sample and data collection. GPZ and JW contributed to the design of the study and interpretation of data. All authors submitted comments on the draft, read, and approved the final manuscript.

Acknowledgments

We would like to thank the Institute of Poultry Sciences of Chinese Academy of Agriculture Sciences, Jiangsu Lihua Co. Ltd., and Xinguang Nongmu Co. Ltd. for the provision of chicken samples. We would like to thank Prof. Jianfeng Liu (China Agricultural University) and Dr. Huamiao Liu (Institute of Special Animal and Plant Sciences of CAAS) for their help on the SNP selection work, and Prof. Martien AM Groenen (Wageningen University & Research) for his good comments and suggestions on the manuscript.

Supplemental files

Supplemental files can be found at https://osf.io/8uzmh/ .

Supplemental table 2.1 The summary of each sequencing pools. The raw reads number, clean reads number, sequencing depth, Q30 percentage, and coverage et al. for each sequencing library were provided.

Supplemental table 2.2 The second group of SNPs which related to the economic traits. The locus and the p-value of SNPs which related to 15 economic traits were provided.

Supplemental table 2.3 The third group of SNPs which related to 861 candidate genes. The information of 861 candidate genes and 118.4 K SNPs selected were provided.

Supplemental table 2.4 The third group of SNPs which related to serum IgY. The loci and allele information of 0.8 K SNPs related to serum IgY were provided.

Supplemental table 2.5 The loci information for the 55K array. The loci, allele information, SNPs frequencies in each breed/line, and overlap information were provided.

Supplemental figure 2.1 The LD decay in whole genome level in Cobb population.

Supplemental figure 2.2 The LD decay in whole genome level in Jingxing-Huang population.

References

Aggrey, S.E., Karnuah, A.B., Sebastian, B., and Anthony, N.B. (2010). Genetic properties of feed efficiency parameters in meat-type chickens. *Genetics Selection Evolution* 42(1), 1-5.

Alexander E, L., Feng, T., Qishan, W., Jason, P., Meng, L., Peter J, B., et al. (2012). GAPIT: genome association and prediction integrated tool. *Bioinformatics* 28(18), 2397.

Anderson, R., McEwan, J., Brauning, R., Kijas, J., Dalrymple, J., Worley, K., et al. (2014). Development of a high density (600K) Illumina Ovine SNP chip and its use to fine map the yellow fat locus. *Plant and Animal Genome XXII Conference*.

BAO Wen-bin, S.J.-t., WANG Cun-bo, ZHANG Hong-xia, Steffen Weigend, CHEN Guo-hong (2008). Investigation on Genetic Diversity and Systematic Evolut ion in Chinese Domestic Fowls and Red Jungle Fowls by Analyzing the mtDNA Control Region. *Acta veterinaria et zootechnica Sinica* 39(11), 1449-1459.

Cui, H.-X., Liu, R.-R., Zhao, G.-P., Zheng, M.-Q., Chen, J.-L., and Wen, J. (2012). Identification of differentially expressed genes and pathways for intramuscular fat deposition in pectoralis major tissues of fast- and slow-growing chickens. *BMC genomics* 13(1), 1-12.

Dash, S., Singh, A., Bhatia, A., Jayakumar, S., Sharma, A., Singh, S., et al. (2017). Evaluation of Bovine High-Density SNP Genotyping Array in Indigenous Dairy Cattle Breeds. *Animal Biotechnology*, 1-7.

Derks, M.F.L., Megens, H.J., Bosse, M., Visscher, J., Peeters, K., Bink, M.C.A.M., et al. (2018). A survey of functional genomic variation in domesticated chickens. *Genetics Selection Evolution* 50(1), 17.

Fan, W., Liu, H., Zheng, M., Liu, R., Li, Q., Wen, J., et al. (2015). Association of BMP15 Gene Polymorphisms with Egg Laying in Wuxing-Yellow Chicken. *Chinese Journal of Animal Science* 51(11), 13-18.

Fu, W., Dekkers, J.C., Lee, W.R., and Abasht, B. (2015). Linkage disequilibrium in crossbred and pure line chickens. *Genetics Selection Evolution* 47(1), 11.

Groenen, M.A., Megens, H.-J., Zare, Y., Warren, W.C., Hillier, L.W., Crooijmans, R.P., et al. (2011). The development and characterization of a 60K SNP chip for chicken. *BMC genomics* 12(1), 1-9.

Gunderson, K., Steemers, F., G, Mendoza, L., and Chee, M. (2005). A genome-wide scalable SNP genotyping assay using microarray technology. *Nature Genetics* 37(5), 549-554.

Houston, R.D., Taggart, J.B., Cézard, T., Bekaert, M., Lowe, N.R., Downing, A., et al. (2014). Development and validation of a high density SNP genotyping array for Atlantic salmon (Salmo salar). *BMC genomics* 15(1), 1-13.

Huang, H., Zhao, G., Liu, R., Li, S., Zhao, Z., Li, Q., et al. (2018). Expression profiles of novel genes and microRNAs involved in lipid deposition in chicken's adipocyte. *Italian Journal of Animal Science* 17(3), 593-598.

Iamartino, D., Nicolazzi, E.L., Van Tassell, C.P., Reecy, J.M., Fritz-Waters, E.R., Koltes, J.E., et al. (2017). Design and validation of a 90K SNP genotyping assay for the water buffalo (Bubalus bubalis). *PloS one* 12(10)**,** e0185220.

Koboldt, D.C., Chen, K., Wylie, T., Larson, D.E., Mclellan, M.D., Mardis, E.R., et al. (2009). VarScan: variant detection in massively parallel sequencing of individual and pooled samples. *Bioinformatics* 25(17)**,** 2283.

Kranis, A., Gheyas, A.A., Boschiero, C., Turner, F., Le, Y., Smith, S., et al. (2013). Development of a high density 600K SNP genotyping array for chicken. *Bmc Genomics* 14(1)**,** 59.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25(16)**,** 2078-2079.

Li, P., Fan, W., Everaert, N., Liu, R., Li, Q., Zheng, M., et al. (2018). Messenger RNA sequencing and pathway analysis provide novel insights into the susceptibility to Salmonella enteritidis infection in chickens. *Frontiers in Genetics* 9**,** 256.

Liu, J., Fu, R., Liu, R., Zhao, G., Zheng, M., Cui, H., et al. (2016). Protein profiles for muscle development and intramuscular fat accumulation at different post-hatching ages in chickens. *PLoS One* 11(8)**,** e0159722.

Liu, J., Liu, R., Wang, J., Zhang, Y., Xing, S., Zheng, M., et al. (2018). Exploring genomic variants related to residual feed intake in local and commercial chickens by whole genomic resequencing. *Genes* 9(2)**,** 57.

Liu, R., Wang, H., Liu, J., Wang, J., Zheng, M., Tan, X., et al. (2017). Uncovering the embryonic development-related proteome and metabolome signatures in breast muscle and intramuscular fat of fast-and slow-growing chickens. *BMC Genomics* 18(1)**,** 816.

Matukumalli, L.K., Lawley, C.T., Schnabel, R.D., Taylor, J.F., Allan, M.F., Heaton, M.P., et al. (2009). Development and Characterization of a High Density SNP Genotyping Assay for Cattle. *Plos One* 4(4)**,** e5350.

Mclaren, W., Gil, L., Hunt, S.E., Riat, H.S., Ritchie, G.R.S., Thormann, A., et al. (2016). The Ensembl Variant Effect Predictor. *Genome Biology* 17(1)**,** 122.

Members, B.D.C. (2018). Database Resources of the BIG Data Center in 2018. *Nucleic Acids Research* 46(Database issue)**,** D14-D20.

Merat, P. (1984). The sex-linked dwarf gene in the broiler chicken industry. *World's Poultry Science Journal* 40(1)**,** 10-18.

Muir, W.M., Wong, G.K., Zhang, Y., Wang, J., Groenen, M.A.M., Crooijmans, R.P.M.A., et al. (2008). Review of the initial validation and characterization of a 3K chicken SNP array. *Worlds Poultry Science Journal* 64(2)**,** 219-226.

Patel, R.K., and Jain, M. (2012). NGS QC Toolkit: A Toolkit for Quality Control of Next Generation Sequencing Data. *Plos One* 7(2)**,** e30619.

Petr, D., Adam, A., Goncalo, A., Albers, C.A., Eric, B., Depristo, M.A., et al. (2011). The variant call format and VCFtools. *Bioinformatics* 27(15)**,** 2156-2158.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., et al. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-

Based Linkage Analyses. *American Journal of Human Genetics* 81(3)**,** 559-575.

Qi, K., Chen, J., Zhao, G., Zheng, M., and Wen, J. (2010). Effect of dietary ω6/ω3 on growth performance, carcass traits, meat quality and fatty acid profiles of Beijing‐you chicken. *Journal of animal physiology and animal nutrition* 94(4)**,** 474-485.

Ramos, A.M., Crooijmans, R.P., Affara, N.A., Amaral, A.J., Archibald, A.L., Beever, J.E., et al. (2009). Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology. *Plos One* 4(8)**,** e6524.

Resources, C.N.C.o.A.G. (2011). *Animal Genetic Resources in China: Poultry.* China agriculture press.

Sun, Y., Zhao, G., Liu, R., Zheng, M., Hu, Y., Wu, D., et al. (2013). The identification of 14 new genes for meat quality traits in chicken using a genome-wide association study. *BMC genomics* 14(1)**,** 458.

Syvänen, A.-C. (2005). Toward genome-wide SNP genotyping. *Nature genetics* 37(6)**,** S5-S10.

Wang, Y., Song, F., Zhu, J., Zhang, S., Yang, Y., Chen, T., et al. (2017). GSA: Genome Sequence Archive. *Genomics Proteomics & Bioinformatics* 15(1)**,** 14-18.

Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis.* Springer Publishing Company, Incorporated.

Zerbino, D.R., Achuthan, P., Akanni, W., Amode, M.R., Barrell, D., Bhai, J., et al. (2018). Ensembl 2018. *Nucleic acids research* 46(D1)**,** D754-D761.

Zhang, B. (2011). China National Commission of Animal Genetic Resources. *Animal genetic resources in China: Horses, Donkeys, Camels. China Agricultural Press, China*.

Zhang, C., Dong, S.-S., Xu, J.-Y., He, W.-M., and Yang, T.-L. (2019). PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics* 35(10)**,** 1786-1788.

3

RNA-seq analysis reveals hub genes involved in chicken intramuscular fat and abdominal fat deposition during development

Siyuan Xing[1,2,†], Ranran Liu[1,†], Guiping Zhao[1], Lu Liu[1], Martien A. M. Groenen[2], Ole Madsen[2], Maiqing Zheng[1], Xinting Yang[1], Richard P. M. A. Crooijmans[2], and Jie Wen[1]

[1]State Key Laboratory of Animal Nutrition, Key Laboratory of Animal (Poultry) Genetics Breeding and Reproduction, Ministry of Agriculture and Rural Affairs, Institute of Animal Sciences, Chinese Academy of Agricultural Sciences, 100193 Beijing, China

[2]Animal Breeding and Genomics, Wageningen University & Research, 6708PB Wageningen, The Netherlands

[†] These authors contributed equally to this work.

Frontiers in Genetics (2020) 11:1009

Abstract

Fat traits are important in the chicken industry where there is a desire for high intramuscular fat and low abdominal fat. However, there is limited knowledge on the relationship between the dynamic status of gene expression and the body fat deposition in chicken. Transcriptome data were obtained from breast muscle and abdominal fat of female chickens from 9 developmental stages (from embryonic day 12 to hatched day 180). In total, 8545 genes in breast muscle and 6824 genes in abdominal fat were identified as developmental dynamic genes. Weighted correlation network analysis was used to identify gene modules and the hub genes. Twenty-one hub genes were identified, e.g., *ENSGALG00000041996* a candidate for high intramuscular fat and *CREB3L1* related to the low abdominal fat weight. The transcription factor *L3MBTL1* and the transcription factor co-factors *TNIP1*, *HAT1*, and *BEND6* showed a correlation to high breast muscle IMF and low abdominal fat weight in this study. Our results provide a resource of developmental transcriptome profiles in chicken breast muscle and abdominal fat. The candidate genes can assist in the selection for increased intramuscular fat content or a decrease in abdominal fat weight which would contribute to the improvement of these traits.

Key words: chicken, intramuscular fat, abdominal fat, transcriptome, tissue development

3.1 Introduction

Lipid metabolism, regulation, and deposition play a very important role not only in relation to obesity in humans but also in livestock production because of its close relationship with tasty and healthy food supply for humans. For global meat consumption, chicken meat is the second largest, providing one fourth of meat resource (http://www.fao.org/faostat/en/#home). In the Chinese meat-type chicken industry, yellow-feathered dwarf chickens are used in one-third of the breeding system. High intramuscular fat (IMF) content contributes to high meat quality, as a result, increasing IMF deposition is a desirable goal in meat-type chicken breeding. Genetic selection, nutritional strategies, and management practices have been shown to enhance fat deposition and IMF in swine (Reiter et al., 2007) and cattle (Pethick et al., 2001). However, unlike the muscle type of pork and beef, the marbling in chicken muscle meat is almost invisible. In chicken, an increased IMF in muscle tissue will result in an increase in abdominal fat (AF) deposition in the chicken body (Jiang et al., 2017). Excess of AF influences animal welfare and becomes a waste product for human consumption after slaughter, therefore resulting in considerable economic losses (Jiang et al., 2017). Thus, an increase of IMF and a reduction of AF deposition are important goals of meat-type chicken production.

Fat can be deposited at different sites in the chicken body: around abdominal tissues (AF, also called visceral fat or central fat), in bones (marrowfat), subcutaneous (subcutaneous fat), and in the muscle (IMF). The IMF content plays a key role in various quality traits of meat and it varies between different chicken breeds/lines, tissue types, and also varies with age, gender, feeding, and even during the season (Hocquette et al., 2010). AF is the most dominant fat tissue in the mature animal body. Fat tissue is composed of adipocytes, which mainly differentiate from mesenchymal stem cells (MSCs) (Pittenger et al., 1999). Adipocyte differentiation should be characterized by 2 phases, the determination phase (hyperplasia) and the terminal differentiation phase (hypertrophy) (Symonds, 2012). Although it has been suggested that the number of adipocytes will not increase after adulthood, in humans, prolonged times of obesity can also result in an increase in the number of adipocytes (Schmitz et al., 2016). For meat-type animals, at the cellular level, the adipocyte number increases most rapidly in the abdominal wall and minimal in the intramuscular depot (Allen, 1976). Chicken fat deposition varies during the different developmental stages. In embryonic stages, the fat deposition starts in the muscle (IMF) before deposition around the abdomen, while at the fast-growing stage this is the other way around.

The gain in fat depends on the adipocyte's ability to synthesize and store lipids. The molecular mechanisms underlying fat deposition and its regulation are still insufficiently understood but there is a close relationship between adipocyte development and expression of specific genes in pre-adipocytes. This involves genes related to adipocyte differentiation, transcription regulators and genes related to lipid metabolism (Ono et al., 1990). The transcription factors (TFs) *PPARγ*, *C/EBPs*, and *ADD1* (*SREBP1*) are involved in the regulation of adipocyte differentiation. *PPARγ* was shown to be a necessary regulator of induced differentiation of adipocytes (Kim et al., 2014). *C/EBPα* plays a very important role in adipocyte differentiation (Tang and Lane, 1999) and activates genes such as *aP2*, *PEPCK*, and *SCD1*, that all contain TF binding sites for *C/EBPα* (Farmer, 2006).

RNA sequencing has been used in studies of chicken pre-adipocytes development at the cellular level (Guo et al., 2018), chicken fat deposition in vivo (Zhuo et al., 2015), and embryonic adipocytes development (Na et al., 2018). Adipogenesis has been shown to be a multi-step process, regulated by both enhancers and inhibitors (Tontonoz et al., 1994; Ross et al., 2000). Several tissues are involved in the regulation of fat deposition and the contribution of these tissues changes during development. Previous studies have mainly focused on IMF and AF separately or mainly focused on one or two developmental stages (Resnyk et al., 2017). Therefore, this study focuses on multiple time points of development and the transcriptome dynamic changes of two different fat-related tissues to achieve more completable knowledge on the molecular mechanism of fat deposition in chicken.

3.2 Materials and methods

Animals Genetic Background, Phenotypes, and Samples Collection

The parental generation used in this study was selected from an inbred dwarf yellow-feathered Jingxing-Huang IMF-up selected chicken line, which is a widely used Chinese local meat-type chicken line (Jiang et al., 2017). Twenty roosters and 60 hens (1 male mated to 3 females) were selected to produce the animals of the experimental generation for the phenotype recording and sample collection. In the experimental generation, 2 batches of eggs were incubated. Sample collection was subsequently performed at the following 9 developmental stages: E12 (embryonic day 12), E17, and D01 (day 1 after hatching), D07, D21, D56, D98, D140, and D180. Chickens were reared with *ad libitum* access to feed and water. Tissue sampling of the animals was approved by the animal ethics committee of the Institute of Animal Sciences, Beijing, China. The following phenotypes were recorded: body weight, breast muscle weight (BMW), and AF weight. The organ growth curve of

breast muscle (BM) and AF was fitted by the Logistic model using the Origin software. BM and AF samples were collected from every animal and developmental stage except for AF from stage E12 to D01, where no obvious AF tissue is observed. From developmental stages E12 to D21, hematoxylin-eosin (HE) and Red Oil O stain were used on the BM samples and the relative amount of BM IMF during these phases was measured by the Red Oil O-stained section. From D21 to D180, the IMF content in BM of the chickens was determined by the Soxhlet extraction method (Soxhlet, 1879). The relative breast muscle IMF content from E12 to D07 was calculated by the IPP software from 10 captured images of Red Oil O-stained sections. The genders of the embryo were determined by a length polymorphism in the intron of the *CHD1* gene by performing a PCR and analysis of the fragments using agarose gel electrophoresis (Griffiths and Korn, 1997). The sequences of the primers are forward primer: 5'-GTTACTGATTCGTCTACGAGA-3', and reverse primer: 5'-ATTGAAATGATCCAGTGCTTG-3'. Finally, three full-sib families were used as experimental chickens. Each full-sib family provided one chicken for samples for RNA-seq in each stage. the middle of AF and the pectoralis major of BM samples from 27 female chickens were used for RNA extraction (Table S3.1). Additionally, in the embryonic period, it is not possible to divide the pectoralis major and pectoralis minor, the whole breast muscle was used for RNA isolation.

RNA sequencing and data quality control

The QIAGEN RNeasy Kit was used to isolate total RNA, and genomic DNA was removed by using the TIANGEN DNase KIT. The RNA concentration and RNA integration number were assessed by Nanophotometer and Nanodrop, respectively. The RNA samples with RIN > 7 were used to isolate mRNA from total RNA by the Dynabeads mRNA DIRECT Kit (Invitrogen) followed by library construction. Un-stranded specific RNA sequencing libraries were sequenced on the Illumina Hiseq2500 (2x125 bp). Library construction and sequencing were commercially performed by Berry Genomics, Beijing, China. Obtained sequences were trimmed for the sequencing adaptors and for low-quality reads (N > 10% in a read) by Trimmomatic (Bolger et al., 2014). The sequence data quality of each sample was controlled by FastQC (Andrews, 2010).

Transcriptome Profiling and differentially expressed genes (DEGs) Detection

All trimmed transcriptome data was aligned to the chicken reference genome (GRCg6a) and annotation file (Gallus.gallus.GRCg6a.95.gtf) by the STAR software (version 2.5.3) (Dobin et al., 2013). Data was assembled by the Stringtie software (version 1.3.3b) (Pertea et al., 2015). Gene and transcript level raw counts were

calculated using the Stringtie provided Python script with the parameter l = 125. The accuracy of the assembled files was evaluated by gffcompare (version 0.10.1). The protein coding genes and the long non-coding genes were all considered in the downstream analysis. Gene expression level normalization was performed by DESeq2 (Love et al., 2014), which is based on the experimental design as Stage + Tissue + Family. The normalized gene expression data were used for downstream analysis. The within tissue PCA plots were performed by the distance of the samples calculated by rlog, the PCA for all samples was performed by the distance of the samples calculated by vst function of DESeq2 (Love et al., 2014). Using the Benjamini-Hochberg method (Benjamini and Hochberg, 1995) with adjusted-p < 0.05, genes with an expression fold change (FC) > 1.5 or FC < 0.67, were considered as DEGs, which is based on the experimental design as Stage.

Pathway Analysis

The KEGG enrichment and GO enrichment were performed by clusterProfiler package version 3.11.1 (Yu et al., 2012) with org.Gg.eg.db package version 3.8.2 (Carlson, 2019) and KOBAS 3.0 (Chen et al., 2011).

Statistical analysis

The student t-test was performed using basic R (version 3.6.0) after using the function of the shapiro.test (for normality test) and the bartlett.test (for homogeneity test of variance). The data sets, which do not fit the normal distribution, were compared by the rank sum test (Kassambara, 2017). LSD-test was performed by agricolae package (version 1.3.1) (De Mendiburu, 2014). Significance was stated at $P < 0.05$.

Development dynamics genes (DDGs) identification in 2 tissues

The normalized gene expression data for the different developmental stages of BM and AF were used in the DDG analysis. Genes with average raw counts lower than 1 were excluded. The DDGs were identified by maSigPro package (Ana et al., 2006; Nueda et al., 2014). By considering the expression distribution as the negative binomial model and the Benjamini and Hochberg procedure to adjust the FDR, significant genes were selected with the forward method using $r^2 > 0.7$. Gene expression patterns in tissues pairwise comparison was performed with the same parameters described above. The list of TFs and TF co-factors (TFCF) was acquired from AnimalTFDB (v.3.0) (Hu et al., 2018).

Weighted gene co-expression network analysis (WGCNA)

All samples were used in the WGCNA except three BM outlier samples (Figure S3.1). The remaining BM samples for the WGCNA did cover all the stages. A weighted genes co-expression analysis was performed by WGCNA package (Langfelder and Horvath, 2008) with default settings and minor modifications. The minModuleSize was set to 100 and mergeCutHeight was set to 0.3 for tissue-stage specific module detection (soft-threshold = 9). A tissue-stage matrix for each RNA-seq sample was built for correlation to identify the modules. For the WGCNA within tissue data set, suitable soft-threshold power values were chosen based on the approximate scale-free topology for each analysis (soft-threshold =18 for BM in E12 to D21 data set, soft-threshold = 10 for BM in D07 to D180 data set, and soft-threshold = 5 for AF data set). By using the step-by-step topology overlap matrix (TOM), module detection, and similar module merging functions (minModuleSize = 30), gene modules co-expression clustering dendrograms were built (Langfelder and Horvath, 2008). The module-traits associations were quantified, and the corresponding correlations were adjusted by the method of Benjamini-Hochberg. The P-value of interested module-traits corresponding correlations less than 0.01 were used for further analysis. To identify the hub genes in the interesting modules, a customized hub gene filtering method was used. The gene network of each module was filtered as follows: a) the edges with weight lower than 0.15 were removed, b) the nodes with a connectivity number smaller than 10 were removed, c) the nodes with an average expression below 10 were removed, for controlling the false positive rate, d) finally, the gene is ranked by the summation of weight value. Genes co-expression networks were performed by Cytoscape software (version 3.7.0) (Shannon et al., 2003) with the edges provided by WGCNA 'exportNetworkToCytoscape' function.

3.3 Results

For the RNA sequenced chickens, the IMF ratio of BM and the AF weight at the 9 different stages were determined (Table S3.2). The fitted growth curve of BM and AF for each of the three full-sib families during development are shown in Figure S3.2, and both the breast muscle weight and abdominal fat weight follows a logistic regression ($R^2 > 0.99$). On the day of hatch (D01), the IMF ratio of BM was relatively high and then dropped to a low level at D07, similar with E12 and E17. From D21 to D98 IMF ratios stayed relatively constant after which it gradually increased to 7.04% at D140 (Figure 3.1K). The HE stained section of BM from embryonic day 12 to day 21 showed that the diameter of muscle fiber increased according to the developmental stage (Figure 3.1A to 3.1E). The AF weight (AFW) weekly gain from D56 to D98 and D98 to D140 was significantly higher than those between the other

developmental stages, whereas D140 to D180 showed a decrease in AFW weekly gain (Figure 3.1L).

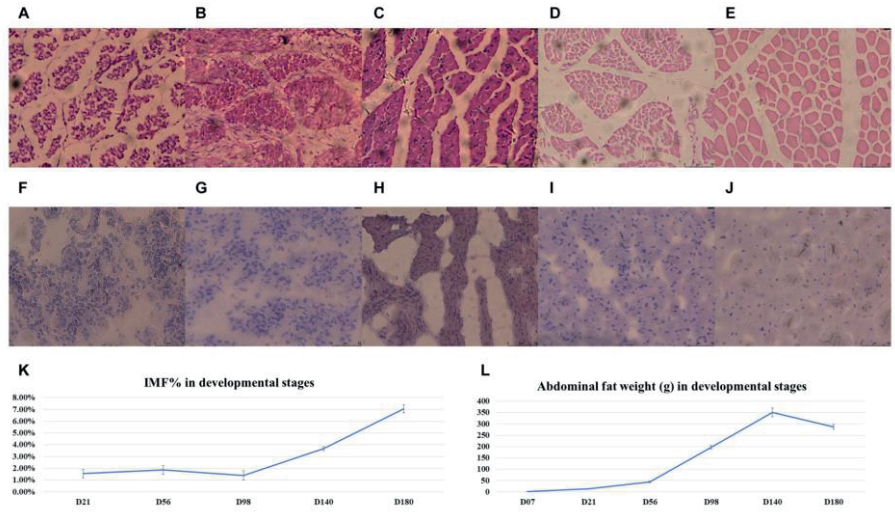


Figure 3.1 The phenotypes of the chickens used for RNA-Seq. (A–E) The HE-stained breast muscle sections in E12, E17, D01, D07, and D21. (F–J) The Red Oil O-stained breast muscle sections in E12, E17, D01, D07, and D21. The (K) IMF percentages and (L) average abdominal fat weight (gram); error bars are the standard deviations.

