FROM HIGH-THROUGHPUT TO HYPOTHESES

AN INTEGRATED, DATA-DRIVEN APPROACH TO CHARACTERIZING HEPATIC DEVELOPMENT AND METABOLIC REPROGRAMMING IN THE POST-HATCH BROILER CHICK

by

Heidi Allison Van Every

A dissertation submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Bioinformatics and Data Science

Spring 2021

FROM HIGH-THROUGHPUT TO HYPOTHESES

AN INTEGRATED, DATA-DRIVEN APPROACH TO CHARACTERIZING HEPATIC DEVELOPMENT AND METABOLIC REPROGRAMMING IN THE POST-HATCH BROILER CHICK

by

Heidi Allison Van Every

Approved: _____

Erin E. Connor, Ph.D.
Chair of the Department of Animal and Food Sciences

Approved: _____

Calvin L. Keeler Jr., Ph.D.
Interim Dean of the College of Agriculture and Natural Resources

Approved: _____

Louis F. Rossi, Ph.D.
Vice Provost for Graduate and Professional Education and
Dean of the Graduate College

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Carl J. Schmidt, Ph.D.
Professor in charge of dissertation

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Abhyudai Singh, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Behnam Abasht, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Darko Stefanovski, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Shawn W. Polson, Ph.D.
Member of dissertation committee

ACKNOWLEDGMENTS

To the countless friends and family, both human and animal, who have patiently and unfailingly supported me through this endeavor; and especially to my Gram, who always believed in me, and would have loved to witness its completion.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Artificial selection of modern meat-producing chickens (broilers) for production characteristics has led to dramatic changes in phenotype, yet the impact of this selection on metabolic and molecular mechanisms is poorly understood. The first three weeks post-hatch represent a critical period of adjustment, during which the yolk lipid is depleted and the bird transitions to reliance on a carbohydrate-rich diet. As the liver is the major organ involved in macronutrient metabolism and nutrient allocation, an approach exploiting two types of high-throughput data (transcriptomics and metabolomics) has been implemented to characterize its development, and molecular fluctuations occurring over this important transition.

- First, with an emphasis on metabolic reprogramming, we compared levels of transcripts and primary metabolites at Day 4 and Day 20 post-hatch, establishing differences in core metabolism. For example, at Day 4, metabolic flexibility allows for efficiency to meet the demands of rapid liver growth under oxygen-limiting conditions. At Day 20, the liver's metabolism has shifted to process a carbohydrate-rich diet that supports the rapid overall growth of the modern broiler.
- Secondly, we have used an integrated correlation network approach to identify clusters of genes, liver metabolites, and plasma metabolites showing synchronized