Transcriptome Profiling

To obtain insight in the transcription of genes during the 9 different developmental stages, transcriptome data was obtained from BM and AF from 3 individuals per stage. In total, 45 RNA-seq libraries were constructed and sequenced (Table S3.2). After trimming adaptors and removing low-quality reads, an average of 28.58 million reads per library were aligned to the chicken reference genome (GRCg6a) with a mean alignment ratio of 92.48% over all libraries (Table S3.3). In total, 21,853 genes were detected among all samples with 20,891 genes expressed in BM and 20,719 genes expressed in AF across all the tested developmental stages. Among which, 90.41% (19757/21853) of genes both expressed in BM and AF. The overlap of the BM and AF expressed genes is shown in Figure S3.3. The genes raw read counts in each library are shown in Table S3.4. The most highly expressed genes in BM and AF are *ACTA1*(actin alpha 1) and *MT-CO1*(Cytochrome c oxidase subunit 1), respectively.

To explore whether the expression profiles correlate with the developmental stages, a combined PCA of BM and AF expressed genes was performed (Figure 3.2)

as well as individual PCA for BM and AF, respectively (Figure S3.4a and Figure S3.4b). As expected, there is a strong separation of the two tissues (Figure 3.2) whereas limited separations were observed for the developmental stages in the two tissues.

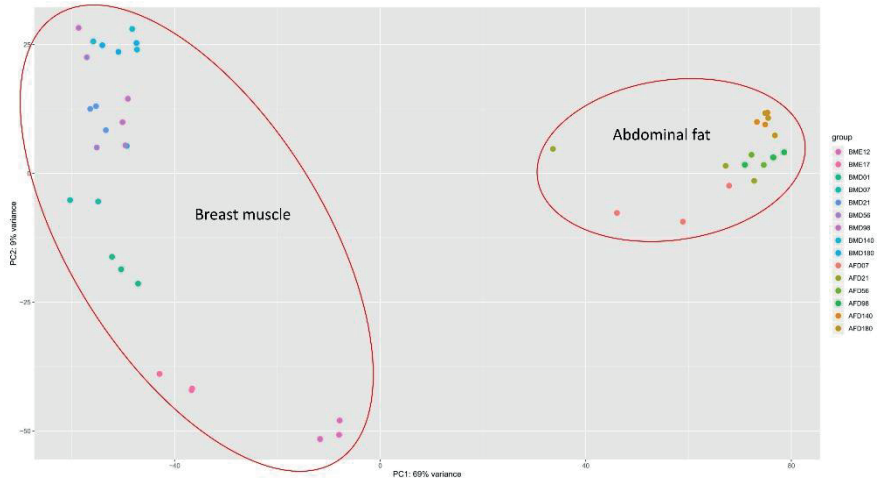


Figure 3.2 PCA plot of BM and AF samples at different developmental stages. The legends represent tissue plus developmental stage. The PCA was calculated by the vst function of DESeq2 package based on the normalized raw gene counts.

The stage specific expressed gene numbers varied from 55 to 708 for BM and from 80 to 694 for AF (Table S3.5, Table S3.6 and Figure S3.5a). The KEGG enriched pathways of the genes specifically expressed in different developmental stages are shown in Figure S3.5b and Figure S3.5c, *e.g.,* the genes specially expressed in D01 enriched for fatty acid related pathways, such as fatty acid elongation, biosynthesis of unsaturated fatty acids, and fatty acid metabolism. The number of differentially expressed genes (DEGs) detected between the adjacent developmental stages varied from 13 to 1432 for BM and from 48 to 1177 for AF (Table S3.7). The number of DEGs between the early stages is higher than between the latter stages. The number of DEGs between D140 and D180 both in BM and AF is very limited. The KEGG enrichment pathways of DEGs between adjacent stages in BM and AF are shown in Figures S3.6a and S3.6b. In the early stages of both BM and AF, the DEGs are enriched in cell cycle and cell adhesion molecules pathways. The DEGs between D21 and D56 in BM are enriched in glycerolipid metabolism. The DEGs between D56 and D98 in AF are enriched in steroid biosynthesis pathway.

Developmentally Dynamic Genes (DDGs)

The genes which showed significant temporal changes in expression were identified in the two tissues and these genes were considered as developmentally dynamic genes. The DDGs reflect the changes across developmental stages in gene expression regulation as well as in biological processes. In the BM data set, 8545 genes were identified as DDGs, including 425 TFs and 392 TF cofactors. In the AF data set, 6824 DDGs were identified including 357 TFs and 305 TF cofactors (Table S3.8). On average around 1/3 of these DDGs overlap between the two tissues (Figure 3.3a to 3.3c). The TFs and TF cofactors of BM and AF DDGs are enriched in similar pathways e.g., cellular senescence and AGE-RAGE signaling pathways in diabetic complications (Figure S3.7). The full list of DDGs is presented in Table S3.9. There is 37.48% and 31.48% of all currently identified TFs dynamically expressed in BM and AF, respectively.

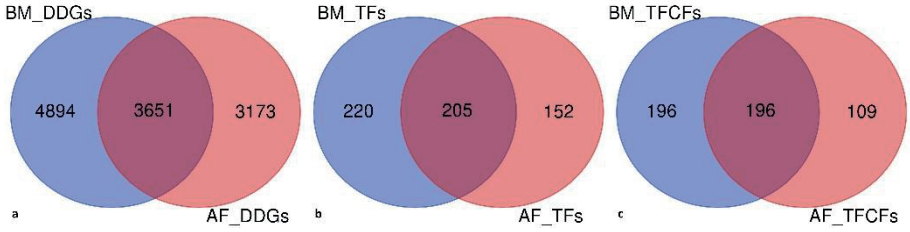


Figure 3.3 BM and AF DDGs overlap. a. Overlap of DDGs between BM and AF. b. Overlap of DDGs TFs between tissues. c. Overlap of DDGs TF co -factors between BM and AF.

Expressed Genes in Different Developmental Stages

From the whole data set, we investigated the stage-specific expressed genes in breast muscle and abdominal fat by weighted gene co-expression network analysis (WGCNA). The genes with similar expression patterns were clustered by the topology overlap matrix. The merged cluster dendrogram is shown in Figure 3.4a. In total, 34 co-expression gene modules were detected. A module can be considered as a group of clustered genes and is color-coded. The module-trait relationships of AF and BM are shown in Figure 3.4b and the co-expressed gene modules were positively correlated with the developmental stages and the tissues. The genes in the cyan, light-yellow, black, and red modules are mainly expressed during the early stages of BM and are enriched for cell cycle, DNA replication, spliceosome, and mismatch repair pathways. The genes in the brown, white, purple, and grey60 modules expressed at embryonic stage day 12 of BM are

enriched for terms like cell cycle, spliceosome, RNA transport, DNA replication, mismatch repair, and homologous recombination.

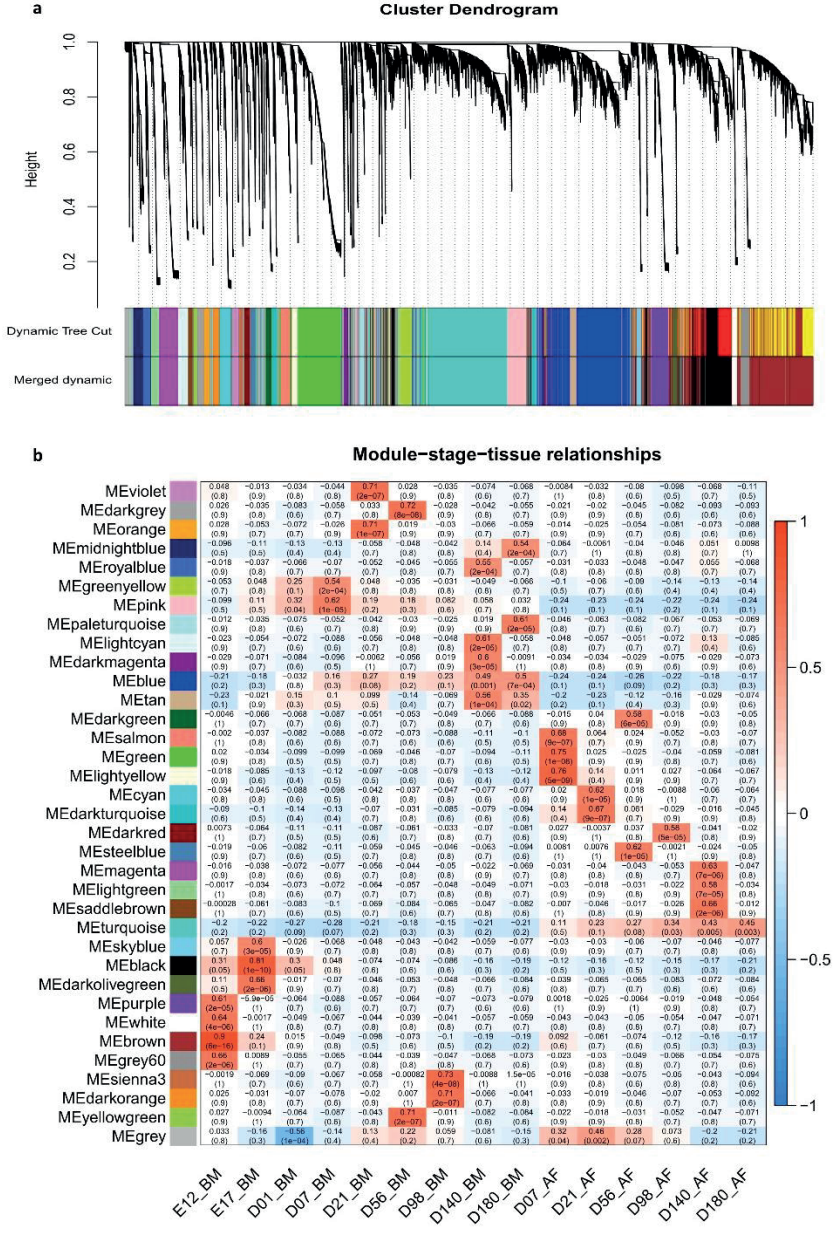
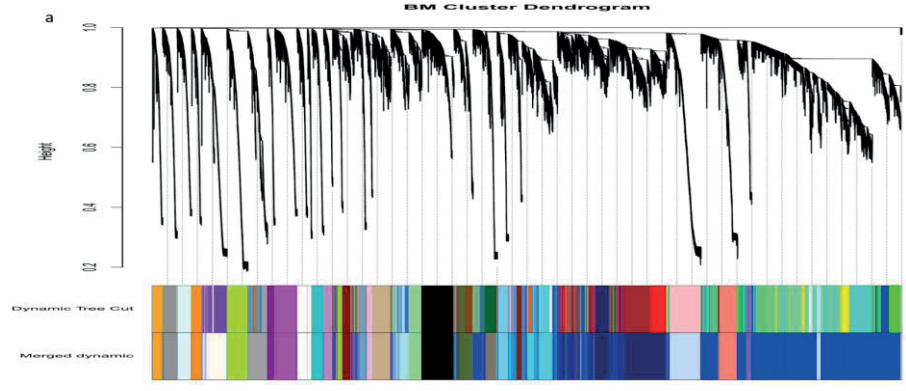


Figure 3.4 WGCNA results of BM and AF as a consensus data set. a. Cluster dendrogram of the BM and AF. b. Module-stages-tissue relationship of BM and AF.

Detection of Hub Genes and Transcription (co) Factors Related to High IMF in Breast Muscle

To detect the hub-genes and the related transcription (co) factors involved in breast muscle IMF deposition, we performed the WGCNA on two different BM data sets. The rational for two separate analyses is that BM has two adipocyte development phases. Thus, a WGCNA analysis on the whole BM data set could result in potential false results. The first phase is from developmental stage E12 to D21 and is mainly related to the adipocyte hyperplasia, and the second phase is from D07 to D180 mainly related to the adipocytes hypertrophy. Consequently, developmental stages D07 and D21 were included in both phases.

The module cluster dendrogram of the WGCNA results of the first phase in the BM data set (E12 to D21) is shown in Figure 3.5a and the module-trait relationship is shown in Figure 3.5b. There are 22 modules clustered in the first phase BM gene co-expression data set (Figure 3.5a). The grey60 and light-yellow modules were significantly positively correlated with BM IMF content (P = 0.002 and P = 0.01, respectively, Figure 3.5b). The network of eigengenes for this data set is shown in Figure S3.8a. The co-expression network of the grey60 module is shown in Figure 3.5c. After filtering the edges with weight, connectivity, and the filtered weight summary of each node, the genes *ENSGALG00000053368*, *COX6A1*, *ATG9B*, and *ENSGALG00000041996* were identified as the hub genes in the grey60 module (Table 3.1). *ENSGALG00000041996* contacted genes are enriched in e.g., carbon metabolism, valine degradation, fatty acid metabolism, 2-Oxocarboxylic acid metabolism, and fatty acid elongation pathways. In the light-yellow module, *LOC107050564* and *ENSGALG00000048510* were identified as the hub genes (Figure 3.5d). The detected TFs *MYCN* and *HOXB1*, the TF-cofactor *RNF168* and *ENSGALG00000008349* are involved in the light-yellow module.



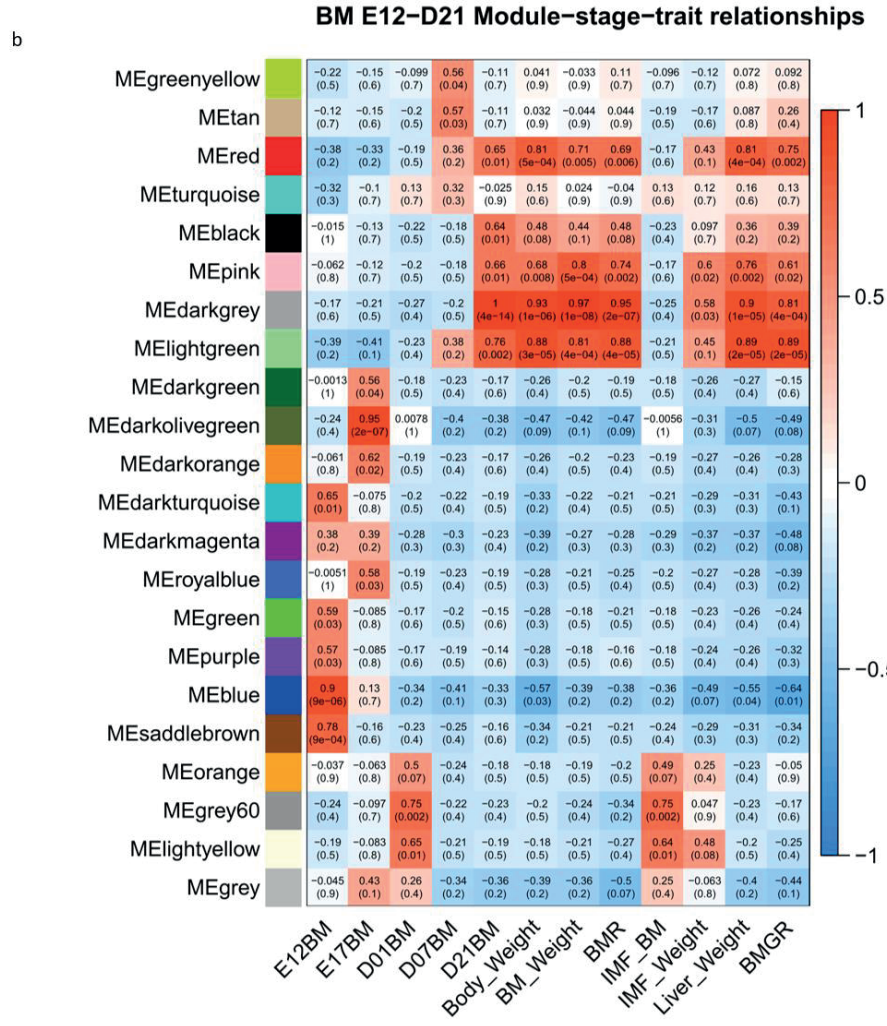
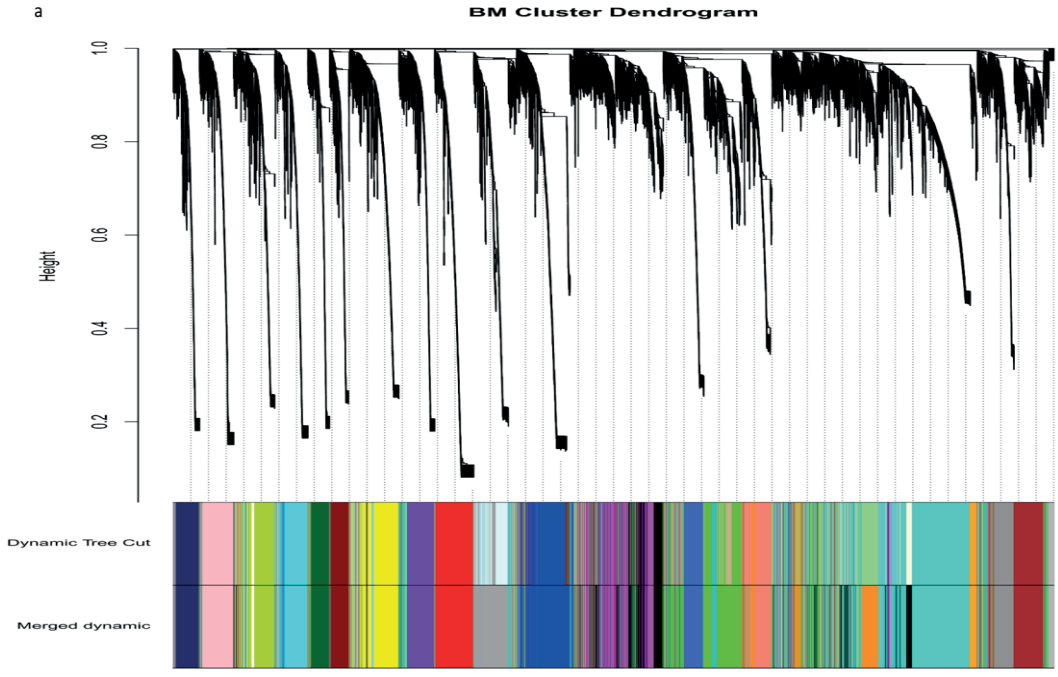

b
BM E12−D21 Module−stage−trait relationships
MEgreenyellow −0.22 (0.5) −0.15 (0.6) −0.099 (0.7) 0.56 (0.04) −0.11 (0.7) 0.041 (0.9) −0.033 (0.9) 0.11 (0.7) −0.096 (0.7) −0.12 (0.7) 0.072 (0.8) 0.092 (0.8)
MEtan −0.12 (0.7) −0.15 (0.6) −0.2 (0.5) 0.57 (0.03) −0.11 (0.7) 0.032 (0.9) −0.044 (0.9) 0.044 (0.9) −0.19 (0.5) −0.17 (0.6) 0.087 (0.8) 0.26 (0.4)
MEred −0.38 (0.2) −0.33 (0.2) −0.19 (0.5) 0.36 (0.2) 0.65 (0.01) 0.81 (5e−04) 0.71 (0.005) 0.69 (0.006) −0.17 (0.6) 0.43 (0.1) 0.81 (4e−04) 0.75 (0.002)
MEturquoise −0.32 (0.3) −0.1 (0.7) 0.13 (0.7) 0.32 (0.3) −0.025 (0.9) 0.15 (0.6) 0.024 (0.9) −0.04 (0.9) 0.13 (0.6) 0.12 (0.7) 0.16 (0.6) 0.13 (0.7)
MEblack −0.015 (1) −0.13 (0.7) −0.22 (0.5) −0.18 (0.5) 0.64 (0.01) 0.48 (0.08) 0.44 (0.1) 0.48 (0.08) −0.23 (0.4) 0.097 (0.7) 0.36 (0.2) 0.39 (0.2)
MEpink −0.062 (0.8) −0.12 (0.7) −0.2 (0.5) −0.18 (0.5) 0.66 (0.01) 0.68 (0.008) 0.8 (5e−04) 0.74 (0.002) −0.17 (0.6) 0.6 (0.02) 0.76 (0.002) 0.61 (0.02)
MEdarkgrey −0.17 (0.6) −0.21 (0.5) −0.27 (0.4) −0.2 (0.5) 1 (4e−14) 0.93 (1e−06) 0.97 (1e−08) 0.95 (2e−07) −0.25 (0.4) 0.58 (0.03) 0.9 (1e−05) 0.81 (4e−04)
MElightgreen −0.39 (0.2) −0.41 (0.1) −0.23 (0.4) 0.38 (0.2) 0.76 (0.002) 0.88 (3e−05) 0.81 (4e−04) 0.88 (4e−05) −0.21 (0.5) 0.45 (0.1) 0.89 (2e−05) 0.89 (2e−05)
MEdarkgreen −0.0013 (1) 0.56 (0.04) −0.18 (0.5) −0.23 (0.4) −0.17 (0.6) −0.26 (0.4) −0.2 (0.5) −0.19 (0.5) −0.18 (0.5) −0.26 (0.4) −0.27 (0.4) −0.15 (0.6)
MEdarkolivegreen −0.24 (0.4) 0.95 (2e−07) 0.0078 (1) −0.4 (0.2) −0.38 (0.2) −0.47 (0.09) −0.42 (0.1) −0.47 (0.09) −0.0056 (1) −0.31 (0.3) −0.5 (0.07) −0.49 (0.08)
MEdarkorange −0.061 (0.8) 0.62 (0.02) −0.19 (0.5) −0.23 (0.4) −0.17 (0.6) −0.26 (0.4) −0.2 (0.5) −0.23 (0.4) −0.19 (0.5) −0.27 (0.4) −0.26 (0.4) −0.28 (0.3)
MEdarkturquoise 0.65 (0.01) −0.075 (0.8) −0.2 (0.5) −0.22 (0.4) −0.19 (0.5) −0.33 (0.2) −0.22 (0.4) −0.21 (0.5) −0.21 (0.5) −0.29 (0.3) −0.31 (0.3) −0.43 (0.1)
MEdarkmagenta 0.38 (0.2) 0.39 (0.2) −0.28 (0.3) −0.3 (0.3) −0.23 (0.4) −0.39 (0.2) −0.27 (0.3) −0.28 (0.3) −0.29 (0.3) −0.37 (0.2) −0.37 (0.2) −0.48 (0.08)
MEroyalblue −0.0051 (1) 0.58 (0.03) −0.19 (0.5) −0.23 (0.4) −0.19 (0.5) −0.28 (0.3) −0.21 (0.5) −0.25 (0.4) −0.2 (0.5) −0.27 (0.4) −0.28 (0.3) −0.39 (0.2)
MEgreen 0.59 (0.03) −0.085 (0.8) −0.17 (0.6) −0.2 (0.5) −0.15 (0.6) −0.28 (0.3) −0.18 (0.5) −0.21 (0.5) −0.18 (0.5) −0.23 (0.4) −0.26 (0.4) −0.24 (0.4)
MEpurple 0.57 (0.03) −0.085 (0.8) −0.17 (0.6) −0.19 (0.5) −0.14 (0.6) −0.28 (0.3) −0.18 (0.5) −0.16 (0.6) −0.18 (0.5) −0.24 (0.4) −0.26 (0.4) −0.32 (0.3)
MEblue 0.9 (9e−06) 0.13 (0.7) −0.34 (0.2) −0.41 (0.1) −0.33 (0.3) −0.57 (0.03) −0.39 (0.2) −0.38 (0.2) −0.36 (0.2) −0.49 (0.07) −0.55 (0.04) −0.64 (0.01)
MEsaddlebrown 0.78 (9e−04) −0.16 (0.6) −0.23 (0.4) −0.25 (0.4) −0.16 (0.6) −0.34 (0.2) −0.21 (0.5) −0.21 (0.5) −0.24 (0.4) −0.29 (0.3) −0.31 (0.3) −0.34 (0.2)
MEorange −0.037 (0.9) −0.063 (0.8) 0.5 (0.07) −0.24 (0.4) −0.18 (0.5) −0.18 (0.5) −0.19 (0.5) −0.2 (0.5) 0.49 (0.07) 0.25 (0.4) −0.23 (0.4) −0.05 (0.9)
MEgrey60 −0.24 (0.4) −0.097 (0.7) 0.75 (0.002) −0.22 (0.4) −0.23 (0.4) −0.2 (0.5) −0.24 (0.4) −0.34 (0.2) 0.75 (0.002) 0.047 (0.9) −0.23 (0.4) −0.17 (0.6)
MElightyellow −0.19 (0.5) −0.083 (0.8) 0.65 (0.01) −0.21 (0.5) −0.19 (0.5) −0.18 (0.5) −0.21 (0.5) −0.27 (0.4) 0.64 (0.01) 0.48 (0.08) −0.2 (0.5) −0.25 (0.4)
MEgrey −0.045 (0.9) 0.43 (0.1) 0.26 (0.4) −0.34 (0.2) −0.36 (0.2) −0.39 (0.2) −0.36 (0.2) −0.5 (0.07) 0.25 (0.4) −0.063 (0.8) −0.4 (0.2) −0.44 (0.1)
E12BM E17BM D01BM D07BM D21BM Body_Weight BM_Weight BMR IMF_BM IMF_Weight Liver_Weight BMGR
1
0.5
0
−0.5
−1
c
d

Figure 3.5 WGCNA results of the first phase BM data set. a. Module cluster dendrogram. b. Module-trait relationship. The upper number in the block is the module's corresponding correlation with the bottom trait and the lower number is the p-value of the corresponding correlation. c. Filtered co-expression network of the grey60 module. d. Filtered co-expression network of the light-yellow module. The yellow nodes in c and d are the identified hub genes, involved TFs and TF co-factors.

Twenty-nine co-expression modules were detected for the second BM phase (D07-D180, Figure 3.6a), and the module-trait-stage relationships are shown in Figure 6b. The significant positive modules for BM IMF percentage are the brown module (P = 0.003), the dark-green module (P = 0.005), and the dark-grey module (P = 0.008). The network of eigengenes for this data set is shown in Figure S3.8b. After within module edges filtering, no hub genes remained in the dark-green module. The genes *GIPC2* and *UBE2V2* in the brown module and *LOC112532140* and *ENSGALG00000053632* in the dark green module were detected as hub genes related to the high breast muscle IMF percentage in phase2. The involved TFs and genes enriched pathways are shown in Table 3.1.



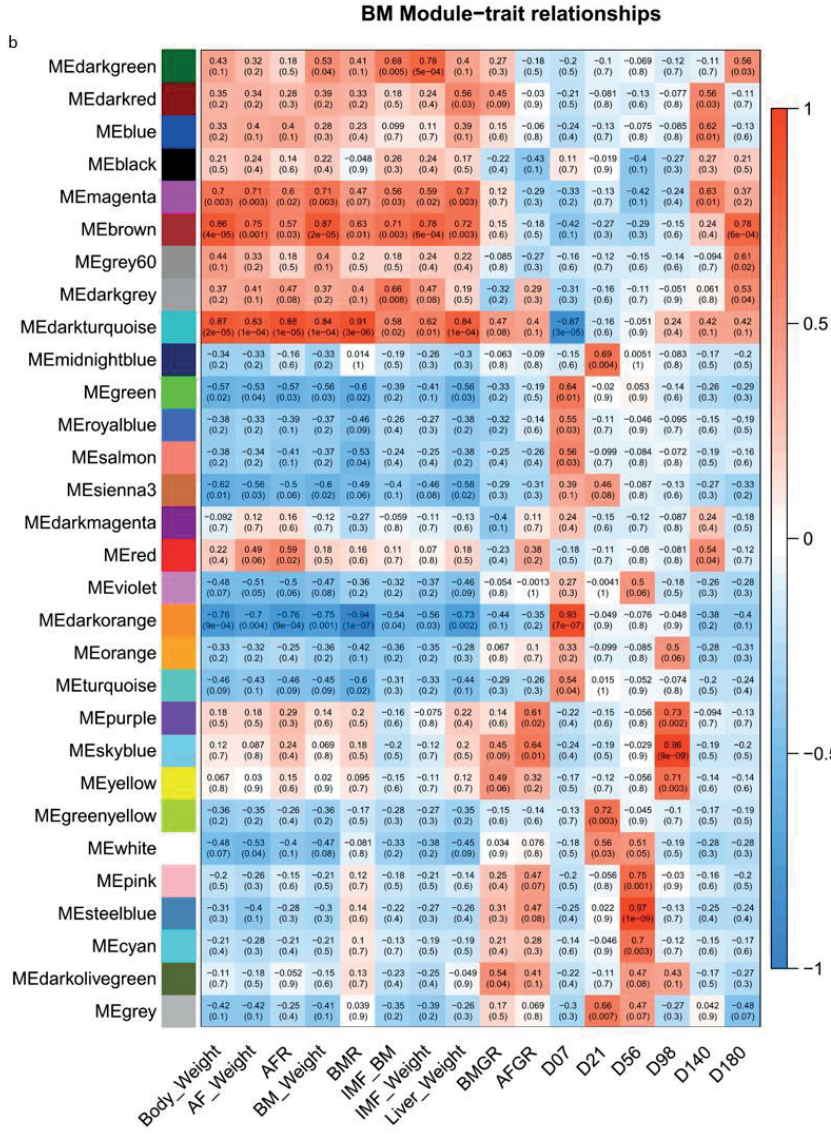


Figure 3.6 WGCNA results of the second phase BM data set. a. Module cluster dendrogram. b. The module-trait relationship.

Table 3.1 Hub genes, TFs, and enriched pathways identified in each phenotype related module of breast muscle.

Data set	Module	Hub genes	TFs*	KEGG pathway**
BM Phase1	grey60	*ATG9B, COX6A1, ENSGALG00000041996, ENSGALG00000053368*	*HOXB1*	Carbon metabolism, leucine and isoleucine degradation, TCA cycle, fatty acid metabolism, fatty acid elongation
	light yellow	*LOC107050564, ENSGALG00000048510*	*MYCN*	-
BM Phase 2	brown	*GIPC2, MLF1, UBE2V2, ENSGALG00000015443, ENSGALG00000030350*	*MYOD1*	Protein processing in endoplasmic reticulum
	dark grey	*ENSGALG00000050515, JCHAIN, LOC112532140*	*EGR2, EGR3, IRF5, KLF4, L3MBTL1, LITAF, PLEK, SMAD7B*	Cytokine-cytokine receptor interaction, neuro active ligand-receptor interaction, C-type lectin receptor signaling pathway

* Detailed TFs and TF co-factors and pathways shows in Table S11.

**No pathway enriched in the light-yellow module in the BM phase1 dataset. The detailed pathway shows in Table S12.

Detection of Hub Genes and Transcription (co) Factors Related to Abdominal Fat Weight

For the abdominal fat genes expression data in the WGCNA, the module cluster dendrogram (Figure S3.9a) and the module-trait relationship (Figure S3.9b) identified 24 modules. The network of eigengenes for AF expression data is shown in Figure S3.9c. The turquoise module (P = 1e-06) is significantly positively related to AF weight. After filtering the within interaction edges, the genes such as *EIF3J*, *EPM2A*, *SH3BGRL*, *ENSGALG00000047756*, and *CHMP4B* were identified as hub genes. The turquoise module membership vs gene significance on AFW is shown in Figure S3.9d.

The yellow module, with 1099 genes, is significantly negatively related to AFW (*P* = 0.009). The yellow module membership vs gene significance for AFW is shown in Figure S3.9e. Several hub genes were identified, such as *MASTL*, *CENPE*, *MCM4*, *CREB3L1*, and *PKB*. The genes of the solute carrier family, such as *SLC1A3*, *SLC2A10*, *SLC3A1*, *SLC6A8*, and *SLC7A2* were also involved in the co-expression network. The complete list of involved TFs and TF co-factors derived from the WGCNA is provided in Table S3.10 and S3.11. The TFs involved and genes enriched in the pathways are shown in Table 3.2.

The TF *L3MBTL1* and the TF co-factors *TNIP1*, *HAT1*, and *BEND6,* all involved in the IMF positively related module of the second phase of the BM data set, overlapped with the AFW negatively related module of the AF data set.

Table 3.2 Hub genes, TFs, and enriched pathways identified in AFW related module of AF.

Module	Correlation	Hub genes	TFs*	KEGG pathway**
turquoise	Positively correlated to AFW	*EIF3J, EPM2A, CALM1*	*PPARD, CEBPB, CEBPD, EGR2, EGR4, FOS, MXD4, RREB1, TWIST2, XBP1*	Endocytosis, Lysosome, mRNA surveillance pathway, and Proteasome
yellow	Negatively correlated to AFW	*MASTL, CENPE, MCM4, CREB3L1*	*FOXM1, CREB3L1, MYBL2, GATA4, TULP3, RFX5, L3MBTL1*	Cell cycle, Protein processing in endoplasmic reticulum

3.4 Discussion

Time course RNA sequencing has been widely used to study cellular differentiation (Ma et al., 2018), tissues development (Cardoso-Moreira et al., 2019), and aging (Baumgart et al., 2016). We provide new insight on the transcriptome changes in chicken between different development stages of breast muscle and abdominal fat. Although the Jingxing-Huang IMF-up selected chicken population is an inbred line, the experimental chickens have similar genetic background. The 3 biological replicates cannot cover all the population transcriptome changes. The PCA result indicates that the transcriptome changes of the late developmental stages are smaller than those of the early stages. The number of stage specific expressed

genes of early stages are higher than those of later stages. The IMF ratio of BM displays a peak around the day of hatch followed by an increase from D98 to D180 which agrees with an earlier study focusing on the stages around the day of hatch (Liu et al., 2017). On the day of hatch, the IMF content is high (13.6%), then the IMF content remains in a low level until D140. This might be caused by the fact that the breast muscle grows a lot during this period. The genes specifically expressed in BM during stage D01 are enriched for fatty acid elongation and biosynthesis of unsaturated fatty acids, indicating that these pathways may contribute to the high IMF phenotype in BM. After hatch, most of the breast muscle IMF deposition starts around D98, while the AF deposition starts from D07 and accelerates from D56.

Genes with significant changes in expression at different developmental stages were considered as DDGs. We used DDGs to reflect the transcriptome of BM and AF changes during development in cell type abundance, gene regulation, and the proportion of cells undergoing division (Pantalacci and Sémon, 2015). The number of DDGs in BM (8545) and AF (6824) is somewhat higher than the average number of DDGs detected in Red jungle fowl (RJF) in the brain, cerebrum, heart, kidney, liver, ovary and testis (Cardoso-Moreira et al., 2019). There are several possible explanations for the observed differences e.g., the tissues, sequencing technology, time points, and species. The number of DDGs in BM is higher than that in AF, showing that BM tissue has more genes that change in expression during the developmental period assessed in this study than AF tissue. This could be due to the higher number of cell types in BM compared to AF. Furthermore, the number of TFs and TF co-factors decreased during development, which is consistent with earlier research in other animals (Bolger et al., 2014). Thus, as the development process proceeds, the required number of TFs becomes lower.

WGCNA is a powerful tool for identifying genes that are associated with the phenotypes under study (Langfelder and Horvath, 2008). WGCNA can also be used to identify tissue- or stage- specifically expressed gene modules (Gao et al., 2018; Ma et al., 2018). To investigate the expressed genes in different developmental stages, we performed the WGCNA for BM and AF tissues as consensus data set. Thirty-four modules were detected, indicating that the gene expression varies a lot between the developmental stages. There are 2157 genes in the salmon, green, and light-yellow modules with more than 39% genes with unknown function. The genes in the turquoise module are expressed higher in AF compared to BM and are enriched in the *PPAR* signaling and fatty acid metabolism pathways, which are known to be involved in fat deposition.

We initially performed the WGCNA on the complete BM data set. We found that IMF positively related modules are like the positively related D01 stage modules, representing the IMF deposition during the early period. This may be an issue to cause potential false-positive errors in the early stages of adipocyte differentiation as well as false-negative errors in the late stage for fat deposition. Therefore, two separate WGCNAs were performed for the hyperplasia and the hypertrophy phases of the BM data. The hyperplasia phase is covered by E12 to D21 and the hypertrophy phase includes the stages D07 to D180. For the abdominal fat dataset, there are only samples from D07 to D180. Hence, no separation of phases was necessary. There are different ways to identify the hub genes for WGCNA results. E.g. the WGCNA package provides a function for hub gene detection (Langfelder and Horvath, 2008), and the genes with kME > 0.95 can also be considered as the hub genes (Gao et al., 2018). However, in this study, there are several large modules, which may be driven by several hub genes. Then, we used the expression level of genes, the weight of connected genes, and the connectivity number of genes as the criteria for the detection of hub genes.

In the WGCNA results of the first phase BM data set, the hub gene *ENSGALG00000041996*, a lncRNA, may regulate *CD36* and *ACADL*. The *ENSGALG00000041996* connected genes in breast muscle phase 1 dataset enriched pathways are shown in Figure S10. This would make sense because fatty acids are transported via fatty acid binding protein (FABP), fatty acid translocase (FAT/CD36), and cell membrane diffusion (Stump et al., 1993). The acyl-CoA dehydrogenase long chain gene (*ACADL*) plays a role in catalyzing the first step of mitochondrial fatty acid beta-oxidation (Indo et al., 1992). Both the *CD36* and *ACADL* belong to the *PPAR* signaling pathway. While, the *PPAR* signaling pathway can also induce and activate the expression of *aP2* and *PEPCK*, which are specifically expressed in fat tissue (Tontonoz et al., 1995). This suggests that the unannotated gene *ENSGALG00000041996* may play a key role in fat deposition during early developmental stages of BM. In the second phase BM data set WGCNA, the brown module is significantly positively correlated with the BM IMF. *MYOD1* is the only TF in the brown module. *MYOD1* is also connected with the hub gene *GIPC2* (Figure 6c). This may indicate that the TF *MYOD1* regulates genes in the brown module through *GIPC2* thereby affecting muscle development and IMF deposition. The remaining hub genes, e.g., *ENSGAL0000005538*, *LOC107050546*, *ENGSGALG00000015443*, and *ENSGALG00000030350* may play so far unidentified roles in IMF deposition. The phenotypes BMW and IMF give a very similar pattern. However, we cannot distinguish genes that associate with muscle development or

IMF deposition. The identified hub genes and involved TFs in the BM data set can be used as candidate genes for high IMF chicken selection.

To identify the hub genes involved in the AF deposition, we performed WGCNA in the AF data set. There is no obvious relationship between the hub genes and the lipid metabolism, e.g., *EIF3J* (eukaryotic translation initiation factor 3 subunit J), *EPM2A* (epilepsy, progressive myoclonus type 2A), and *CALM1* (calmodulin 1). However, the TFs *PPARD* and *CEBPB* are involved in the turquoise module, which is positively correlated to AFW. The gene *PPARD* is expressed in multiple tissues in adult mouse (Higashiyama et al., 2007) and regulates glucose metabolism and insulin sensitivity (Chih-Hao et al., 2006). *CEBPB* seems to be synergistic in promoting lipogenesis in AF of cockerels (Resnyk et al., 2017). The TF *CREBP3L1* (cAMP responsive element binding protein 3 like 1) is involved in the AF weight negatively correlated yellow module. The gene *CREBP* can reduce lipogenesis as well as glycolysis in mice (Katsumi et al., 2004). The genes which belong to the solute carrier family, were not present in the center of the yellow module. This may indicate that the genes of the solute carrier family play some role downstream of lipid metabolism.

In chicken breeding there is a desire of producing chicken with high IMF and low AF. From our study we found some promising candidate genes. Especially the TF *L3MBTL1* and the TF co-factors *TNIP1*, *HAT1*, and *BEND6*, which were identified as significantly positively related to the high IMF and significantly negatively related to the low AFW could be relevant biomarkers for chicken breeding. RT-q-PCR of the four TFs/TFCFs in breast muscle of the Jingxing-Huang chickens and Cobb chickens in a large number of individuals, showed that the expression of *TNIP1* and *HAT1* in the high IMF group is significantly higher than in the low IMF group (Li et al., 2020).

3.5 Conclusion

In this chapter, the transcriptome dynamics of chicken BM and AF in different developmental stages is described. This is an important resource for studying IMF and AF in chicken. Developmental dynamics genes and involved TFs were identified, which may play key roles in tissue development. In addition, we identified several regulatory hub genes that potentially can be used in breeding to improve IMF content in muscle and meanwhile reduce the AFW.

3.6 Declarations

Abbreviations

IMF: intramuscular fat, BM: breast muscle, AF: abdominal fat, DDGs: developmental dynamic genes, MSCs: mesenchymal stem cells, *PPARγ*: peroxisome proliferator activated receptor γ, *CREBP3L1*: cAMP responsive element binding protein 3 like 1, DEGs: differentially expressed genes, *ACADL*: acyl-CoA dehydrogenase long chain, RJF: red jungle fowl, *MYOD1*: myogenic differentiation 1, TF: transcription factors, TOM: topological overlap matrix, WGCNA: weight gene co-expression network analysis, BMGR: breast muscle growth rate, AFGR: abdominal fat growth rate, AFW: abdominal fat weight.

Ethics approvement

The animal experiments were conducted at the chicken farm in the IAS of CAAS, and the animal experiments have been approved by the animal ethics committee of Institute of Animal Sciences, Chinese Academy of Agricultural Sciences in 2016 (approved code IASCAAS-AE-02).

Data accessibility

The datasets analyzed for this study can be found in the Genome Sequence Archive [53] in BIG Data Center [54], Beijing Institute of Genomics (BIG), Chinese Academy of Sciences, under project number PRJCA001192 and accession number CRA001334 that are publicly accessible at http://bigd.big.ac.cn/gsa.

Author contributions

J.W., G.Z., R.L., R.P.M.A.C, M.A.M.G, and S.X. designed the study; S.X., L.L., and M.Z. performed the animal experiments and samples collection; L.L., S.X., and X.Y. tested the phenotypes; S.X. performed data analysis; R.L., R.P.M.A.C., O.M., M.A.M.G, and S.X. discussed the results; S.X. wrote the manuscript; O.M., R.L., R.P.M.A.C, and M.A.M.G. provide valuable suggestion and comments to improve the manuscript with contributions from all other authors.

Funding

The study was supported by the National Key Research and Development Program of China (No. 2018YFD0500401), the Modern Agro-industry Technology Research System (CARS-41) and Agricultural Science and Technology Innovation Program (ASTIPIAS04; ASTIP-IAS-TS-15).

Acknowledgements

We thank Ms. Zhou Wu, Mr. Langqing Liu, and Dr. Evert van Schothorst (Wageningen University & Research), Dr. Qinghe Li, Mr. Jie Wang, Mr. Fan Ying (Institute of Animal Sciences of Chinese Academy of Agricultural Sciences), Prof. Xiaoxiang Hu and Prof. Jianfeng Liu (China Agricultural University) for technical assistance and good comments. We thank the Berry Genomics company for RNA-seq library construction and sequencing. We thank Dr. Mark Camara for providing the professional native English revision.

Supplemental files

The supplemental files for this article can be found online.

Table S3.1 Phenotypes of RNA-sequenced chickens.
Table S3.2 RNA-sequenced samples number in different developmental stages of BM and AF.
Table S3.3 Alignment summary for each sequenced sample.
Table S3.4. Gene expression raw counts matrix.
Table S3.5 BM stage specifically expressed gene list.
Table S3.6 AF stage specifically expressed gene list.
Table S3.7 The number of DEGs between each developmental stage in BM and AF. The number of up- and down-regulated gene number between adjacent developmental stages in each tissue from D07 to D180. The comparison performed as former stage vs latter stage.
Table S3.8 BM and AF expressed TFs of DDGs number in different stages.
Table S3.9 Developmental dynamic gene list of BM and AF.
Table S3.10 Transcription factors list of DDGs and interested modules.
Table S3.11 Transcription factors co-factors of DDGs and interested modules.
Table S3.12 Details of KEGG enriched pathways of interested module of WGCNA results.
Figure S3.1 The overlap between BM and AF expressed genes.
Figure S3.2 a. The PCA plot of BM samples, b. the PCA plot of AF samples.
Figure S3.3 a. The stage specifically expressed gene number of BM and AF. b. BM stage specifically expressed genes enriched KEGG pathways. c. AF stage specifically expressed genes enriched KEGG pathways.
Figure S3.4 The BM IMF and AFW fitted curve.
Figure S3.5 The TFs and TF cofactors of BM and AF DDGs enriched pathways.
Figure S3.6 a. BM samples outliers. b. The network of eigengene for the WGCNA result of the first phase of the BM data set. c. The network of eigengenes for the WGCNA result of the second phase of the BM data set.

Figure S3.7 a. WGCNA result of modules dendrogram. b. module traits relationship. c. AFW eigengene dendrogram and adjacency heatmap. d. Turquoise module membership vs gene significance on AFW. e. Yellow module membership vs gene significance on AFW.

Figure S3.8 The KEGG pathway of *ENSGALG00000041996* related genes in the grey60 module.

References

Allen, C. (1976). Cellularity of adipose tissue in meat animals. *Federation proceedings* 35(11)**,** 2302-2307.

Ana, C., María José, N., Alberto, F., and Manuel, T. (2006). maSigPro: a method to identify significantly differential expression profiles in time-course microarray experiments. *Bioinformatics* 22(9)**,** 1096-1102.

Andrews, S. (2010). *Babraham bioinformatics-FastQC a quality control tool for high throughput sequence data* [Online].

Baumgart, M., Priebe, S., Groth, M., Hartmann, N., and Cellerino, A. (2016). Longitudinal RNA-Seq Analysis of Vertebrate Aging Identifies Mitochondrial Complex I as a Small-Molecule-Sensitive Modifier of Lifespan. *Cell Systems* 2(2)**,** 122-132.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)* 57(1)**,** 289-300.

Bolger, A.M., Lohse, M., and Usadel, B.J.B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. 30(15)**,** 2114-2120.

Cardoso-Moreira, M., Halbert, J., Valloton, D., Velten, B., Chen, C., Shao, Y., et al. (2019). Gene expression across mammalian organ development. *Nature* 571(7766)**,** 505-509.

Carlson, M. (2019). Genome wide annotation for chicken org.Gg.eg.db R version 3.8.2.

Chen, X., Xizeng, M., Jiaju, H., Yang, D., Jianmin, W., Shan, D., et al. (2011). KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. 39(Web Server issue)**,** 316-322.

Chih-Hao, L., Peter, O., Andrea, H., Isaac, M., Ling-Wa, C., Olefsky, J.M., et al. (2006). PPARdelta regulates glucose metabolism and insulin sensitivity. *Proceedings of the National Academy of Sciences of the United States of America* 103(9)**,** 3444-3449.

De Mendiburu, F. (2014). Agricolae: statistical procedures for agricultural research. *R package version* 1(1).

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: ultrafast universal RNA-seq aligner. 29(1)**,** 15-21.

Farmer, S.R.J.C.m. (2006). Transcriptional control of adipocyte formation. 4(4)**,** 263-273.

Gao, H., Wang, Y., Li, W., Gu, Y., Lai, Y., Bi, Y., et al. (2018). Transcriptomic comparison reveals genetic variation potentially underlying seed developmental evolution of soybeans. *Journal of experimental botany* 69(21)**,** 5089-5104.

Griffiths, R., and Korn, R.M. (1997). A CHD1 gene is Z chromosome linked in the chicken Gallus domesticus. *Gene* 197(1–2)**,** 225-229.

Guo, L., Cui, H., Zhao, G., Liu, R., Li, Q., Zheng, M., et al. (2018). Intramuscular preadipocytes impede differentiation and promote lipid deposition of muscle satellite cells in chickens. *BMC genomics* 19(1)**,** 838.

Higashiyama, H., Billin, A.N., Okamoto, Y., Kinoshita, M., and Asano, S. (2007). Expression profiling of Peroxisome proliferator-activated receptor-delta (PPAR-delta) in mouse tissues using tissue microarray. *Histochemistry & Cell Biology* 127(5)**,** 485-494.

Hocquette, J.F., Gondret, F., Baéza, E., Médale, F., Jurie, C., and Pethick, D.W. (2010). Intramuscular fat content in meat-producing animals: development, genetic and nutritional control, and identification of putative markers. *Animal An International Journal of Animal Bioscience* 4(2)**,** 303-319.

Hu, H., Miao, Y.-R., Jia, L.-H., Yu, Q.-Y., Zhang, Q., and Guo, A.-Y. (2018). AnimalTFDB 3.0: a comprehensive resource for annotation and prediction of animal transcription factors. *Nucleic acids research* 47(D1)**,** D33-D38.

Indo, Y., Yang-Feng, T., Glassberg, R., and Tanaka, K. (1992). Molecular cloning and nucleotide sequence of cDNAs encoding human long-chain acyl-CoA dehydrogenase and assignment of the location of its gene (ACADL) to chromosome 2 ☆. *Genomics* 12(3)**,** 609-620.

Jiang, M., Fan, W.L., Xing, S.Y., Wang, J., Li, P., Liu, R.R., et al. (2017). Effects of balanced selection for intramuscular fat and abdominal fat percentage and estimates of genetic parameters. *Poult Sci* 96(2)**,** 282-287.

Kassambara, A. (2017). ggpubr:"ggplot2" based publication ready plots. *R package version 0.1* 6.

Katsumi, I., Bruick, R.K., Guosheng, L., Horton, J.D., and Kosaku, U. (2004). Deficiency of carbohydrate response element-binding protein (ChREBP) reduces lipogenesis as well as glycolysis. *Proc Natl Acad Sci U S A* 101(19)**,** 7281-7286.

Kim, J., Park, K., Lee, E., Jang, W., Seo, J., Shin, S., et al. (2014). Suppression of PPARγ through MKRN1-mediated ubiquitination and degradation prevents adipocyte differentiation. 21(4)**,** 594.

Langfelder, P., and Horvath, S.J.B.b. (2008). WGCNA: an R package for weighted correlation network analysis. 9(1)**,** 559.

Li, J., Xing, S., Wang, X., Li, Q., Zhao, G., Zhang, Y., et al. (2020). Screening of Candidaate Genes Related to Intramuscular Fat Deposition in Chickens *China Animal Husbandry & Veterinary Meidcine* 47(6)**,** 1826-1836.

Liu, R., Wang, H., Liu, J., Wang, J., Zheng, M., Tan, X., et al. (2017). Uncovering the embryonic development-related proteome and metabolome signatures in breast muscle and intramuscular fat of fast-and slow-growing chickens. *BMC Genomics* 18(1)**,** 816.

Love, M.I., Huber, W., and Anders, S.J.G.b. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. 15(12)**,** 550.

Ma, B., Lee, T.-L., Hu, B., Li, J., Li, X., Zhao, X., et al. (2018). Molecular characteristics of early-stage female germ cells revealed by RNA sequencing of low-input cells and analysis of genome-wide DNA methylation. *DNA Research* 26(2)**,** 105-117.

Na, W., Wu, Y.Y., Gong, P.F., Wu, C.Y., Cheng, B.H., Wang, Y.X., et al. (2018). Embryonic transcriptome and proteome analyses on hepatic lipid metabolism in chickens divergently selected for abdominal fat content. *Bmc Genomics* 19(1)**,** 384.

Nueda, M.J., Tarazona, S., and Conesa, A.J.B. (2014). Next maSigPro: updating maSigPro bioconductor package for RNA-seq time series. 30(18)**,** 2598.

Ono, M., Aratani, Y., Kitagawa, I., and Kitagawa, Y.J.E.c.r. (1990). Ascorbic acid phosphate stimulates type IV collagen synthesis and accelerates adipose conversion of 3T3-L1 cells. 187(2)**,** 309-314.

Pantalacci, S., and Sémon, M. (2015). Transcriptomics of developing embryos and organs: A raising tool for evo-devo. *J Exp Zool B Mol Dev Evol* 324(4)**,** 363-371.

Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.C., Mendell, J.T., and Salzberg, S.L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol* 33(3)**,** 290-295.

Pethick, D.W., Harper, G.S., and Oddy, V.H. (2001). Growth, development and nutritional manipulation of marbling in cattle. *Crc for Cattle & Beef Quality*.

Pittenger, M.F., Mackay, A.M., Beck, S.C., Jaiswal, R.K., Douglas, R., Mosca, J.D., et al. (1999). Multilineage Potential of Adult Human Mesenchymal Stem Cells. *Science* 284(5411)**,** 143-147.

Reiter, S.S., Halsey, C.H., Stronach, B.M., Bartosh, J.L., Owsley, W.F., Bergen, W.G.J.C.B., et al. (2007). Lipid metabolism related gene-expression profiling in liver, skeletal muscle and adipose tissue in crossbred Duroc and Pietrain pigs. 2(3)**,** 200-206.

Resnyk, C.W., Carré, W., Wang, X., Porter, T.E., Simon, J., Bihan-Duval, E.L., et al. (2017). Transcriptional analysis of abdominal fat in chickens divergently selected on bodyweight at two ages reveals novel mechanisms controlling adiposity: validating visceral adipose tissue as a dynamic endocrine and metabolic organ. *Bmc Genomics* 18(1)**,** 626.

Ross, S.E., Hemati, N., Longo, K.A., Bennett, C.N., Lucas, P.C., Erickson, R.L., et al. (2000). Inhibition of adipogenesis by Wnt signaling. *Science* 289(5481)**,** 950-953.

Schmitz, J., Evers, N., Awazawa, M., Nicholls, H.T., Brönneke, H.S., Dietrich, A., et al. (2016). Obesogenic memory can confer long-term increases in adipose tissue but not liver inflammation and insulin resistance after weight loss. 5(5)**,** 328-339.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., et al. (2003). Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. 13(11)**,** 2498.

Soxhlet, F. (1879). Die gewichtsaiialytische Bestimmung des Milchfettes; von.

Stump, D.D., Zhou, S.L., and Berk, P.D. (1993). Comparison of plasma membrane FABP and mitochondrial isoform of aspartate aminotransferase from rat liver. *Am J Physiol* 265(5 Pt 1)**,** G894.

Symonds, M.E. (2012). *Adipose Tissue Biology.*

Tang, Q.Q., and Lane, M.D. (1999). Activation and centromeric localization of CCAAT/enhancer-binding proteins during the mitotic clonal expansion of adipocyte differentiation. *Genes Dev* 13(17)**,** 2231-2241.

Tontonoz, P., Hu, E., Devine, J., Beale, E.G., and Spiegelman, B.M. (1995). PPAR gamma 2 regulates adipose expression of the phosphoenolpyruvate carboxykinase gene. *Mol.cell.biol* 15(1)**,** 351-357.

Tontonoz, P., Hu, E., and Spiegelman, B.M. (1994). Stimulation of adipogenesis in fibroblasts by PPAR gamma 2, a lipid-activated transcription factor. 79(7)**,** 1147-1156.

Yu, G., Wang, L.G., Han, Y., and He, Q.Y.J.O.-a.J.o.I.B. (2012). clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *Omics-a Journal of Integrative Biology* 16(5)**,** 284-287.

Zhuo, Z., Lamont, S.J., Lee, W.R., and Abasht, B. (2015). RNA-seq analysis of abdominal fat reveals differences between modern commercial broiler chickens with high and low feed efficiencies. *PloS one* 10(8)**,** e0135810.

4

Time course transcriptomic study reveals the gene regulation during liver development and the correlation with abdominal fat weight in chicken

Siyuan Xing[1, 2], Ranran Liu[1], Guiping Zhao[1], Martien A. M. Groenen[2], Ole Madsen[2], Lu Liu[1], Maiqing Zheng[1], Qiao Wang[1], Zhou Wu[2], Richard P. M. A. Crooijmans[2], Jie Wen[1],

[1]State Key Laboratory of Animal Nutrition, Key Laboratory of Animal (Poultry) Genetics Breeding and Reproduction, Ministry of Agriculture and Rural Affairs, Institute of Animal Sciences, Chinese Academy of Agricultural Sciences, 100193 Beijing, China

[2]Animal Breeding and Genomics, Wageningen University & Research, 6708PB Wageningen, The Netherlands

Frontiers in Genetics (2021) :1593

Abstract

Background: Liver is the central metabolic organ of animals. In chicken, knowledge on the relationship between gene expression in the liver and fat deposition during development is still limited. A time-course transcriptomic study from the embryonic (day 12) to the egg-producing period (day 180 after hatch) was performed to profile slow-growing meat type chicken liver gene expression and to investigate its correlation with abdominal fat deposition.

Results: The transcriptome profiles showed a separation of the different developmental stages. In total, 13,096 genes were ubiquitously expressed at all the tested developmental stages. The analysis of differentially expressed genes between adjacent developmental stages showed that the biosynthesis of unsaturated fatty acids pathway was enriched from Day 21 to Day 140 after hatch. The correlation between liver gene expression and the trait abdominal fat weight (AFW) was analysed by weighted gene co-expression network analysis. The genes *MFGE8*, *HHLA1*, *CKAP2*, and *ACSBG2* were identified as hub genes in AFW positively correlated modules, which suggested important roles of these genes in the lipid metabolism in chicken liver.

Conclusion: Our results provided a resource of developmental transcriptome profiles in chicken liver and suggested that the gene *ACSBG2* among others can be used as a candidate gene for selecting low abdominal fat weight chickens.

Key words: chicken, liver, time course, transcriptome, abdominal fat

4.1 Introduction

In chicken, transcriptional analysis on fat deposition were mainly aimed to understand the mechanisms of fat deposition in different depots *e.g.*, visceral fat (Resnyk et al., 2015) and intramuscular fat (Liu et al., 2019). Liver is the central metabolic organ and it provides many essential endocrine and exocrine functions including fat synthesis (Zorn, 2008). In mammals, the liver provides around 70% of *de novo* synthesized fatty acid, whereas, in chickens, the liver provides around 90% of *de novo* synthesized fatty acid (O'Hea and Leveille, 1969). Our knowledge of chicken liver gene expression during different developmental stages and how it regulates the lipid deposition is still limited.

In the last decades, new insights into the liver-visceral adipose axis grew rapidly (Cornide-Petronio et al., 2019). An important factor influencing the liver lipid flux is the adipose tissue (Azzu et al., 2020). In the fasted state, lipolysis is the main contributor to the increased fatty acid (FA) turnover rate, whereas, in the fed state, both the disability of adipose tissue to take up lipids and the failure of insulin to suppress lipolysis can increase the FA turnover rates (Azzu et al., 2020). In addition, the liver can also facilitate lipolysis of adipose tissue (Mandard et al., 2006).

Liver is mainly composed of hepatocytes and biliary epithelial cells, which differentiated from the endoderm (Zorn, 2008). Gene expression differed in chicken liver at 5 embryonic stages between chickens divergently selected for abdominal fat (AF) and showed enrichment of the fatty acid metabolism and the peroxisome proliferator-activated receptor (PPAR) signaling pathways (Na et al., 2018). Chen *et al.* (2019) studied the transcriptome in the chicken liver after 2 weeks of high-fat feeding and found the differentially expressed genes (DEGs) mainly enriched in the cell cycle and PPAR signaling pathways (Chen et al., 2019).

However, gene expression during liver development and its relationship with adipose deposition in chicken has been investigated only to a limited extent. Here we present the results of gene expression in chicken liver at different embryonic (from embryo day 12) until egg-production (up to day 180 after hatch) stages and find potential regulator genes for abdominal fat deposition by combining the time course, co-expression, and genomic analyses.

4.2 Material & Methods

Chicken phenotypes and samples collection

In this study we used chickens from the 16th generation of an intramuscular fat-up selected line. The selection line originates from the Jingxing-Huang chicken, which is a yellow-feathered slow-growing dwarfism line. The genetic background of these experimental chickens has been described in our previous studies (Zhao et al., 2006; Jiang et al., 2017). At each of the following 9 developmental stages: E12 (embryonic day 12), E17, and D1 (day 1 after hatching), D7, D21, D56, D98, D140, and D180, liver samples of 3 female chickens were collected. Chickens were reared with *ad libitum* access to feed and water. The chickens were slaughtered without fasting to avoid activation of the fasting-feeding cycle of gene expression regulation. The growth curve of body weight (BW), liver weight (LW), and abdominal fat weight (AFW) were fitted by the Logistic model using the Origin software (version 2018). Abdominal fat percentage (AFP) was calculated by AFW/BW. Abdominal fat growth rate (AFGR) was calculated as (AFW_{later} - AFW_{former})/Time. The lower margin of the liver was collected for RNA isolation and RNA-sequencing of the 27 female chickens. Oil Red O stain was used on the liver sections for developmental stages E12, E17, D1, and D21.

RNA sequencing and data quality control

The QIAGEN RNeasy Kit was used to isolate total RNA, and genomic DNA was removed by the TIANGEN DNase KIT. The RNA concentration was assessed by Nanophotometer. RNA integrity number (RIN) was assessed by Nanodrop analysis. The RIN value of all total RNA sample was larger than 7 and RNA library construction was performed by Berry Genomics (Beijing, China). Poly-A enriched RNA samples were isolated by Dynabeads mRNA DIRECT Kit (Invitrogen). The non-stranded specific RNA libraries were sequenced on the Illumina Hiseq2500 (paired end at125 bp). After trimming of the sequencing adaptors and low quality reads (N > 10% in a read) by Trimmomatic (version 0.39) with default parameters (Bolger et al., 2014), the quality of the sequencing data was assessed by FastQC (version 0.11.5) (Andrews, 2010).

Transcriptome profiling and differentially expressed genes (DEGs) detection

The transcriptome data were aligned to the chicken reference genome (GRCg6a) and annotation file (Gallus.gallus.GRCg6a.95.gtf) by STAR (version 2.5.3) (Dobin et al., 2013) and assembled with Stringtie (version 1.3.3b) (Pertea et al., 2015). Raw gene counts were performed by using a Python script provided by Stringtie with

parameter l=125 (Pertea et al., 2016). Gene expression level normalization was performed by DESeq2 (version 1.22.2) (Love et al., 2014) in R (version 3.6.1), based on the experimental design as Family + Stage. The normalized gene expression data were used for all downstream analyses. The list of transcription factors (TF) and transcription co-factors were extracted from AnimalTFDB (v.3.0) (Hu et al., 2018). Transcriptome PCA plots were performed by sample distances calculated by *rlog* function of DESeq2 (Love et al., 2014). Genes with the expression fold change (FC) > 1.5 or FC < 0.67 and with the Benjamini-Hochberg method (Benjamini and Hochberg, 1995) adjusted-p < 0.05 were considered as DEGs.

Development dynamics genes (DDGs) identification and genes expression pattern

The DDGs were identified by the maSigPro package (version 1.46.0) (Ana et al., 2006; Nueda et al., 2014) applying a negative binomial model for the expression distribution and using the Benjiamini and Hochberg procedure to adjust the false discovery rate. Significant genes were selected by the forward method with r^2 > 0.7. Gene expression pattern analysis followed the design of a single series time course. The parameters used for gene pattern clustering: counts = TRUE, min.obs=10, and rsq=0.6. The k.mclust = TRUE was used to calculate the optimal clusters number.

Weighted gene co-expression network analysis (WGCNA)

A weighted genes co-expression analysis was performed by the WGCNA (version 1.41) package (Langfelder and Horvath, 2008) with minor modified parameters. By using the step-by-step topology overlap matrix (TOM) construction (soft-threshold=8) and setting the minModuleSize to 30 for module detection. The co-expression network of a given module were filtered by edges with weight < 0.15. Finally, genes with edge number less than or equal to 10 were filtered out. Gene co-expression networks were performed by the Cytoscape software (version 3.6.0) (Shannon et al., 2003) with the edges provided by the WGCNA 'exportNetworkToCytoscape' function. The genes with the highest Σweight were identified as hub genes. The time course impulse expression of *ACSBG2* were performed by ImpulseDE (version 3.11) (Sander et al., 2017).

QTL information

The chicken AFW related QTL regions were collected from chicken QTL data base (release 41) (Hu et al., 2019). The UCSC tool lift-over (http://genome.ucsc.edu/cgi-bin/hgLiftOver) was used to transform the chicken AFW QTL regions from galGal-5.0 to GRRCg6a. The candidate gene detected in this study (*ACSBG2*) and AFW

related QTL region were visualized by Gviz (version 1.34.1) and related packages in R.

Pathway analysis

KEGG enrichment analysis resulting in dot plots and bar plots was performed using clusterProfiler package (version 3.11.1) with p.adj < 0.05 as significant (Yu et al., 2012) and org.Gg.eg.db package (version 3.8.2) (Carlson, 2019).

Statistical analysis

Student's t-test was performed by R (version 3.6.0) using the functions shapiro.test for normality test and bartlett.test to test for homogeneity of the variance. The data sets which did not fit the normal distribution were compared by the rank-sum test. The Least-Significant-Difference-test were performed by the agricolae package (version 1.3.1)(Mendiburu, 2017). All significances were set at $P < 0.05$.

4.3 Results

The phenotype of liver, BW, and AFW during development

BW, LW, and AFW were obtained from the 27 chickens used for RNA-seq. The fitted curves for body weight and liver weight are shown in Figure 4.1 A and B. The BW and liver weight fitted a logistic regression model. Compared with the stages from D07 to D56, the AF deposition increased from D56 to D140, with more than 8 grams per week (Figure 4.1C). The lipid analysis of the early developmental stages using the Oil Red O-stained section of the liver (Figure 4.1D to 4.1G) show that there is no obvious lipid staining at E12. From E17 to the first day after hatching, the lipid started to deposit in the liver. While at D21, very limited lipid is seen in the hepatocyte, adipocytes have appeared. The phenotypic data is provided in supplementary Table S4.1.

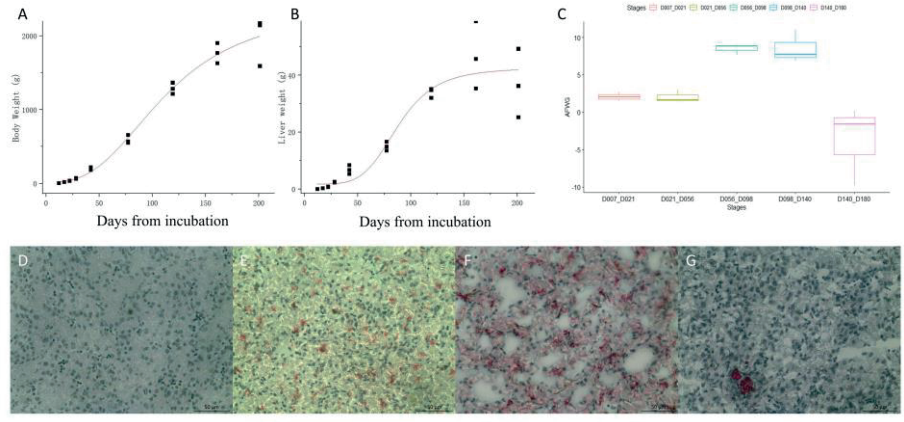


Figure 4.1 The phenotypes of BW and liver weight during development. A. The RNA-sequenced chicken body weight fitted curve. B. The sequenced chicken liver weight fitted curve. C. to F. The Oil-Red-O-stained sections of the liver in E12, E17, D1, and D21, respectively. Red color are the stained lipid or adipocytes.

Transcriptome profiling during liver development

To observe the difference at the transcriptomic level for the 9 different developmental stages, 27 RNA-seq libraries were constructed and sequenced. On average 30.61±4.39 million trimmed reads were obtained per library and the mean uniquely mapped alignment ratio was 93.78% (Table S4.2). In total, 20,496 out of 24,356 genes (Gallus.gallus.GRCg6a.95.gtf) were detected as expressed (read count >1) across the 9 developmental stages of which 13,096 were ubiquitously expressed at all stages. The total number of genes expressed at each stage ranged from 15,373 in D180 to 17,222 in E12 (Table 4.1).

To explore whether the gene expression profiles correlated with the developmental stages, we performed a PCA plot (Figure 4.2). In general, the 3 samples of each stage clustered together. The resolution at D7 to D98 is less distinct compared to the other time points. The gene expression level of each sample is provided in Table S4.3.

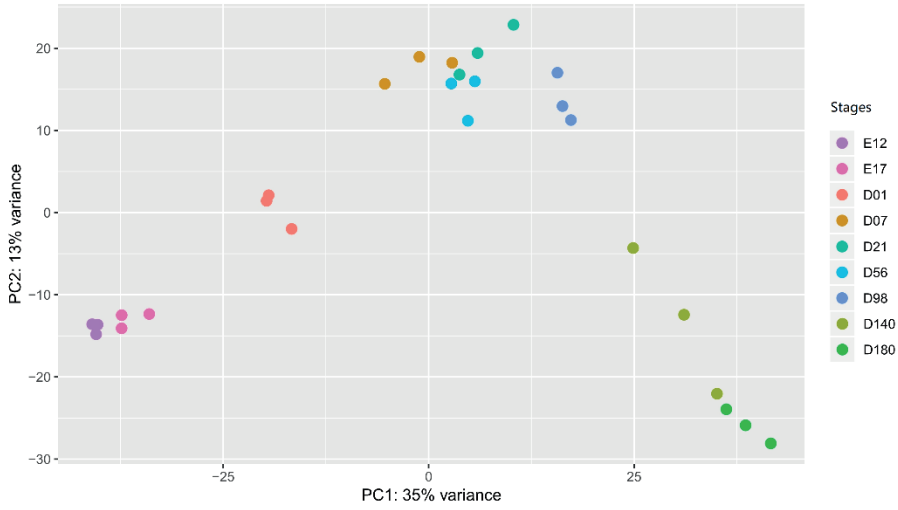


Figure 4.2 The PCA plot of liver samples in different developmental stages. Different colors represent different development stages.

The number of stage-specific expressed genes varied from 63 to 343 (Table 4.1 and Table S4.5). Stage specific expressed genes enriched KEGG pathways are shown in Figure S4.1. The number of switched on/off genes are presented in Table 4.1. Switched on genes varied from 796 (D1) to 1,505 (D98), whereas for the switched off genes, it varied from 942 (D7) to 1,989 (D140). The KEGG enrichment results for the switched on and off genes are presented in Figure S4.2 and S4.3, respectively. At stage D98, the switched-on genes were enriched in alpha-linolenic acid metabolism, arachidonic acid metabolism, and ether lipid metabolism pathways. The switched-off genes at D140 are enriched in ether lipid metabolism, alpha-linolenic acid metabolism, and melanogenesis. Most of the TFs for the stage-specific expressed genes and switched-on/off genes belong to the Homeobox family (Table S4.5 to S4.7).

Table 4.1 The number of expressed, specific expressed, and switched-on/off genes.

Developmental stage	Expressed genes	Specific expressed genes*	Switched-on genes*	Switched-off genes*
E12	17,222	343	**	**
E17	17,131	311	1,169	1,261
D1	15,963	108	796	1,963
D7	16,429	118	1,408	942
D21	16,225	279	1,219	1,423
D56	16,399	136	1,406	1,232
D98	16,764	324	1,505	1,140
D140	15,615	87	840	1,989
D180	15,373	63	943	1,185

* The list of specifically expressed gene and switched-on/off genes and the involved TFs and TFCFs in each stage are presented in Table S4.5 to S4.7.
** There are no switched-on/off genes at the E12.

Differentially expressed genes (DEGs) between adjacent stages of liver development

The numbers of DEGs between the developmental stages varied from 45 (D21 vs D56) to 4411 (E12 vs D180), and detailed information of the genes is shown in Table S4.8. The number of DEGs between adjacent stages are highest for D01 vs D07 and D56 vs D98 (Figure 4.3A). KEGG pathway enrichment analysis results of the detected DEGs are presented in Figure 4.3B. There are no KEGG enriched pathways detected for the E12 vs E17 nor between the D7 vs D21 comparisons. The biosynthesis of unsaturated fatty acids pathway was enriched for the comparison of D21 vs D56, D56 vs D98, and D98 vs D140. Fatty acid degradation and PPAR signaling pathways were enriched for the comparison of D56 vs D98 and D98 vs D140.

Figure 4.3 The number of DEGs and KEGG enrichment. A. The number of DEGs between adjacent developmental stages. The red and blue stand for up- and down-regulated gene numbers, respectively. B. The KEGG enriched pathways of DEGs between different developmental stages. The color represents adjusted p-value. The spot size represents the enriched gene ratio.

Developmentally dynamic genes (DDGs) and gene expression patterns in liver

To study gene expression changes during liver development, genes with significant temporal changes (DDGs) were clustered. Across all the tested stages, 8974 genes were identified as DDGs, including 340 TFs and 261 TF co-factors (Table S4.9). The top 20 enriched GO terms for these TFs of DDGs are shown in Figure 4.4. Processes like DNA templated transcription, RNA biosynthetic process, and regulation of transcription were enriched.

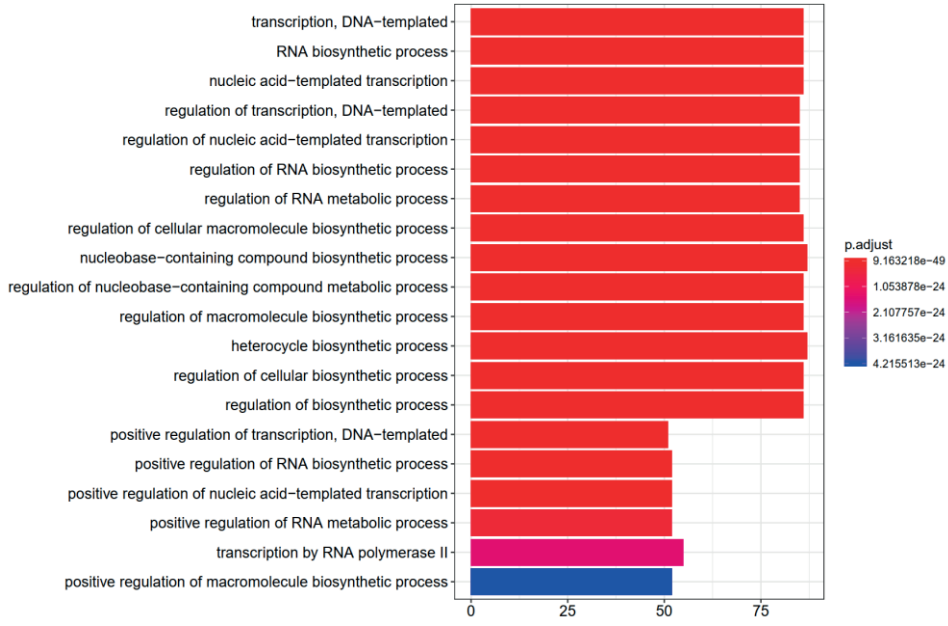


Figure 4.4 Top 20 enriched GO terms for TFs of DDGs.

Hub genes in liver development and lipid deposition

To detect the hub-genes involved in liver growth, lipid metabolism, and abdominal fat deposition, WGCNA were performed (Figure S4.4A and S4.4B). In total, 30 co-expression modules were obtained (Figure S4.4C). The module-trait relationship is presented in Figure 4.5. The red module, involving 847 genes, is significantly positively correlated with liver weight (p=2e-04), AFW (p=7e-06) and AFP (p=0.001). A group of 323 genes within the royal-blue module is significantly positively correlated to AFW (p=0.003) and AFP (p=0.003). The turquoise module, with in total 4,852 genes, is significantly negatively correlated to liver weight (p=2e-04), AFW (p=0.002), and AFP (p=5e-05). The turquoise module is also positively correlated to the embryonic period E12 (p=1e-04) and E17 (p=0.006). The dark-orange module which include 216 genes significantly correlated to liver weight (p=0.001), AFW (p=0.01), AFP (p=0.002), and AFGR (p=0.01).

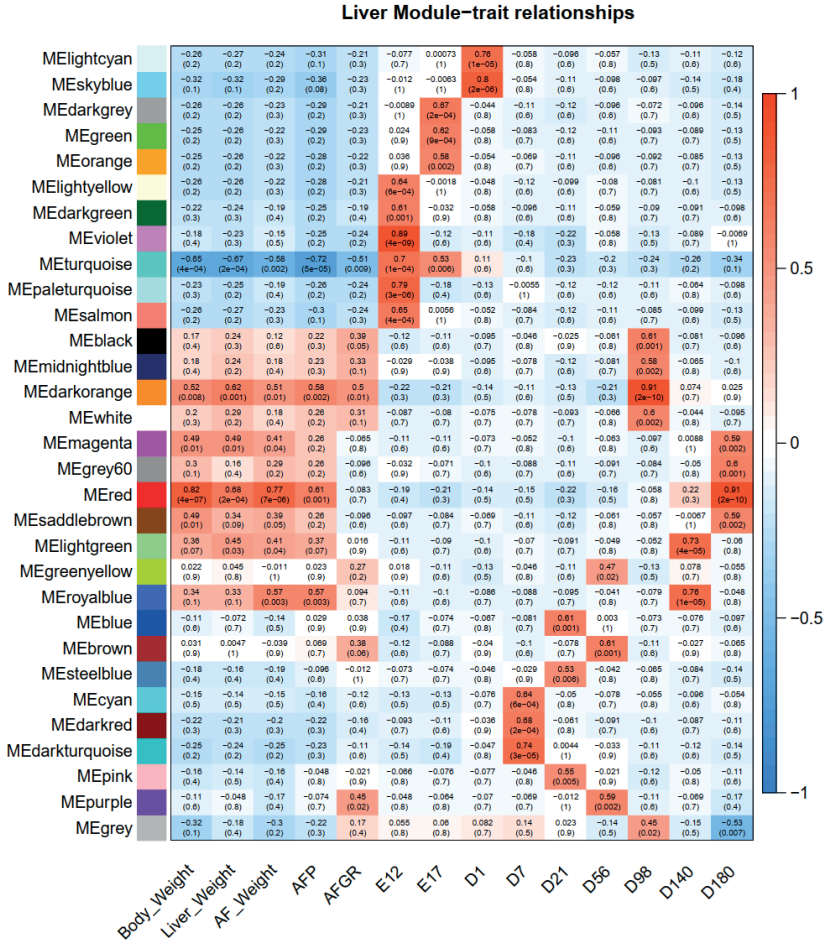


Figure 4.5 The liver expressed genes WGCNA module-trait relationship. The color represents the correlation between module and trait. The upper-number and lower-number are the corresponding correlations and *p* value.

The *MFGE8* (milk fat globule-EGF 8 protein), *HHLA1* (HERV-H LTR-associating 1), *CKAP2* (cytoskeleton associated protein 2), and *ACSBG2* (Acyl-CoA synthetase bubblegum family member 2) genes were identified as hub genes in these 4 modules, respectively. Pathway enrichment analysis of the genes in the four modules are presented in Table 4.2. The protein processing in endoplasmic reticulum, Protein export, Cell cycle, DNA replication, and Fanconi anemia pathways are enriched. The co-expression network with the detected hub genes is shown in Figure S4.5 A to D. We found chicken QTL 24370 (chr28:1,761,021-1761061) and QTL 24371 (chr28:1,751,075-1,751,115) (D'Andre et al., 2013)

associated to chicken AFW that overlap with the *ACSBG2* gene (chr28:1,746,737-1,763,012). The expression pattern of *ACSBG2* may indicate that the *ACSBG2* was impulse regulated at D98 stage (Figure S4.6).

Table 4.2 The identified hub genes, and enriched pathways.

Module	Positively correlated to	Hub genes	Significantly enriched KEGG pathway
Red	AFW	*MFGE8*	Protein processing in endoplasmic reticulum and Protein export
Royal blue	AFW	*HHLA1*	No pathway enriched
Turquoise	AFW	*CKAP2*	Cell cycle, DNA replication, and Fanconi anemia pathway
Dark orange	AFW and AFGR	*ACSBG2*	No pathway enriched

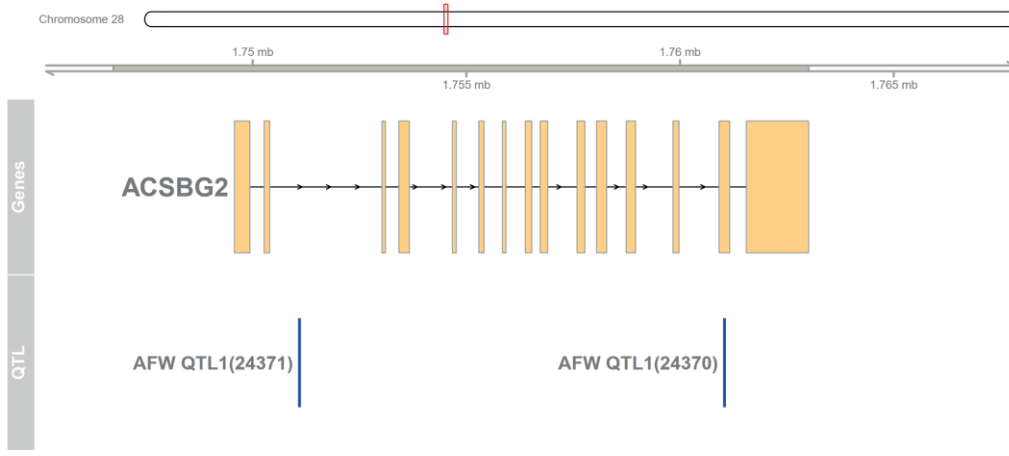


Figure 4.6 The region of *ACSBG2* covers 2 chicken AFW QTLs at chromosome 28.

4.4 Discussion

We profiled the transcriptome during liver development from early embryonic stages to the egg-producing period in chicken. The relatively large differences of the transcriptomes during development shows a large variation of the number of expressed genes in liver. Nevertheless, the changes in the transcriptome from D7 to D56 are relatively small. The KEGG enrichment result of temporal and ubiquitously expressed genes suggest that different biological processes are active during development. For example, D56 specific expressed genes are enriched for several fatty acid related pathways, *e.g.,* linoleic acid metabolism. Furthermore, D98 switched on genes also are enriched in linoleic acid metabolism pathway but switched off again in D140. This may be caused by changes in the feed composition.

But it also may indicate that during D56 to D98, the linoleic acid metabolism becomes more important in liver lipid metabolism.

We identified 8,974 DDGs in this study, a number that is higher to the DDGs detected in the Red jungle fowl using both genders (Cardoso-Moreira et al., 2019). The reasons of this difference can be the different library construction methods used (single-end vs paired-end), different reference genomes used (galGal4 vs GRCg6a), different genders, as well as different time points that were used.

WGCNA is a powerful tool to analyze complex transcriptome data sets (Langfelder and Horvath, 2008). From the modules correlating with these traits, we identified *MFGE8*, *HHLA1*, *CKPA2*, and *ACSBG2* as hub genes. The first three hub genes, however, do not seem to be directly related to lipid metabolism. Therefore, more research is needed to determine the roles of these hub genes in liver fat metabolism in chicken, but it should be noted that some of these hub genes may be false positives caused by the positive correlation between liver weight, AFW, and BW.

Interestingly, although the corresponding correlation to AF percentage of the dark-orange module is lower than the red and the royal-blue modules, when we focused on the AF growth rate, the hub gene of dark-orange module, *ACSBG2,* shows more potential relationship with lipid metabolism. *ACSBG2* encodes the acyl-CoA synthetase bubblegum family member 2 protein and can catalyze hexadecenoic acid to the hexadecanoyl-CoA. It is involved in the fatty acid metabolism, fatty acid degradation, adipocytokine signaling, PPAR signaling, and thermogenesis pathways. It also plays an important role downstream of *FAT/CD36*, which can acquire the free fatty acid from the outside of the cell. The *ACSBG2* gene was first cloned and identified in human in 2006 and shown to be specifically expressed in testis and the brainstem (Pei et al., 2006). In chicken, it is expressed in many tissues like *e.g.* brain, cerebellum, heart, kidney and ovary, and highly expressed in testis (Cardoso-Moreira et al., 2019) and liver (Figure S4.6). The expression of *ACSBG2* was tested in the liver and hypothalamus tissues of fast- and slow-growing chicken by using the Affymetrix Genechip ® Chicken Genome array (D'Andre et al., 2013). They found 2 SNPs in the gene to be significantly associated with abdominal fat weight. This suggests that *ACSBG2* might be a good candidate gene for selection for slim chicken. A study comparing the transcriptomes of the intestine and muscle, between divergent feed efficient broilers, showed that *ACSBG2* influences the muscular lipid utilization and was among the highest expressed genes in muscle (Henry et al., 2018). We however did not find high expression of *ACSBG2* in breast

muscle at the nine different developmental stages of the same 27 slow growing chicken used in this study (Xing et al., 2020). This difference between the studies may be caused by using different chicken breeds that differ in their growth rate.

4.5 Conclusion

In the current study, we provided a useful gene expression data resource for chicken liver during development. The results suggest that the candidate gene *ACSBG2* among potentially other detected genes can further contribute to chicken breeding with the aim of low abdominal fat weight.

4.6 Declarations

Sequencing Data Accession Numbers

The raw sequence data reported in this paper have been deposited in the Genome Sequence Archive (Wang et al., 2017) in BIG Data Center (Members, 2018), Beijing Institute of Genomics (BIG), Chinese Academy of Sciences, under accession numbers CRA001334 that are publicly accessible at http://bigd.big.ac.cn/gsa .

Author contributions

J.W., G.P., M.A.M.G., R.L., R.P.M.A.C., and S.X. designed the study; S.X., L.L., M.Z., and Q.W. performed the animal experiments and samples collection; S.X. and L.L. tested the phenotypes; S.X. performed data analysis; S.X., R.L., R.P.M.A.C., M.A.M.G., O.M., and Z.W. discussed the results; S.X. wrote the manuscript; R.L., R.P.M.A.C., O.M., M.A.M.G., Z.W., and J.W. provided valuable suggestion and comments to improve the manuscript with contributions from all other authors.

Acknowledgments

We thank Dr. Qinghe Li and Dr. Huanxian Cui for project comments, Mr. Jie Wang and Mr. Fan Ying, Mr. Bo Zhu, Mr. Xiaodong Tan, and Dr. Fei Wang for assistance in samples collection (Institute of Animal Sciences of Chinese Academy of Agricultural Sciences). We thank Dr. Langqing Liu (Munich University) for assistance in data analysis. We thank the Berry Genomics company for RNA-seq library construction and sequencing.

Funding

The study was supported by the National Natural Science Foundation of China (31772591), the Project of State Key Laboratory of Animal Nutrition (2004DA125184G1607), the China Agriculture Research System of MOF and MARA (CARS-41), Agricultural Science and Technology Innovation Program (ASTIP-IAS04), and China Scholarship Council (201903250098).

Ethics statement

The animal experiments were conducted at the chicken farm in the IAS of CAAS, and the animal experiments had been approved by the animal ethics committee of Institute of Animal Sciences, Chinese Academy of Agricultural Sciences in 2016 (approved code IASCAAS-AE-02).

Abbreviations

AFGR: abdominal fat growth rate, AFP: abdominal fat percentage, AFW: abdominal fat weight, DDGs: developmental dynamic genes, DEGs: differentially expressed genes, QTL: quantitative trait loci, RIN: RNA integrity number, TOM: topological overlap matrix, WGCNA: weighted gene co-expression network analysis.

Supplemental files can be found at https://osf.io/8uzmh/ .

Table S4.1 Phenotypes of RNA-sequenced chickens.
Table S4.2 Alignment ratio for each sequenced library.
Table S4.3 Gene expression raw counts matrix of each library.
Table S4.4 The ubiquitously expressed gene list.
Table S4.5 The stage-specific expressed gene list from E12 to D180.
Table S4.6 The switch-on gene list from E17 to D180.
Table S4.7 The switch-off gene list from E17 to D180.
Table S4.8 The number of DEGs between each developmental stage.
Table S4.9 The DDGs and involved TFs list.
Figure S4.1 Liver stage-specific expressed genes enriched pathways.
Figure S4.2 Liver switched-on genes enriched pathways.
Figure S4.3 Liver switched-off genes enriched pathways.
Figure S4.4 WGCNA processing figures. A. Outliers. B. Optimal soft-threshold selection. C. Cluster dendrogram.
Figure S4.5 Module networks. Yellow nodes stand for hub genes.
Figure S4.6 ACSBG2 impulse expressed in D98 stage.

References

Ana, C., María José, N., Alberto, F., and Manuel, T. (2006). maSigPro: a method to identify significantly differential expression profiles in time-course microarray experiments. *Bioinformatics* 22(9), 1096-1102.

Andrews, S. (2010). "FastQC: a quality control tool for high throughput sequence data".).

Azzu, V., Vacca, M., Virtue, S., Allison, M., and Vidal-Puig, A. (2020). Adipose Tissue-Liver Cross Talk in the Control of Whole-Body Metabolism: Implications in Nonalcoholic Fatty Liver Disease. *Gastroenterology* 158(7), 1899-1912.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)* 57(1), 289-300.

Bolger, A.M., Lohse, M., and Usadel, B.J.B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. 30(15), 2114-2120.

Cardoso-Moreira, M., Halbert, J., Valloton, D., Velten, B., Chen, C., Shao, Y., et al. (2019). Gene expression across mammalian organ development. *Nature* 571(7766), 505-509.

Carlson, M. (2019). Genome wide annotation for chicken org.Gg.eg.db R version 3.8.2.

Chen, J.-H., Ren, X.-Y., Li, L.-M., Lu, S.-Y., and Luo, W. (2019). The cell cycle pathway regulates chicken abdominal fat deposition as revealed by transcriptome sequencing. *Hereditas (Beijing)* 41(10), 962-973.

Cornide-Petronio, M.E., Jimenez-Castro, M.B., Gracia-Sancho, J., and Peralta, C. (2019). New Insights into the Liver-Visceral Adipose Axis During Hepatic Resection and Liver Transplantation. *Cells* 8(9).

D'Andre, H.C., Paul, W., Shen, X., Jia, X., Zhang, R., Sun, L., et al. (2013). Identification and characterization of genes that control fat deposition in chickens. *Journal of animal science and biotechnology* 4(1), 1-16.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: ultrafast universal RNA-seq aligner. 29(1), 15-21.

Henry, Reyer, Barbara, Metzler-Zebeli, Nares, Trakooljul, et al. (2018). Transcriptional shifts account for divergent resource allocation in feed efficient broiler chickens. *Scientific Reports* 8, 12903.

Hu, H., Miao, Y.-R., Jia, L.-H., Yu, Q.-Y., Zhang, Q., and Guo, A.-Y. (2018). AnimalTFDB 3.0: a comprehensive resource for annotation and prediction of animal transcription factors. *Nucleic acids research* 47(D1), D33-D38.

Hu, Z.-L., Park, C.A., and Reecy, J.M. (2019). Building a livestock genetic and genomic information knowledgebase through integrative developments of Animal QTLdb and CorrDB. *Nucleic acids research* 47(D1), D701-D710.

Jiang, M., Fan, W.L., Xing, S.Y., Wang, J., Li, P., Liu, R.R., et al. (2017). Effects of balanced selection for intramuscular fat and abdominal fat percentage and estimates of genetic parameters. *Poult Sci* 96(2), 282-287.

Langfelder, P., and Horvath, S.J.B.b. (2008). WGCNA: an R package for weighted correlation network analysis. 9(1), 559.

Liu, L., Cui, H., Xing, S., Zhao, G., and Wen, J. (2019). Effect of Divergent Selection for Intramuscular Fat Content on Muscle Lipid Metabolism in Chickens. *Animals* 10(1).

Love, M.I., Huber, W., and Anders, S.J.G.b. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. 15(12)**,** 550.

Mandard, S., Zandbergen, F., van Straten, E., Wahli, W., Kuipers, F., Muller, M., et al. (2006). The fasting-induced adipose factor/angiopoietin-like protein 4 is physically associated with lipoproteins and governs plasma lipid levels and adiposity. *J Biol Chem* 281(2)**,** 934-944.

Members, B.D.C.J.N.A.R. (2018). Database Resources of the BIG Data Center in 2018. 46(Database issue)**,** D14-D20.

Mendiburu, F.D. (2017). agricolae: Statistical Procedures for Agricultural Research. *Journal of the American Statal Association* 80(390).

Na, W., Wu, Y.Y., Gong, P.F., Wu, C.Y., Cheng, B.H., Wang, Y.X., et al. (2018). Embryonic transcriptome and proteome analyses on hepatic lipid metabolism in chickens divergently selected for abdominal fat content. *Bmc Genomics* 19(1)**,** 384.

Nueda, M.J., Tarazona, S., and Conesa, A.J.B. (2014). Next maSigPro: updating maSigPro bioconductor package for RNA-seq time series. 30(18)**,** 2598.

O'Hea, E.K., and Leveille, G.A. (1969). Lipid biosynthesis and transport in the domestic chick (Gallus domesticus). *Comp Biochem Physiol* 30(1)**,** 149-159.

Pei, Z.T., Jia, Z.Z., and Watkins, P.A. (2006). The second member of the human and murine "bubblegum" family is a testis- and brainstem-specific acyl-CoA synthetase. *Journal of Biological Chemistry* 281(10)**,** 6632-6641.

Pertea, M., Kim, D., Pertea, G.M., Leek, J.T., and Salzberg, S.L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nature Protocols* 11(9)**,** 1650-1667.

Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.C., Mendell, J.T., and Salzberg, S.L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol* 33(3)**,** 290-295.

Resnyk, C.W., Chuming, C., Hongzhan, H., H., W.C., Jean, S., Elisabeth, L.B.-D., et al. (2015). RNA-Seq Analysis of Abdominal Fat in Genetically Fat and Lean Chickens Highlights a Divergence in Expression of Genes Controlling Adiposity, Hemostasis, and Lipid Metabolism. *Plos One* 10(10)**,** e0139549-.

Sander, J., Schultze, J.L., and Yosef, N. (2017). ImpulseDE: detection of differentially expressed genes in time series data using impulse models. *Bioinformatics* 33(5)**,** 757-759.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., et al. (2003). Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. 13(11)**,** 2498.

Wang, Y., Song, F., Zhu, J., Zhang, S., Yang, Y., Chen, T., et al. (2017). GSA: Genome Sequence Archive*. 15(1)**,** 14-18.

Xing, S., Liu, R., Zhao, G., Liu, L., Groenen, M.A., Madsen, O., et al. (2020). RNA-Seq analysis reveals hub genes involved in chicken intramuscular fat and abdominal fat deposition during development. *Frontiers in genetics* 11.

Yu, G., Wang, L.G., Han, Y., and He, Q.Y.J.O.-a.J.o.I.B. (2012). clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *Omics-a Journal of Integrative Biology* 16(5)**,** 284-287.

Zhao, G., Wen, J., Chen, J., Zheng, M., and Li, X. (2006). Selection response and estimation of the genetic parameters for intramuscular fat in a quality chicken line. *J Acta veterinaria et zootechnica Sinica* 37(9)**,** 870-873.

Zorn, A.M. (2008). Liver development. *Stembook*. Publisher:Cambridge (MA): Harvard Stem Cell Institute, Oct 31

5

Exploring gene expression correlated to fat deposition in chicken breast muscle and abdominal fat

Siyuan Xing[1, 2], Ranran Liu[1], Guiping Zhao[1], Martien A. M. Groenen[2], Jinghui Li[1], Xicai Wang[1], Qiao Wang[1], Maiqing Zheng[1], Richard P. M. A. Crooijmans[2], Jie Wen[1]

[1] 1. State Key Laboratory of Animal Nutrition, Key Laboratory of Animal (Poultry) Genetics Breeding and Reproduction, Ministry of Agriculture and Rural Affairs, Institute of Animal Sciences, Chinese Academy of Agricultural Sciences, Beijing, China

2. Animal Breeding and Genomics, Wageningen University & Research, Wageningen, The Netherlands

In preparation

Abstract

Background: High intramuscular fat (IMF) in chicken breast muscle (BM) improves meat flavor, while abdominal fat (AF) is a waste after slaughter. Although differences in expression levels of specific genes have been shown to affect fatness phenotype, the knowledge on the correlations between gene expression and multiple fatness traits is still limited. To explore gene expression regulation on IMF and AF deposition in chicken, we used BM and AF transcriptome data of high-IMF-low-AF and low-IMF-high-AF chickens from a slow growing dwarf line at marketing time.

Results: In total, 13,344 and 19,582 genes were expressed in BM and AF, respectively. Weighted gene co-expression analysis (WGCNA) resulted in 2 modules that are significantly positively correlated to triglyceride content of BM ($p < 0.05$) and significantly negatively correlated to AF percentage ($p < 0.05$). We identified *ACSM3* and *CYP2AB1* as hub genes in the BM and AF datasets, respectively.

Conclusion: Our results provide new insight on gene expression both influencing breast IMF and AF deposition in chicken. *ACSM3* and *CYP2AB1* can be considered as candidate genes for increased breast muscle IMF content and decreased AF weight selection in chickens.

Key words: chicken, RNA-seq, co-expression, intramuscular fat, abdominal fat

5.1 Introduction

Chicken meat is a very important resource in animal protein intake of humans (Xin et al., 2021). Intramuscular fat (IMF) content in chicken breast muscle (BM) influences the meat quality e.g. flavour and juiciness (Chizzolini et al., 1999), while abdominal fat (AF) is a waste after slaughter, decreasing the economic efficiency (Fouad and El-Senousey, 2014). In the chicken industry, increased IMF and decreased AF is a long-term preferred breeding goal. Lipid droplets, composed of triglycerides (TGs), hold most of the volume of adipocytes (Wang et al., 2017a). TGs have been used as an indicator for breast muscle IMF content (Chen et al., 2017; Liu et al., 2019).

Differences in gene expression have been shown to influence the fat deposition (Resnyk et al., 2013). By comparing transcriptome data from an abdominal weight selection fat-line and a lean-line chicken, a large number of differentially expressed genes (DEGs) involved in the synthesis, metabolism, and transport of lipids were identified in AF (Resnyk et al., 2013). Similar results were obtained by comparing AF DEGs between high- and low-growth chickens (Resnyk et al., 2017). A comparison of DEGs in the major pectoralis between high and low TGs chickens, showed that the peroxisome proliferator activated receptor (PPAR) signaling, fatty acid degradation, and the steroid biosynthesis pathways may regulate the differential lipid deposition in BM (Liu et al., 2019). The DEGs between IMF- and AF-derived preadipocytes and DEGs enriched pathways were also analyzed and PPAR signaling and Extracellular Matrix-receptor interaction pathways were shown to be enriched in preadipocytes differentiation (Zhang et al., 2019). The different DEGs of AF and BM in these studies are the result of differences in genetic background, sample numbers, transcriptome library construction technologies, and data analytical methods. To limit the influence of variation in these external factors, a consensus co-expression network analysis, commonly used to analyze multiple data set e.g., abdominal fat RNA-seq datasets of different chicken breeds, was performed in multiple broiler lines to identify genes correlated with abdominal fat deposition in chicken (Yuan and Lu, 2021). These previous studies focused on the gene expression influencing AF and IMF deposition independently. Knowledge on gene expression correlated to both BM IMF and AF deposition, however, is lacking.

Weighted co-expression analysis is a powerful tool to study the correlation between gene expression and phenotypes (Langfelder and Horvath, 2008). In the present study therefore, we aimed to use this approach to identify genes influencing fat deposition in both breast muscle and abdominal fat in chicken.

5.2 Material & Methods

Animals and phenotyping

The chickens used in this study are from the Chinese locally developed Jingxing-Huang chicken, a dwarf yellow feathered meat-type chicken line. Chickens were reared with *ad libitum* access to feed and water. There were 240 male chickens slaughtered at 98 days after hatching and phenotyped for body weight, slaughtering weight (SW), breast muscle weight (BMW), and abdominal fat weight (AFW). Part of the major pectoralis of breast muscle and abdominal fat was snap-frozen and stored in liquid nitrogen. For breast muscle IMF content, we measured the triglyceride (TG) value of the major pectoralis, which represents the main content of IMF. The remaining major pectoralis samples were stored at -20 ℃ for analysis with the Triglyceride Assay Kit (Nanjing Jiancheng Bioengineering Institute) with 3 measurement replicates for each sample. TG content is shown as mmol/L.

Library construction and sequencing data quality control

To maximize the variation both in IMF and AF, we selected 8 individuals from the high-IMF-low-AF group and 10 individuals from the low-IMF-high-AF group. The BM and AF samples from these 18 birds with significant differences (p<0.01) in both BM TG content and AFW, were used for transcriptome library construction. Total RNA of the BM and AF samples were isolated using the RNAprep Pure Tissue Kit (TIANGEN). The RNA concentration was measured by Nanophotometer. Oligo(dT) enriched transcriptome library construction was commercially performed by IGENECODE, Beijing, China. The un-stranded specific RNA libraries were sequenced on the NovaSeq 6000 (2×150 bp). Sequencing adaptors were trimmed by Trimmomatic (version 0.39) (Bolger et al., 2014). We assessed the quality of the obtained sequencing data by FastQC (version 0.11.5) (Andrews, 2010). Venn diagram (http://bioinformatics.psb.ugent.be/webtools/Venn/) was used to calculate the intersection between the analyses and to draw Venn diagrams of the output.

Transcriptome data alignment, assembly, and gene expression profiling

An index was built using the chicken reference genome (GRCg6a) and annotation file (Gallus.gallus.GRCg6a.102.gtf) for read alignment. The transcriptome data were aligned to the chicken reference genome using STAR (version 2.5.3) (Dobin et al., 2013) and assembled with Stringtie (version 1.3.3b) (Pertea et al., 2015). Raw gene counts were performed by using the Python script provided by Stringtie with parameter l=150 (Pertea et al., 2016). Genes with average raw counts > 1 in BM and AF were considered as expressed genes. PCA plots of gene expression in BM and AF

were performed by sample distances calculated by *vst* function of DESeq2 (version 1.30.1) (Love et al., 2014).

Gene co-expression network construction

To study the correlation between phenotypes and gene expression, weighted gene co-expression analysis was performed using the WGCNA package (version 1.70-3) (Langfelder and Horvath, 2008). Raw gene counts of all expressed genes in BM and AF were used in this analysis. Genes with low expression level were filtered by the *gsg* function with default parameter. To construct the scale free co-expression network, a suitable soft threshold (power) is needed. By raising the soft threshold, the weighted gene co-expression network weakens the low correlations and emphasizes high correlations (Langfelder and Horvath, 2008). When the R^2 of the scale free topology model fit reaches 0.9, the soft threshold (power) was used for downstream topology overlap matrixes (TOM) construction. The TOMs were constructed with minModuleSize = 50. Similar modules were merged with height=0.25. The remaining parameters were default.

Module-traits relationship analysis

The Benjiamini-Hochberg method was used on module-trait association quantification. Gene co-expression networks were performed by the Cytoscape software (version 3.6.0) (Shannon et al., 2003) using the 'exportNetworkToCytoscape' function with the edges provided by WGCNA. As previously described (Xing et al., 2020), the gene with the highest Σweight was considered as hub gene in each module. The biological function networks of hub genes were provided by STRING database (version 11) (Szklarczyk et al., 2019).

5.3 Results

Chicken phenotypes

The body weight (BW), breast muscle weight (BMW), abdominal Fat weight (AFW), and triglyceride (TG) phenotypes obtained from the 18 chickens used for RNA-seq were used to compare these phenotypes between the two groups of Jingxing Huang chicken. Significant differences (p<0.01) are observed for TG and AFW between the high-IMF-low-AF and low-IMF-high-AF groups (Figure 5.1). And there is no significant difference between groups in BW (p=0.13, Figure S5.1) and BMW (p=0.76, Figure S5.2). The detailed phenotypes for individuals used in this study, are provided in Table S5.1.

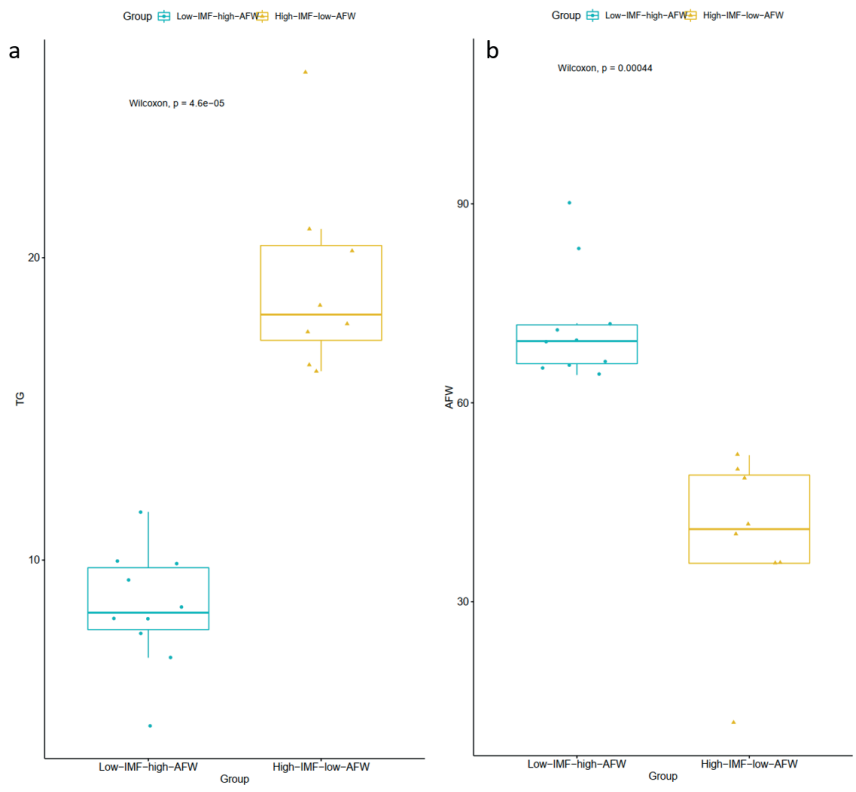


Figure 5.1 The TG content and AFW of chickens. The TG content (a) and AFW (b) are significantly different between the two groups (p=4.6e-5 and p=4.4e-5 respectively).

Transcriptome profiling and gene expression

To analyze BM and AF gene expression in the two groups, 36 RNA-seq libraries were constructed and sequenced from the 18 individuals (Table S5.2). On average, 21.27 million trimmed reads were obtained for each library. After alignment, the mean uniquely mapped alignment ratio was 84.45% (Table S5.3). We detected 13,344 genes expressed in the BM dataset and 19,582 expressed in the AF dataset. There are 13,041 genes expressed in both tissues (Figure S5.3). The PCA of the expression data (Figure 5.2), shows a strong separation of BM and AF in the first dimension explaining 88% of the variation. The BM samples cluster together while the AF samples are divided in the second dimension. However, the second dimension only explains 2% of the variation. Also, no clustering is observed for the samples derived from the two different phenotype groups (Figures S5.4 and S5.5). The gene expression level of each sample is provided in Table S4.

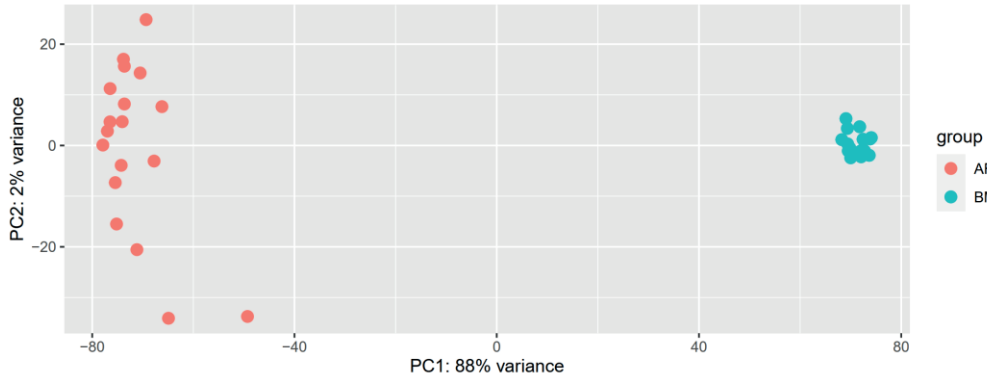


Figure 5.2 PCA plot of expressed genes per sample showing strong separation of BM and AF samples. The PCA plots of BM and AF showing the selection of the individual samples are provided in Figures S5.4 and S5.5, respectively.

BM and AF gene co-expression network construction

The soft threshold for the BM (soft threshold=8) and AF (soft threshold=10) data sets were selected to construct the scale-free network for the two datasets (Figure 5.3). Then we constructed gene co-expression networks for the BM and AF datasets, respectively. Constructing the topology overlap matrixes (TOMs), for the gene co-expression data sets, resulted in 25 gene modules for BM (Figure S5.5) and 26 gene modules for AF (Figure S5.6), respectively.

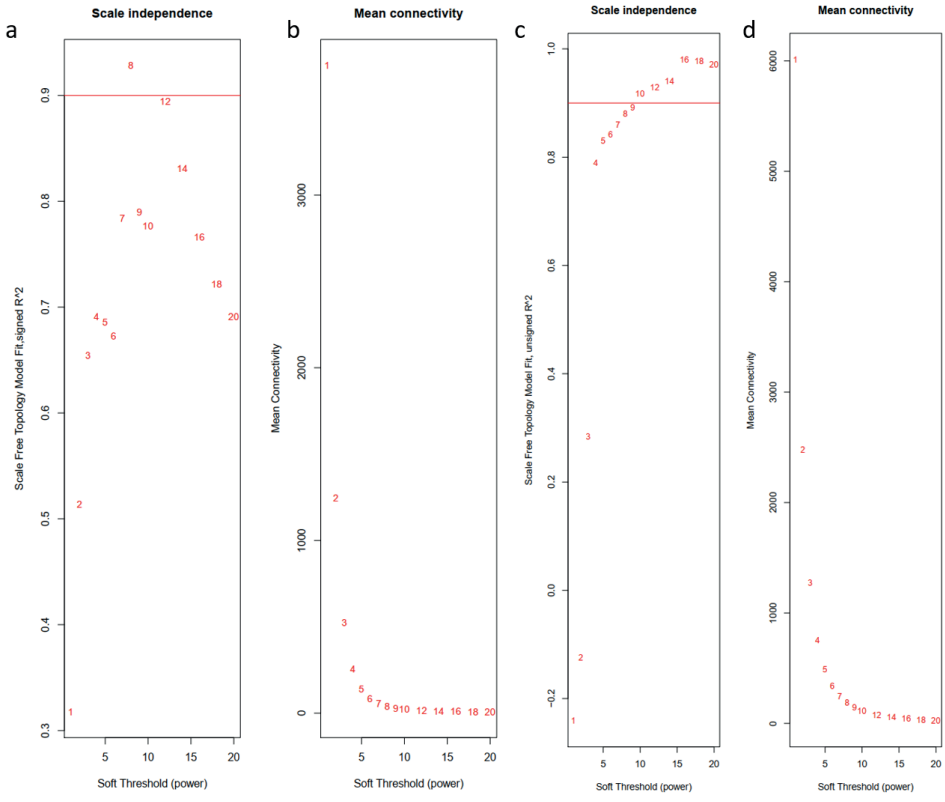


Figure 5.3 Soft threshold selection for BM and AF datasets. R^2>0.9 (red lines in a and c) for scale free independence fit index (y-axis,) for BM (a) and AF (c), and the soft threshold was set to 8 (BM) and 10 (AF), respectively. The mean connectivity (degree, y-axis) showed the mean nodes connectivity in the scale free network for different soft thresholds for BM (b) and AF (d).

BM gene expression module-trait relationships

To reveal the corresponding correlation between phenotypes and gene expression modules in BM, we determined the module-trait relationships (Figure 5.4). For the BM data set, a group of 522 genes within the cyan module is significantly negatively correlated with AFW (corresponding correlation = -0.56, p = 0.02) and significantly positively correlated with TG (corresponding correlation = 0.54, p = 0.02) (Figure 5.4). The cyan module gene list is provided in Table S5.5.

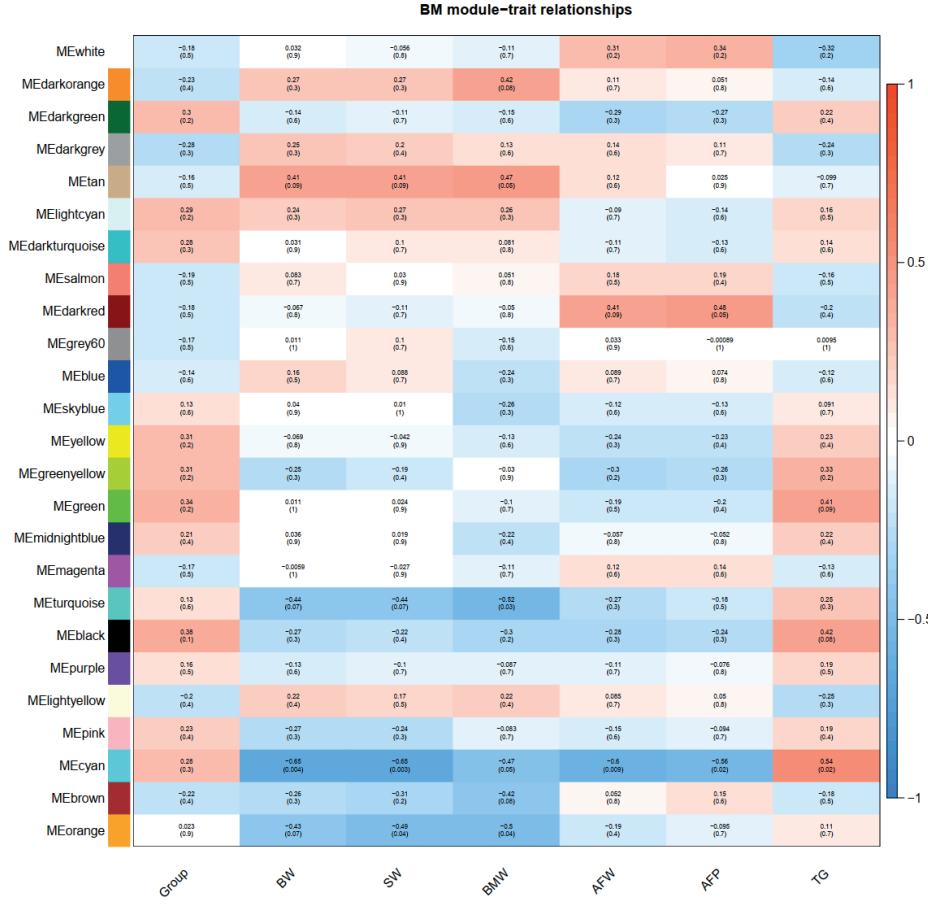


Figure 5.4 The module-trait relationship of BM. The color represents the correlation between module and trait. The upper number and lower number are the corresponding correlations and *p* value, respectively.

AF gene expression module-trait relationships

For the AF WGCNA results, a group of 498 genes within the light-cyan module is significantly negatively correlated with AFW (corresponding correlation = -0.60, p = 0.01). This module is also significantly positively correlated with breast muscle TG content (corresponding correlation = 0.56, p = 0.02) (Figure 5.5). The light-cyan module gene list is provided in Table S5.6.

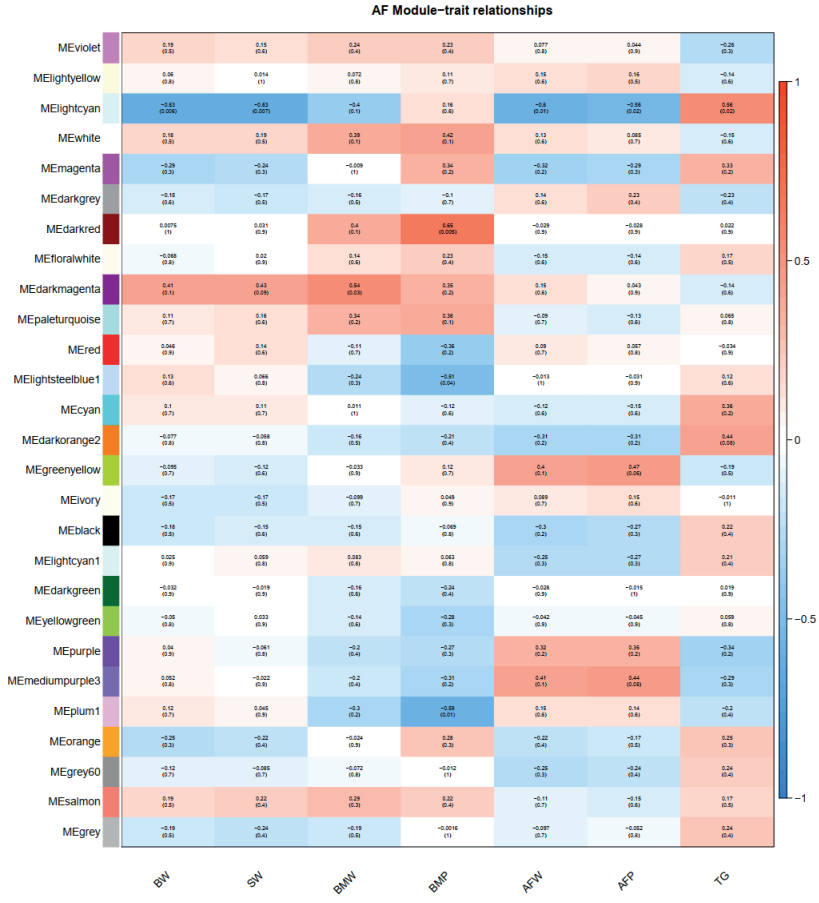


Figure 5.5 The module-trait relationship of AF. The upper number and lower number are the corresponding correlations and *p* value, respectively.

Hub genes in BM and AF datasets

Hub genes are expected to play an important role within networks. In the cyan module of the BM dataset, the *Acyl-CoA synthetase medium-chain family member 3* gene (*ACSM3*) is identified as a hub gene because it has the highest connectivity to other genes. In the light-cyan module of the AF dataset, the *cytochrome P450 family 2 subfamily AB polypeptide 1* gene (*CYP2AB1*) is identified as a hub gene. The networks of *ACSM3* and *CYP2AB1* are shown in Figure 5.6.

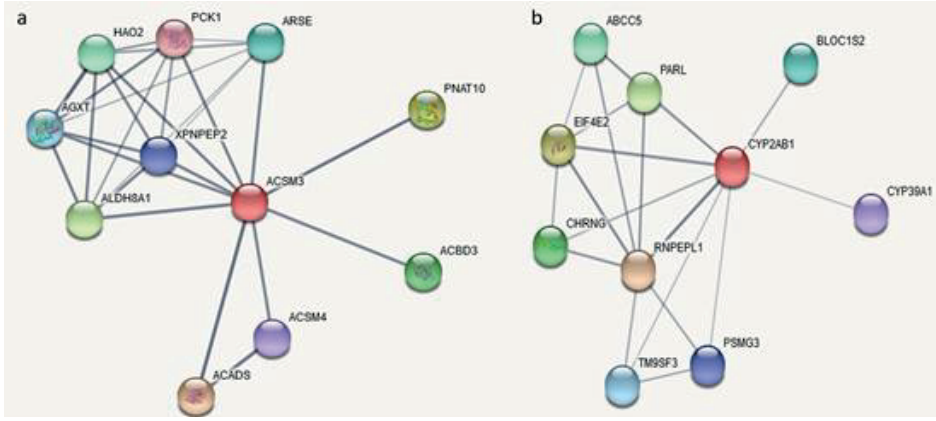


Figure 5.6 The networks of *ACSM3* and *CYP2AB1*.

5.4 Discussion

Chicken fatness traits are economically important traits. Chickens with high skeletal muscle IMF and low abdominal fat are desired. Yuan et al. (2020) used chickens with extreme AFW in fat and lean groups to detect the relationship between gene co-expression and AFW. In this study, we selected birds from high-IMF-low-AF and low-IMF-high-AF groups to increase the phenotypic variation for the gene co-expression networks analysis. No significant difference for BW and BMW is seen between the 2 groups, limiting the influence of BW and BMW on breast muscle IMF and AF. The large difference between the BM and AF expression profiles prohibited performing a consensus co-expression network analysis.

The detected gene *ACSM3* is involved in the fatty acid biosynthesis, fatty acid elongation, butanoate metabolism, glycerophospholipid metabolism, and glycerolipid metabolism pathways (Kanehisa and Goto, 2000). Although *ACSM3* is not very well studied in chicken, in human it has been shown to catalyze the activation of medium-chain length fatty acids (Watkins et al., 2007). It has been shown that when providing perfluorooctane sulfonate to fertilized eggs, which is similar to naturally occurring fatty acids, *ACSM3* expression in the embryo was downregulated (Jacobsen et al., 2018). *ACSM3* showed higher expression in primordial germ cells than in gonadal stromal cells and embryonic fibroblasts (Rengaraj et al., 2013). A study in cattle indicated that the differential expression of *ACSM3* is related to differences in fatty acid metabolism (Berton et al., 2016). This suggests that *ACSM3* is a potential candidate gene for selection of IMF and AF deposition in chicken.

The exact function of *CYP2AB1* is still not known. In the mouse gene expression atlas, *CYP2AB1* is highly expressed in adult lung, mammary gland, spleen and subcutaneous fat (Yue et al., 2014). *CYP2AB1* belongs to cytochromes P450 super family. In mammals, the cytochromes P450 enzymes oxidize steroids, fatty acids, and xenobiotics (De Montellano, 2005). In human, the CYP2 sub-family proteins are oxygenases, which catalyze mainly the metabolism of synthesis of cholesterol, steroids, and other lipids (Murray, 2016). *CYP2J2* decreased expressed in adipose tissue of obese individuals (Wamberg et al., 2013) and mesenteric arteries of obese rats (Zhao et al., 2005). This suggested that *CYP2AB1* may play an important role in chicken fat deposition.

5.5 Conclusion

In this study, we detected 2 gene co-expression modules in chicken AF and BM, that are significantly positively correlated to TG content and negatively correlated to AFW. Our results point to *ACSM3* and *CYP2AB1* as potential interesting candidates to select for increased breast muscle IMF content and decreased AFW in chickens.

5.6 Declarations

Sequencing Data Accession Numbers

The raw sequence data reported in this paper have been deposited in the Genome Sequence Archive (Wang et al., 2017b) in BIG Data Center (Members, 2018), Beijing Institute of Genomics (BIG), Chinese Academy of Sciences, under project accession numbers PRJCA005314 that is publicly accessible.

Author contributions

J. Wen, G.P., M.A.M.G., R.L., R.P.M.A.C., and S.X. designed the study; S.X., Q.W., and X.W. performed the animal experiments and samples collection; S.X., J.L., and X.W. tested the phenotypes; S.X. and J. Wang performed data analysis; S.X., R.L, R.P.M.A.C., and M.A.M.G., discussed the results; S.X. wrote the manuscript; R.L., R.P.M.A.C., M.A.M.G., and G.Z. provided valuable suggestion and comments to improve the manuscript with contributions from all other authors.

Acknowledgments

We thank Dr. Qinghe Li and Huanxian Cui (Institute of Animal Sciences, Chinese Academy of Agricultural Sciences) for valuable comments. We thank Dr. Zhou Wu (University of Edinburgh) for providing comments in RNA-seq SNP calling. We thank COMPASS Biotechnology for DNA genotyping and IGENECODE for transcriptome library construction and sequencing.

Funding

The study was supported by the National Natural Science Foundation of China (31772591), the China Agriculture Research System of MOF and MARA (CARS-41), Agricultural Science and Technology Innovation Program (ASTIP-IAS04), and China Scholarship Council (201903250098).

Ethics statement

The animal experiments were conducted at the chicken farm in the Institute of Animal Sciences of Chinese Academy of Agricultural Sciences, Beijing, China. The animal experiments had been approved by the animal ethics committee of Institute of Animal Sciences, Chinese Academy of Agricultural Sciences in 2016 for genotyping data (permission No. IASCAAS-AE-03) and in 2019 for transcriptome data (permission No. IAS2019-20).

Supplemental files can be found at https://osf.io/8uzmh/.

Table S5.1 Phenotypes of RNA-seq chickens.
Table S5.2 Alignment details of RNA-seq samples.
Table S5.3 Gene raw counts of BM samples.
Table S5.4 Gene raw counts of AF samples.
Table S5.5 The gene list of cyan module in BM dataset.
Table S5.6 The gene list of light-cyan module in AF dataset.
Figure S5.1 Breast muscle scale independence and mean connectivity for WGCNA.
Figure S5.2 AF scale independence and mean connectivity for WGCNA.
Figure S5.3 AF RNA-seq samples PCA plot.
Figure S5.4 BM RNA-seq samples PCA plot.
Figure S5.5 The distribution of significant ASE variants detected in AF.
Figure S5.6 The distribution of significant ASE variants detected in BM.
Figure S5.7 WGCNA cluster dendrogram of AF.

References

Andrews, S. (2010). *Babraham bioinformatics-FastQC a quality control tool for high throughput sequence data* [Online]. [Accessed July 17, 2017].

Berton, M.P., Fonseca, L.F., Gimenez, D.F., Utembergue, B.L., Cesar, A.S., Coutinho, L.L., et al. (2016). Gene expression profile of intramuscular muscle in Nellore cattle with extreme values of fatty acid. *BMC genomics* 17(1), 1-16.

Bolger, A.M., Lohse, M., and Usadel, B.J.B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. 30(15), 2114-2120.

Chen, X., Niu, J., and Geng, Z. (2017). Gene expression and plasma lipid content in relation to intramuscular fat in Chinese indigenous Wuhua chicken. *Journal of Applied Poultry Research* 26(3), 391-400.

Chizzolini, R., Zanardi, E., Dorigoni, V., and Ghidini, S. (1999). Calorific value and cholesterol content of normal and low-fat meat and meat products. *Trends in Food Science & Technology* 10(4-5), 119-128.

De Montellano, P.R.O. (2005). *Cytochrome P450: structure, mechanism, and biochemistry*. Springer.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: ultrafast universal RNA-seq aligner. 29(1), 15-21.

Fouad, A., and El-Senousey, H. (2014). Nutritional factors affecting abdominal fat deposition in poultry: a review. *Asian-Australasian journal of animal sciences* 27(7), 1057.

Jacobsen, A.V., Nordén, M., Engwall, M., and Scherbak, N. (2018). Effects of perfluorooctane sulfonate on genes controlling hepatic fatty acid metabolism in livers of chicken embryos. *Environmental Science and Pollution Research* 25(23), 23074-23081.

Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research* 28(1), 27-30.

Langfelder, P., and Horvath, S.J.B.b. (2008). WGCNA: an R package for weighted correlation network analysis. 9(1), 559.

Liu, L., Liu, X., Cui, H., Liu, R., Zhao, G., and Wen, J. (2019). Transcriptional insights into key genes and pathways controlling muscle lipid metabolism in broiler chickens. *BMC genomics* 20(1), 1-10.

Love, M.I., Huber, W., and Anders, S.J.G.b. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. 15(12), 550.

Members, B.D.C.J.N.A.R. (2018). Database Resources of the BIG Data Center in 2018. 46(Database issue), D14-D20.

Murray, M. (2016). CYP2J2–regulation, function and polymorphism. *Drug metabolism reviews* 48(3), 351-368.

Pertea, M., Kim, D., Pertea, G.M., Leek, J.T., and Salzberg, S.L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nature Protocols* 11(9), 1650-1667.

Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.C., Mendell, J.T., and Salzberg, S.L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol* 33(3), 290-295.

Rengaraj, D., Lee, B.R., Jang, H.-J., Kim, Y.M., and Han, J.Y. (2013). Comparative metabolic pathway analysis with special reference to nucleotide metabolism-related genes in chicken primordial germ cells. *Theriogenology* 79(1)**,** 28-39.

Resnyk, C.W., Carré, W., Wang, X., Porter, T.E., Simon, J., Bihan-Duval, E.L., et al. (2017). Transcriptional analysis of abdominal fat in chickens divergently selected on bodyweight at two ages reveals novel mechanisms controlling adiposity: validating visceral adipose tissue as a dynamic endocrine and metabolic organ. *Bmc Genomics* 18(1)**,** 626.

Resnyk, C.W., Carré, W., Wang, X., Porter, T.E., Simon, J., Le Bihan-Duval, E., et al. (2013). Transcriptional analysis of abdominal fat in genetically fat and lean chickens reveals adipokines, lipogenic genes and a link between hemostasis and leanness. *BMC genomics* 14(1)**,** 1-26.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., et al. (2003). Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. 13(11)**,** 2498.

Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., et al. (2019). STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic acids research* 47(D1)**,** D607-D613.

Wamberg, L., Christiansen, T., Paulsen, S., Fisker, S., Rask, P., Rejnmark, L., et al. (2013). Expression of vitamin D-metabolizing enzymes in human adipose tissue—the effect of obesity and diet-induced weight loss. *International journal of obesity* 37(5)**,** 651-657.

Wang, G., Kim, W.K., Cline, M.A., and Gilbert, E.R. (2017a). Factors affecting adipose tissue development in chickens: A review. *Poultry science* 96(10)**,** 3687-3699.

Wang, Y., Song, F., Zhu, J., Zhang, S., Yang, Y., Chen, T., et al. (2017b). GSA: Genome Sequence Archive*. 15(1)**,** 14-18.

Watkins, P.A., Maiguel, D., Jia, Z., and Pevsner, J. (2007). Evidence for 26 distinct acyl-coenzyme A synthetase genes in the human genomes. *Journal of lipid research* 48(12)**,** 2736-2750.

Xin, X., Zheng, M., Wen, J., and Wang, J. (2021). The situation analysis, future prospect and countermeasures of meat type chicken industry in China in 2020. *Chinese Journal of Animal Science* 57(3)**,** 217-222.

Xing, S., Liu, R., Zhao, G., Liu, L., Groenen, M.A., Madsen, O., et al. (2020). RNA-Seq analysis reveals hub genes involved in chicken intramuscular fat and abdominal fat deposition during development. *Frontiers in genetics* 11.

Yuan, H., and Lu, J. (2021). Consensus module analysis of abdominal fat deposition across multiple broiler lines. *BMC genomics* 22(1)**,** 1-14.

Yue, F., Cheng, Y., Breschi, A., Vierstra, J., Wu, W., Ryba, T., et al. (2014). A comparative encyclopedia of DNA elements in the mouse genome. *Nature* 515(7527)**,** 355-364.

Zhang, M., Li, F., Ma, X.-f., Li, W.-t., Jiang, R.-r., Han, R.-l., et al. (2019). Identification of differentially expressed genes and pathways between intramuscular and abdominal fat-derived preadipocyte differentiation of chickens in vitro. *BMC genomics* 20(1)**,** 1-15.

Zhao, X., Dey, A., Romanko, O.P., Stepp, D.W., Wang, M.-H., Zhou, Y., et al. (2005). Decreased epoxygenase and increased epoxide hydrolase expression in the mesenteric artery of obese Zucker rats. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* 288(1)**,** R188-R196.

6

General discussion

6.1 Introduction

Our world is facing a problem of increased demands for food. The chicken industry has seen a steady increase in production because of the different farm animals, chickens have the lowest feed conversion ratio and are therefore ideal to provide animal protein requirements of humans at low cost. Fat deposited in different tissues of chicken influences both the production efficiency and product quality. This thesis focused on different fat related tissues and aimed to reveal the genes and variants affecting fat deposition in chicken. The results presented in this thesis provided further insights into gene regulation during developmental stages in the liver, breast muscle and abdominal fat (AF). Finally, I also presented the results of a new SNP genotyping array and its use for selection on fatness traits.

The influence of subcutaneous fat (SF) on chicken meat quality is lower than that of intramuscular fat (IMF), while it is also lower in relation to feed efficiency than AF. Moreover, due to methodological difficulties of SF measurements, limited studies on SF are performed (Zhao et al., 2021). Therefore, SF is not further studied in my thesis. The chicken abdominal fat weight (AFW) is commonly considered to be a high heritability trait (Le Bihan-Duval et al., 1998; Zerehdaran et al., 2004; Jiang et al., 2017). In this general discussion, I will focus on the relationship between IMF and AF, the difficulties in chicken IMF genetic studies, and the current methodological limitations of genotyping arrays. Furthermore, I discuss the opportunities and challenges with transcriptome profiling in terms of chicken fat deposition studies from a developmental perspective. Finally, I discuss future possibilities to select chicken fatness traits using new technologies, and advantages of international cooperative projects.

6.2 Complexity of IMF genetics

As quantitative traits, fatness deposition in different tissues of the chicken is influenced by many different factors. These include, adipocyte differentiation and development, nutritional supply e.g. the amount of daily dietary fat digestion and absorption, lipid metabolism including the synthesis, degradation , and transport of fatty acids in and between cells (Griffin et al., 1982), diseases like non-alcoholic fatty liver syndrome (Tan et al., 2020), environmental factors like temperature (Lu et al., 2007) and behavior, as well as the interactions between these different factors (Leenstra, 1986). To obtain progress in optimizing chicken fat deposition on the short-term in can be obtained by adjusting the feed contents and limiting feed intake, a breeding approach is the preferred long-term approach. In this section, I will discuss the genetics underlying IMF deposition and the correlation between

IMF and AF. I also discuss the challenges that researchers face when working on a low heritability trait like IMF.

6.2.1 Intramuscular fat vs abdominal fat

Chicken IMF can only be measured after slaughter and the IMF measurement procedure with Soxhlet extraction (AOAC, 1990) is time and labor consuming, explaining the limited number of IMF heritability studies. Chicken AF traits include abdominal fat weight (AFW) and abdominal fat percentage (AFP), where AFP is the AFW to body weight ratio of an individual. Until now, there are not a lot of research (table 6.1) reported the genetic and phenotypic correlations between IMF and AF in chickens.

Table 6.1 Overview of some IMF percentage heritability studies in chicken breast muscle IMF.

Source breeds	Bird No.	Gender	heritability	Genetic correlations with AFW	Phenotypic correlations with AFW	Ref#
White Plymouth Rock	1,467	Both	0.08	0.02	0.05	a
Jingxing-Huang	1,428	Both	0.22	0.26	*	b
Beijing-You	1,069	Male	0.11	0.66	-0.09	c
Slow growing	1022	Both	0.18	*	*	d
Jingxing-Huang	1,335	Male	0.16	0.61	0.27	e

Notes: * Not given.
Ref: a: (Zerehdaran et al., 2004); b: (Zhao et al., 2006); c: (Chen et al., 2008); d: (Chabault et al., 2012), the breast muscle IMF genetic correlation with AFP is 0.32±0.21; e: (Jiang et al., 2017)

The reported heritability of chicken breast muscle IMF content varied from 0.08 to 0.22 in the different lines or breeds used (Table 6.1). The results indicate that the different genetic backgrounds of chicken could lead to different results for correlations between IMF and AF. Furthermore, Zerehdaran et al., (2004) tested the IMF percentage in minor pectoralis muscle, while Zhao et al., (2006) and Jiang et al., (2017) used the major pectoralis (Figure 6.1) for this phenotype, which likely is another reason for the different heritability for breast muscle IMF reported in these studies.

Either way, the chicken breast muscle IMF percentages can be considered as a low heritability trait. This causes challenges in the increased emphasis of selection on chicken breast muscle IMF.

Figure 6.1 The major pectoralis (left) and the minor pectoralis (right) of Jingxing-Huang chicken at marketing age (120 days).

The correlation between IMF and AFW is another important aspect. If the two traits were negatively correlated both in genetic and phenotypic aspects, the selection for increased IMF content and decreased AFW would be easier. Zerehdaran et al. (2004) reported that the genetic and phenotypic correlations of IMF percentage and AFW in 2 outcross broiler dam lines, which originated from the White Plymouth Rock, are very low (0.02 and 0.05, respectively). These results differ from the study in the IMF-up selection line and balanced selection line of Jingxing-Huang chicken with a genetic correlation of 0.67 and a phenotypic correlation of 0.27 (Jiang et al., 2017). Jiang et al. (2017) also reported that after 5 generations of selection, the IMF percentage in the IMF-up selection line increased 17.6%, while the AFP significantly increased 18.7%. This contradicts the requirements of the chicken industry which want to increase IMF and a decrease or unchanged AFP mainly because more AF means more energy wasted from feed. However, in the balanced line of Jingxing-Huang chicken, after 5 generations' selection, the IMF percentages increased 11.4% and the AFP decreased 1.5%. While the 1.5% decrease in the balanced selection line seems not a significant result, moving from

an increase of 19% to a decrease almost 2% suggests that it is possible to select chickens with high IMF and low AFP, and thus control the amount of IMF and AF.

6.2.2 IMF, a low heritability quantitative trait

Until now, 132 traits/disorders have been identified that show Mendelian inheritance and of which some have been used in modern chicken breeding (OMIA, 2021). However, none of these are related to fat deposition traits. According to animal QTL database, 489 QTLs have been reported to be associated with chicken fatness traits, of which 25 are associated with IMF percentage and 233 with AFW, respectively (Hu, 2021). A study based on an advanced intercross line, was the first to report suggestive evidence of IMF percentage QTLs detected on chicken chromosomes 1 and 27 (Jennen et al., 2005).

Since then, 12 QTLs on chicken breast muscle IMF percentage have been identified (Hu, 2021), but confirmation of QTLs between studies is hardly found. Potential reasons can be differences between the minor allele frequencies of causative or representative SNPs in the populations studied, different degrees of linkage disequilibrium (LD) within these populations, differences in the number of phenotyped individuals, and differences in the genetic structure of the cohorts, thereby affecting the power of the experiment. I would like to highlight that to improve the power in future research on heritability of chicken breast muscle IMF an advanced crossed population should be used, because of its higher phenotypic variation and shorter LD compared to the previously used pure lines (Table 6.1).

In the other farm animal species IMF, AF, and back fat is well studied in pigs. A functional regulatory structural variant in the porcine *MYH3* gene was identified to influence IMF (Cho et al., 2019), and a causative SNP in the porcine *MYH4* gene was detected to increase meat IMF content (Xiong, 2017). The *MYH* family drives the mammalian skeletal and cardiac muscle contraction (Weiss et al., 1999). Although, the *MYH3* and *MYH4* genes do not exist in chicken, the predicted *myosin heavy chain 1G* and predicted *myosin heavy chain 1F* are the homologous genes of the mammalian *MYH3* and *MYH4 genes*, respectively. I suppose that these genes may play an important role in chicken IMF content. Without doubt further identifications are needed.

6.3 Genotyping arrays: Advantages and limitations to study IMF

Since the publication of the first chicken reference genome sequence (International Chicken Genome Sequencing Consortium, 2004) and the initial identification of 2.8

million SNPs (International Chicken Polymorphism Map Consortium, 2004), a variety of genotyping arrays with different SNP densities have been produced (Muir et al., 2008; Groenen et al., 2011; Kranis et al., 2013; Liu et al., 2019a; Liu et al., 2021) that have been used extensively in studies focusing on genome wide associations (Luo et al., 2013; Li et al., 2021), linkage disequilibrium (Fu et al., 2015), population structure (Fleming et al., 2016), and for genomic selection (Liu et al., 2019b). In this section, I discuss the results of genome wide association studies (GWAS) and a genomic selection application for the chicken breast muscle IMF trait.

6.3.1 GWAS for breast muscle IMF content

GWAS have been widely used in studying the relationships between population genetics and traits since 2007 (Visscher et al., 2017). The different GWAS results for chicken breast muscle IMF are derived using different populations (Table 6.2).

Table 6.2 SNPs significantly associated with breast muscle IMF.

Breeds	Population size	SNPs No.*	Located Chromosomes	Ref#
F2 of Beijing-You × Cobb	367	5	1, 3, 4, 5, and Z	a
Jinghai Yellow	200	3	2	b
Beijing-You	724	2	2 and 5	c
Slow growing line	595	3	5	d
F2 of reciprocal crosses	272	NA	14	e

Notes: * GWAS identified significant trait associated SNPs number; # Reference: a (Sun et al., 2013); b (Zhang et al., 2015); c: The significance of two SNPs associated with breast muscle IMF did not reach the genome wide value (Liu et al., 2013); d: (Allais et al., 2019); e: F2 of reciprocal crosses between the genetically and phenotypically extreme inbred chicken lines New Hampshire and White Leghorn (Nassar et al., 2013).

The development of 60K genotyping array (Groenen et al., 2011), has enabled performing GWAS of breast muscle IMF in chicken. In 5 different studies (Table 6.2), a total of 13 SNPs were reported to be associated with chicken breast muscle IMF. The SNP density is too low to identify the causative mutations underlying the identified QTL and additional whole genome sequencing (WGS) is needed for fine mapping and identifying the actual causal variants. As the price of next generation sequencing has decreased considerably over the past 10-15 years, an increasing number of studies has used sequencing data to detect the potential candidate genes of fat deposition related traits in chicken (Moreira et al., 2018).

Genotyping arrays have several advantages and disadvantages compared to WGS. First, the cost of GWAS is an important aspect when performing a GWAS. Until now, using a medium-density SNP genotyping array is usually considered as more

cost efficient than using WGS. In addition to the direct sequencing costs of WGS, the data processing also is more complicated and time consuming than that for array data. Furthermore, the analysis pipelines of SNP genotyping data are well developed. The repeatability of genotyping array data is also better than that of WGS. Additionally, the repeatability of WGS also depends on the library construction, sequencing depth, data processing software, and SNP filtering criteria. Even so, the SNP density of WGS is much higher than using a genotyping array and insertions and deletions can be considered. Importantly, the causative mutations are more likely to be found using WGS data. To obtain DNA information from a large number of chickens, array data combined with WGS and imputation might be an alternative, cost-effective approach.

6.3.2 Genomic selection in chicken breast muscle IMF

Genomic selection was first suggested in 2001 (Meuwissen et al., 2001), and became feasible when methods were developed that allowed the simultaneous genotyping of tens of thousands of SNPs. This in turn made it possible for the method to be massively applied by the breeding industry. In chicken, genomic selection has been assessed for a wide variety of traits including growth and carcass traits (Liu et al., 2014b), feed efficiency (Liu et al., 2017), antibody response (Liu et al., 2014a), and reproduction traits (Wolc et al., 2011).

It has been shown that genomic selection with re-training could reduce 75% of the rearing and 82% of the phenotype recording in layers in achieving the breeding goal (Wolc et al., 2011). This implicates that, genomic selection can be cost-effective to be used in the selection of IMF, because measuring the IMF content test is labor and time consuming. Genomic selection was used in Beijing-You chicken using the breast muscle IMF content as the sole trait (Liu et al., 2019b). The results showed that compared with the control group (random mating), IMF improved 9.2%, after 1 generation of genomic selection, while IMF improved 10.38% in the family-based selection group (Liu et al., 2019b). The difference in selection response might be caused by the limited size of the training population used. Additionally, the low heritability of IMF might have resulted in a low genomic prediction accuracy.

The new developed chicken 55K SNP genotyping array has been used in several genomic selection progress in yellow-feathered meat-type chickens in China. Now, the genotyping array has been updated to the third version with more economic traits related locus. It accelerated the Chinese chicken breeding progress and played a good demonstration effect.

6.4 Developmental and tissue specific regulation of gene expression

According to the central dogma, gene expression is the bridge between DNA and protein (Crick, 1970). The FarmGTEx (Farm Animal Genotype-Tissue Expression) cooperation project (Liu et al., 2020) that has recently been initiated, is aiming to provide a comprehensive atlas of tissue-specific gene expression and genetic regulation in livestock. One major objective of my thesis was to identify the candidate genes influencing IMF and AF using gene expression data obtained from relevant tissues at different developmental stages. In this section, I discuss the gene expression data as a future resource for fatness trait studies and the use of the identified correlations between gene expression and quantitative traits.

There is a gene expression database related to the development of multi organs in red jungle fowl (Cardoso-Moreira et al., 2019). This database provides gene expression data for brain, cerebellum, heart, kidney, liver, ovary, and testis during development. The data is illustrative for the transcriptome complexity between tissues. The differences in the timing of key events during the development of the gonads were identified. And the result of gene expression range gradually decrease during development is similar with the results obtained in chapter 3 and 4. The resource of 7 tissues however is still limited and transcriptome data from additional tissues are needed.

Gene expression data can be a useful resource for detecting causative mutations underlying important phenotypic traits. For instance, Xiong et al., (2017) found a 526 Kb QTL region significantly associated with longissimus dorsi IMF content in pig. This region harbors 9 genes but 3 of these are not expressed. Furthermore, only the *MYH4* gene expression level was significantly correlated with IMF content. Finally, a SNP in the *MYH4* gene was identified as a causative SNP and shown to be regulating the binding of nuclear regulatory proteins. In this thesis, I provide the transcriptome profiling of breast muscle, liver, and abdominal fat during development as well as marketing time. The data can assist further candidate genes identification of fatness traits in chicken.

6.5 Future of chicken fatness traits selection

6.5.1 The reference genome

A primary goal of genome research in the agricultural field is to use genomic information to improve selection (Rexroad et al., 2019). Recently, in human, a complete reference from telomere to telomere of all human chromosomes has

now been assembled (Miga et al., 2020; Nurk et al., 2021). The genome information of previous unresolved gaps made the integrated phenotype-genotype research. Chicken was the first agricultural animal for which a reference genome was assembled (International Chicken Genome Sequencing Consortium, 2004). The reference genome of chicken was derived from a female of UCD001, an inbred line of red jungle fowl. In the latest assembled chicken reference genome, all the micro-chromosomes (chromosome 29 to 39) were assembled (NCBI genome, 2021). Compared with the previously assembled reference genome (build galGal6), the higher level of completeness can provide further insight between the genotype and phenotype. By updating the SNPs array with probes focusing on the more complete reference genome including the newly assembled micro chromosomes and improvement of the other chromosomes, will be beneficial to find potential associations with traits which were not possible in the past.

6.5.2 Integrated analysis of multi omics data

Chicken fatness traits are complex and quantitative. Filling the gap between chicken fatness genotype and fatness phenotype not only requires accurate phenotypes but also a complete functional genome annotation. The Functional Annotation of ANimal Genomes project (FAANG) aims to provide a comprehensive insight of functional elements by integrating multi omics assays e.g., assay for transposase-accessible chromatin sequencing (ATAC-seq), chromatin immunoprecipitation sequencing (ChIP-seq), and whole genome bisulfite sequencing (WGBS) (The FAANG Consortium et al., 2015; Clark et al., 2020). For instance, skeletal muscle and adipose are among the target tissues for these assays. Obtaining regulatory elements maps associated to genes including variation in these elements are important to detect variation in gene expression related to certain traits, like enhancer were provided. Another important improvement is getting access to long read sequencing to get information on sequence haplotypes and in case of RNA of complete sequenced transcripts (Kuo et al., 2017). These long read sequences are also very important to detect structural variation present in the genomes which might be related to traits (Giuffra et al., 2019). By using data provided by the FAANG project, researchers can obtain deeper insight in the transcribed loci and elements involved in their regulation. The data provided by FAANG includes knowledge of modified histones, DNA methylation, chromatin accessibility, and spatial conformation of chromatin. In the future, these data can help researchers identify the candidate genes or causative mutations related to fatness traits.

6.5.3 Gene editing

Gene editing combined with genomic selection has been suggested to be used in cattle to reduce the dehorning costs (Carlson et al., 2016). Unlike mammals, gene editing in poultry is much more difficult because of the difficulties for accessing and manipulating the zygote (Mizushima et al., 2010) and therefor lag far behind those in mammals. Two studies reported the generation of single SNP modified live chickens by using Transcription Activator-Like Nuclease (TALEN) and Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR), respectively (Park et al., 2014; Oishi et al., 2016). Although the success rates of both gene editing methods are still low in chicken, I can imagine that it must be popular and improvements in the technology and methods will make it more accessible soon. I think it is easier to perform gene editing at causative variations for monogenetic traits. But for the polygenetic quantitative traits, like IMF content and abdominal fat, the causative variations are still difficult to identify. What is more, even if the causative variations have been identified, the number of causative variations likely run in the hundreds perhaps even thousands for the low heritability traits. Until now, it is not possible to edit the multiple target mutations together in one operable round.

6.6 Conclusions

In this thesis, I got information on differential expression of genes during development in different tissues. By detecting hup genes related to fat deposition, variation in those genes can be used with existing variation of known genes related to fat deposition and fat formation to improve IMF in chicken breast muscle. Adding these variants to existing SNP arrays will further enrich chip information, so that it can play a greater role in the selection of fatness traits in chicken.

References

Allais, S., Hennequet-Antier, C., Berri, C., Salles, L., Demeure, O., and Le Bihan-Duval, E. (2019). Mapping of QTL for chicken body weight, carcass composition, and meat quality traits in a slow-growing line. *Poultry science* 98(5), 1960-1967.

AOAC (1990). Official Methods of Analysis: Fat or Ether Extract in Meat. *Association of Official Analytical Chemists*.

Cardoso-Moreira, M., Halbert, J., Valloton, D., Velten, B., Chen, C., Shao, Y., et al. (2019). Gene expression across mammalian organ development. *Nature* 571(7766), 505-509.

Carlson, D.F., Lancto, C.A., Zang, B., Kim, E.-S., Walton, M., Oldeschulte, D., et al. (2016). Production of hornless dairy cattle from genome-edited cell lines. *Nature biotechnology* 34(5), 479-481.

Chabault, M., Baéza, E., Gigaud, V., Chartrin, P., Chapuis, H., Boulay, M., et al. (2012). Analysis of a slow-growing line reveals wide genetic variability of carcass and meat quality-related traits. *BMC genetics* 13(1), 1-8.

Chen, J., Zhao, G., Zheng, M., Wen, J., and Yang, N. (2008). Estimation of genetic parameters for contents of intramuscular fat and inosine-5′-monophosphate and carcass traits in Chinese Beijing-You chickens. *Poultry Science* 87(6), 1098-1104.

Cho, I.-C., Park, H.-B., Ahn, J.S., Han, S.-H., Lee, J.-B., Lim, H.-T., et al. (2019). A functional regulatory variant of MYH3 influences muscle fiber-type composition and intramuscular fat content in pigs. *PLoS genetics* 15(10), e1008279.

Clark, E.L., Archibald, A.L., Daetwyler, H.D., Groenen, M.A., Harrison, P.W., Houston, R.D., et al. (2020). From FAANG to fork: application of highly annotated genomes to improve farmed animal production. *Genome Biology* 21(1), 1-9.

Crick, F. (1970). Central dogma of molecular biology. *Nature* 227(5258), 561-563.

Fleming, D., Koltes, J., Markey, A., Schmidt, C., Ashwell, C., Rothschild, M., et al. (2016). Genomic analysis of Ugandan and Rwandan chicken ecotypes using a 600 k genotyping array. *BMC genomics* 17(1), 1-16.

Fu, W., Dekkers, J.C., Lee, W.R., and Abasht, B. (2015). Linkage disequilibrium in crossbred and pure line chickens. *Genetics Selection Evolution* 47(1), 1-12.

Giuffra, E., Tuggle, C.K., and Consortium, F. (2019). Functional Annotation of Animal Genomes (FAANG): current achievements and roadmap. *Annual review of animal biosciences* 7, 65-88.

Griffin, H., Whitehead, C., and Broadbent, L. (1982). The relationship between plasma triglyceride concentrations and body fat content in male and female broilers—a basis for selection? *British Poultry Science* 23(1), 15-23.

Groenen, M.A., Megens, H.-J., Zare, Y., Warren, W.C., Hillier, L.W., Crooijmans, R.P., et al. (2011). The development and characterization of a 60K SNP chip for chicken. *BMC genomics* 12(1), 1-9.

Hu, Z.-L.P., Carissa; Reecy, James M. (2021). *Animal QTLdb* [Online]. Available: https://www.animalgenome.org/QTLdb [Accessed June 15th 2021].

International Chicken Genome Sequencing Consortium (2004). Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* 432(7018)**,** 695-716.

International Chicken Polymorphism Map Consortium (2004). A genetic variation map for chicken with 2.8 million single nucleotide polymorphisms. *Nature* 432(7018)**,** 717.

Jennen, D.G., Vereijken, A.L., Bovenhuis, H., Crooijmans, R.M., van der Poel, J.J., and Groenen, M.A. (2005). Confirmation of quantitative trait loci affecting fatness in chickens. *Genetics Selection Evolution* 37(3)**,** 1-14.

Jiang, M., Fan, W.L., Xing, S.Y., Wang, J., Li, P., Liu, R.R., et al. (2017). Effects of balanced selection for intramuscular fat and abdominal fat percentage and estimates of genetic parameters. *Poult Sci* 96(2)**,** 282-287.

Kranis, A., Gheyas, A.A., Boschiero, C., Turner, F., Le, Y., Smith, S., et al. (2013). Development of a high density 600K SNP genotyping array for chicken. *Bmc Genomics* 14(1)**,** 59.

Kuo, R.I., Tseng, E., Eory, L., Paton, I.R., Archibald, A.L., and Burt, D.W. (2017). Normalized long read RNA sequencing in chicken reveals transcriptome complexity similar to human. *BMC Genomics* 18(1)**,** 323.

Le Bihan-Duval, E., Mignon-Grasteau, S., Millet, N., and Beaumont, C. (1998). Genetic analysis of a selection experiment on increased body weight and breast muscle weight as well as on limited abdominal fat weight. *British Poultry Science* 39(3)**,** 346-353.

Leenstra, F. (1986). Effect of age, sex, genotype and environment on fat deposition in broiler chickens—A review. *World's Poultry Science Journal* 42(1)**,** 12-25.

Li, W., Zheng, M., Zhao, G., Wang, J., Liu, J., Wang, S., et al. (2021). Identification of QTL regions and candidate genes for growth and feed efficiency in broilers. *Genetics Selection Evolution* 53(1)**,** 1-17.

Liu, R., Sun, Y., Zhao, G., Wang, F., Wu, D., Zheng, M., et al. (2013). Genome-wide association study identifies loci and candidate genes for body composition and meat quality traits in Beijing-You chickens. *PLoS One* 8(4)**,** e61172.

Liu, R., Xing, S., Wang, J., Zheng, M., Cui, H., Crooijmans, R.P., et al. (2019a). A new chicken 55K SNP genotyping array. *BMC genomics* 20(1)**,** 1-12.

Liu, R., Zheng, M., Wang, J., Cui, H., Li, Q., Liu, J., et al. (2019b). Effects of genomic selection for intramuscular fat content in breast muscle in Chinese local chickens. *Animal genetics* 50(1)**,** 87-91.

Liu, S., Gao, Y., Canela-Xandri, O., Wang, S., Yu, Y., Cai, W., et al. (2020). A comprehensive catalogue of regulatory variants in the cattle transcriptome. *bioRxiv*.

Liu, T., Luo, C., Wang, J., Ma, J., Shu, D., Lund, M.S., et al. (2017). Assessment of the genomic prediction accuracy for feed efficiency traits in meat-type chickens. *Plos one* 12(3)**,** e0173620.

Liu, T., Qu, H., Luo, C., Li, X., Shu, D., Lund, M.S., et al. (2014a). Genomic selection for the improvement of antibody response to Newcastle disease and avian influenza virus in chickens. *PloS one* 9(11)**,** e112685.

Liu, T., Qu, H., Luo, C., Shu, D., Wang, J., Lund, M.S., et al. (2014b). Accuracy of genomic prediction for growth and carcass traits in Chinese triple-yellow chickens. *BMC genetics* 15(1)**,** 1-8.

Liu, Z., Sun, C., Yan, Y., Li, G., Li, X.C., Wu, G., et al. (2021). Design and evaluation of a custom 50K Infinium SNP array for egg-type chickens. *Poultry Science* 100(5)**,** 101044.

Lu, Q., Wen, J., and Zhang, H. (2007). Effect of chronic heat exposure on fat deposition and meat quality in two genetic types of chicken. *Poultry science* 86(6)**,** 1059-1064.

Luo, C., Qu, H., Ma, J., Wang, J., Li, C., Yang, C., et al. (2013). Genome-wide association study of antibody response to Newcastle disease virus in chicken. *BMC genetics* 14(1)**,** 1-9.

Meuwissen, T.H., Hayes, B.J., and Goddard, M.E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157(4)**,** 1819-1829.

Miga, K.H., Koren, S., Rhie, A., Vollger, M.R., Gershman, A., Bzikadze, A., et al. (2020). Telomere-to-telomere assembly of a complete human X chromosome. *Nature* 585(7823)**,** 79-84.

Mizushima, S., Takagi, S., Ono, T., Atsumi, Y., Tsukada, A., Saito, N., et al. (2010). Novel method of gene transfer in birds: intracytoplasmic sperm injection for green fluorescent protein expression in quail blastoderms. *Biology of reproduction* 83(6)**,** 965-969.

Moreira, G.C.M., Boschiero, C., Cesar, A.S.M., Reecy, J.M., Godoy, T.F., Pértille, F., et al. (2018). Integration of genome wide association studies and whole genome sequencing provides novel insights into fat deposition in chicken. *Scientific reports* 8(1)**,** 1-14.

Muir, W., Wong, G., Zhang, Y., Wang, J., Groenen, M., Crooijmans, R., et al. (2008). Review of the initial validation and characterization of a 3K chicken SNP array. *World's Poultry Science Journal* 64(2)**,** 219-226.

Nassar, M., Goraga, Z., and Brockmann, G. (2013). Quantitative trait loci segregating in crosses between N ew H ampshire and W hite L eghorn chicken lines: III. Fat deposition and intramuscular fat content. *Animal genetics* 44(1)**,** 62-68.

NCBI genome (2021). Available: https://www.ncbi.nlm.nih.gov/genome/?term=chicken [Accessed June 22 2021].

Nurk, S., Koren, S., Rhie, A., Rautiainen, M., Bzikadze, A.V., Mikheenko, A., et al. (2021). The complete sequence of a human genome. *bioRxiv*.

Oishi, I., Yoshii, K., Miyahara, D., Kagami, H., and Tagami, T. (2016). Targeted mutagenesis in chicken using CRISPR/Cas9 system. *Scientific reports* 6(1)**,** 1-10.

OMIA (2021). *Online Mendelian Inheritance in Animals* [Online]. Available: https://www.omia.org/home/ [Accessed June 12, 2021].

Park, T.S., Lee, H.J., Kim, K.H., Kim, J.-S., and Han, J.Y. (2014). Targeted gene knockout in chickens mediated by TALENs. *Proceedings of the National Academy of Sciences* 111(35)**,** 12716-12721.

Rexroad, C., Vallet, J., Matukumalli, L.K., Reecy, J., Bickhart, D., Blackburn, H., et al. (2019). Genome to phenome: improving animal health, production, and well-being–a new USDA blueprint for animal genome research 2018–2027. *Frontiers in genetics* 10**,** 327.

Sun, Y., Zhao, G., Liu, R., Zheng, M., Hu, Y., Wu, D., et al. (2013). The identification of 14 new genes for meat quality traits in chicken using a genome-wide association study. 14(1)**,** 458.

Tan, X., Liu, R., Xing, S., Zhang, Y., and Wen, J. (2020). Genome-Wide Detection of Key Genes and Epigenetic Markers for Chicken Fatty Liver. *International Journal of Molecular Sciences* 21(5)**,** 1800.

The FAANG Consortium, Andersson, L., Archibald, A.L., Bottema, C.D., Brauning, R., Burgess, S.C., et al. (2015). Coordinated international action to accelerate genome-to-phenome with FAANG, the Functional Annotation of Animal Genomes project. *Genome biology* 16(1)**,** 1-6.

Visscher, P.M., Wray, N.R., Zhang, Q., Sklar, P., McCarthy, M.I., Brown, M.A., et al. (2017). 10 years of GWAS discovery: biology, function, and translation. *The American Journal of Human Genetics* 101(1)**,** 5-22.

Weiss, A., Schiaffino, S., and Leinwand, L.A. (1999). Comparative sequence analysis of the complete human sarcomeric myosin heavy chain family: implications for functional diversity. *Journal of molecular biology* 290(1)**,** 61-75.

Wolc, A., Stricker, C., Arango, J., Settar, P., Fulton, J.E., O'Sullivan, N.P., et al. (2011). Breeding value prediction for production traits in layer chickens using pedigree or genomic relationships in a reduced animal model. *Genetics Selection Evolution* 43(1)**,** 1-9.

Xiong, X. (2017). *Identification and Mechanism of Causal Gene and Its Causal Mutation on Porcine Chromosome 12 Increasing Intramuscular Fat Content.* Jiangxi Agricultural University.

Zerehdaran, S., Vereijken, A.J., Van Arendonk, J., and Van der Waaijt, E. (2004). Estimation of genetic parameters for fat deposition and carcass traits in broilers. *Poultry science* 83(4)**,** 521-525.

Zhang, T., Fan, Q., Wang, J., Zhang, G., Gu, Y., and Tang, Y. (2015). Genome-wide association study of meat quality traits in chicken. *Genet Mol Res* 14**,** 10452-10460.

Zhao, G., Wen, J., Chen, J., Zheng, M., and Li, X. (2006). Selection response and estimation of the genetic parameters for intramuscular fat in a quality chicken line. *J Acta veterinaria et zootechnica Sinica* 37(9)**,** 870-873.

Zhao, H., Wu, M., Tang, X., Li, Q., Yi, X., Wang, S., et al. (2021). Function of chick subcutaneous adipose tissue during the embryonic and posthatch period. *Frontiers in Physiology* 12**,** 831.

Summary

Domesticated animal meat production and quality are closely related to muscle growth and fat deposition. Fat deposition is determined by adipocyte differentiation and development. Fatness related genes and the regulation of their expression play an important role in fat deposition. In this thesis, I describe chicken breast muscle intramuscular fat (IMF) and abdominal fat (AF) deposition during development. I provide gene expression profiles in breast muscle, liver, and abdominal fat. By using transcriptome analyses to obtain new insights into the correlation between gene expression and fat in chickens, I identified new fatness related candidate genes. The new understanding of gene expression and chicken fat deposition can potentially be used in chicken breeding.

In **Chapter 2**, I describe the development of a new chicken SNP genotyping array. The SNPs on the array are based on i) whole-genome sequencing from 5 Chinese traditional local breeds and 3 commercial lines, ii) top significant associated SNPs for 15 traits of interest, which were identified using the 60 K chicken SNP array, iii) SNPs from candidate genes e.g., differentially expressed genes for IMF in fast- and slow-growing chickens, iv) SNPs related to feed efficiency, and v) SNPs from a commercial 600 K genotyping array. The genotyping results of the newly developed SNP array in 13 different breeds/lines show a high genotyping rate and the breeds/lines can be clearly distinguished. The SNP array has been utilized for genomic selection in chicken, genome-wide association studies, and to characterize population structure and diversity of different selection lines.

Chapter 3 focuses on a study to identify candidate genes of breast muscle IMF and abdominal fat deposition. Described are the IMF and abdominal fat deposition during development in a slow-growing chicken line. Fat accumulation in breast and abdomen both accelerate from day 56 after hatching. Transcriptome profiling of chicken breast muscle and abdominal fat during developmental stages resulted in the clustering of developmentally dynamic genes. Hub genes affecting these traits were identified e.g., *ENSGALG00000041996*, a candidate for high breast muscle IMF, and *CREB3L1*, related to low abdominal fat weight. The transcription factor *L3MBTL1* and the transcription factor co-factors *TNIP1*, *HAT1*, and *BEND6* showed a correlation to both high breast muscle IMF and low abdominal fat weight.

The objective of **Chapter 4** is to provide a comprehensive understanding of the gene regulation in chicken liver during development. RNA-seq data of liver samples from the embryonic stage to the egg-laying stage, which were derived from slow-growing female chickens, were generated to identify differentially expressed genes between adjacency stages. The differentially expressed genes were enriched in

pathways for fatty acid metabolism, biosynthesis of unsaturated fatty acids, fatty acid degradation, and PPAR signalling. To identify hub genes during liver development, a cluster analysis of the developmentally dynamic genes was performed. This resulted in the identification of *ACSBG2* as a candidate gene correlated to abdominal fat weight. The *ACSBG2* gene overlaps with 2 abdominal fat weight QTLs on chromosome 28.

In **chapter 5**, we further explore gene expression regulation of IMF and abdominal fat deposition in chicken. We used breast muscle and abdominal fat RNA-seq transcriptome data of high-IMF-low-AF and low-IMF-high-AF chickens from a slow growing dwarf line at marketing time. To avoid external influencing factors, we employed the weighted gene co-expression analysis on the transcriptome data. Two hub genes, *ACSM3* in breast muscle and *CYP2AB1* in abdominal fat, were identified which are both significantly positively correlated to IMF and significantly negatively correlated to low abdominal fat weight.

Finally, in **chapter 6**, the general discussion, I discuss the complexity of chicken IMF genetics. I address the advantages and limitations to study chicken IMF deposition by genotyping arrays. Concerning the correlation between gene expression and fat deposition, the potential usage of RNA-seq on fatness studies is discussed. I end my general discussion by emphasizing the future of chicken fatness traits selection.

Acknowledgements

During the five years of my PhD project, I had the pleasure to meet many great people. I would like to take this opportunity to thank all those people who have contributed to this thesis and helped me.

I would like to firstly express my gratitude to my promoter, Martien, you are so supportive throughout my PhD project. Thank you for allowing me to study at ABG. Thank you for the guidance on my research. Thank you for your timely feedback on my manuscripts and endless patience in the meetings. Looking forward to seeing you in physical life soon and in the future.

Richard! How lucky am I to have you as my daily supervisor! I still remember the first morning I arrived at Wageningen, and you took me into Radix, it just likes yesterday. I do admire the meetings, discussions, and of course beers in the past several years. Thank you for saying that “Let it go” and “Make it open”. Without you encourage me during the work from home period, I cannot finish my thesis. Thank you!

I also would like to express my gratitude to Ole. Thank you very much for teaching me how to use HPC and the valuable discussions on my RNA-seq papers. I would like to thank Mirte, Hendrik-Jan and all my ABG colleagues and office mates: Vinicius, Chiara, Henri, Lim, Pascal, Floor, Maulik, Tom, Michael, Harmen, and Rodrigo.

I would like to thank CAAS-WUR sandwich PhD project coordinator, Claudius, you are my favorite hello and hardest goodbye in Wageningen. I learned a lot from you and still learning from you. Your experience and attitude make me understand more about what the hell is science.

I would like to thank Bruce Currie for the kind help in PhD position applications and encourage.

I would like to express my special thank you for my paranymphs. Martijn, thank you very much for your help in my PhD journey and defence. I do appreciate all those discussions on genomics, codes, R, working, family, life, and future. 感谢张陈，非常荣幸能交到你这个朋友，你对科学的认识、我们的彻夜长谈让我受益匪浅。

致谢

时光如白驹过隙，转眼间博士研究生生活已接近尾声，回首在课题组六年多来的点点滴滴，心中感概万千，满是不舍。

感谢我的导师文杰研究员，感谢您付出的大量心血和指导，您以身作则的科研作风，您独特的见解和思路总能让我受益匪浅。您对科研严谨认真的态度也时刻感染着我，鼓舞着我，让我对学习和工作充满了严谨与责任。谢谢您！

感谢我的日常指导老师刘冉冉研究员，您渊博的学识和严谨的态度，是我学习的榜样，更能白明白自己要努力的方向。您不仅在学习上给予我们指导，还以身作则教给了我们很多做人的道理，让我们在以后的路上成为更好的自己。

感谢赵桂苹研究员，感谢您在生活和学习中给予的照顾，您在工作中的严谨和雷厉风行以及生活中的豁达与平易近人，都为我们树立了学习的榜样，激励我不断进步。

感谢鸡遗传育种创新团队郑麦青老师、李庆贺老师、崔焕先老师、刘丽老师、王巧老师，张永宏博士后、孙玮玮博士后、原小雅博士后在我学习和生活中给予的指导、帮助和照顾。

感谢鸡遗传育种创新团队师兄李鹏，师姐郭立平，同学宋娇、李韦、王巧，师弟刘振、蒋明、营凡、王杰、谭晓冬、朱波、李森、王海龙、何政肖、丁纪强、杜永旺、蔡克奇，师妹刘璐、李京徽、王希彩、王菲、陈颖、金雨锡、刘晓静、张锦、杨欣婷、张琪、骆娜、孙嘉宏、王永丽、白露等在工作和学习中的帮助、关心和支持。感谢中国畜牧兽医研究昌平试验基地全体工作人员，感谢您们的付出使得试验顺利进行。

感谢中国农业科学院研究生院国际教育处张明军副处长。张老师！谢谢您！

感谢中国农业大学胡晓湘教授在博士申请过程中的帮助，感谢刘建峰教授、赵鹏举博士、冯文博士，在芯片设计工作中的帮助。感谢中国农科院特产研究所师兄张铁涛博士、杨童奥博士、刘华淼博士在我博士期间的鼓励和支持。

感谢我的硕士导师杨福合研究员和本科论文指导老师晏向华教授带领我走入科研的世界。

当然还要感谢在瓦村的诸位中国同胞。感谢我的 buddy 小包，带我逛校园逛吃乌特勒支、鹿特丹……从一起买肉建立起友谊的吴袜子和牛郎，感谢提供各种遗传育种基因组的抱大腿服务，但看前面黑黑洞洞，定是那贼人巢穴，待我赶上前去，杀他个干干净净；感谢哎必鸡的路海博、Mandy、老司机、刘文哲、

黄其通、余小飞、喻运、王卓识、王芳。感谢房东黄佳琦，室友黄一帆、于宏威，廊友渠清博、王然，上铺兄弟李云雷的不杀之恩。感谢瓦赫宁根的同学们：蹭饭狂魔张成龙、孟凡荣、孙秀路、徐蔚、陈敏杰、周礼斌、杨宁、胡义信、王天鹏、陈浩、宫晓倩、崔宏瑞、缪秋红、张嘉琪、周波、张雪、华登科、巫可佳、陈玫君、程喆……感谢你们对我的帮助和照顾！

感谢我的母亲谢新红副教授、我的妻子张琼、女儿邢一苇，和所有亲人们这些年来的支持。我爱你们！

邢思远，2021 年 9 月于荷兰瓦赫宁根

Curriculum Vitae

Siyuan Xing was born on July 24th 1988 in Changzhi, Shanxi, China. He developed interest during his MSc period. In 2007, he studied Animal Science at Huazhong Agricultural University (HZAU). He obtained his bachelor degree in 2011 supervised by Professor Xianghua Yan. He studied rearing special economic animals in the Institute of Special Animal and Plant Sciences of the Chinese Academy of Agricultural Sciences (CAAS). He got his MSc degree in 2014, supervised by Professor Fuhe Yang. He joined the CEVA-HUADU company until August 2015. Then he worked as a supporter in the Institute of Animal Sciences of CAAS. From 2016, he started his sandwich PhD journey with thesis titled “Differential Deposition of Intramuscular and Abdominal Fat in Chicken”.

Training and Supervision Plan (TSP)	**Graduate School WIAS**
A. The Basic Package	
WIAS Introduction Day **(mandatory)**	2017
Science integrity & Course on philosophy of science and/or ethics **(mandatory)**	2018
Course on essential skills (Frank Little) *(recommended)*	2017
Applied Ethics	2017
B. Disciplinary Competences	
proposals	2017
ASReml	2018
Machine learning with Python	2020
C. Professional Competences	
Efficient and Effective Academic Development	2016
Academic Writing and Presenting in English	2016
Writing Scientific Proposals	2017
Survival Guide of Peer Review	2017
Data management planning	2018
Project and Time Management	2018
Supervising BSc & MSc thesis	2018
Reviewing a scientific manuscript	2020
Writing the propositions	2020
Writing the General Introduction and Discussion	2020
D. Presentation Skills *(maximum 4 credits)*	
Egg&Meat, Turkey, Oral Presentation	2019
WIAS science day, Poster	2018
E. Teaching competences *(max 6 credits)*	
Supervising an MSc major	2019
Supervising an MSc major	2020
Supporting course (English writing)	2018
Supporting course (Scientific development)	2019
Total	**30.3**

Publications and patent:

Siyuan Xing*, Ranran Liu*, Guiping Zhao, Martien Groenen, Ole Madsen, et al. Time course transcriptomic study reveals the gene regulation during liver development and the correlation with abdominal fat weight in chicken. *Frontiers in genetics* (2021).

Siyuan Xing*, Ranran Liu*, Guiping Zhao, Lu Liu, Martien AM Groenen, et al. RNA-Seq Analysis Reveals Hub Genes Involved in Chicken Intramuscular Fat and Abdominal Fat Deposition During Development. *Frontiers in genetics* (2020).

Liu, Ranran*, **Siyuan Xing***, Jie Wang, Maiqing Zheng, Huanxian Cui, et al. A New Chicken 55k SNP Genotyping Array. *BMC genomics* (2019).

Jinghui Li, **Siyuan Xing**, Guiping Zhao, Maiqing Zheng, Xinting Yang, et al. Identification of diverse cell populations in skeletal muscles and biomarkers for intramuscular fat of chicken by single-cell RNA sequencing. *BMC Genomics* (2020). (Supervised MSc student's paper)

Xiaodong Tan*, Ranran Liu*, **Siyuan Xing**, Yonghong Zhang, Qinghe Li, et al. Genome-Wide Detection of Key Genes and Epigenetic Markers for Chicken Fatty Liver. *International Journal of Molecular Sciences* (2020). (Supervised MSc student's paper)

Lu Liu, Huanxian Cui, **Siyuan Xing**, Guiping Zhao, Jie Wen. Effect of Divergent Selection for Intramuscular Fat Content on Muscle Lipid Metabolism in Chickens. *Animals* (2019).

Liu, Jie, Ranan Liu, Jie Wang, Yonghong Zhang, **Siyuan Xing**, et al. Exploring Genomic Variants Related to Residual Feed Intake in Local and Commercial Chickens by Whole Genomic Resequencing. *Genes* (2018).

M. Jiang*, W. L. Fan*, **S. Y. Xing**, J. Wang, P. Li, et al. Effects of Balanced Selection for Intramuscular Fat and Abdominal Fat Percentage and Estimates of Genetic Parameters. *Poultry Science* (2017).

Siyuan Xing, Tong'ao Yang, Qiang Guo, Fuhe Yang. Correlation Analysis of TYR Gene Mutation and Coat Color of Mustela vison. *Journal of Jilin Agricultural University*. In Chinese (2015).

Zongyue Liu, **Siyuan Xing**, Dawei Hu, Shuming Wang, Xiumei Xing, et al. Estimation of Genetic Parameter for Growth and Fur Trait in Silver Blue Mink. *China Animal Husbandry & Veterinary Medicine*. In Chinese. (2013)

J Wen, R Liu, G Zhao, M Zheng, Q Li, H Cui, **S Xing**. Chicken whole-genome snp chip and use thereof. US Patent App. 16/309,763 (Executed 2020)

This project falls under the “WU-CAAS Cooperative PhD Programme” that is funded by Chinese Academy of Agricultural of Sciences (CAAS) and Wageningen University (WU). The design of the study and collection of data and interpretation of data were supported by the Chinese National Key Research and Development Program of China (No. 2018YFD0500401), the earmarked fund for the modern agroindustry technology research system (CARS-41), and the National Natural Science Foundation of China (31772591).

Siyuan Xing was sponsored by the Institute of Animal Sciences of Chinese Academy of Agricultural Sciences (CAAS) and the China Scholarship Council (Grant Number: 201903250098), as well as personally by Associate Prof. Xinhong Xie and Mrs. Zhang Qiong.

Cover design: Siyuan Xing

Printed by Digiforce | ProefschriftMaken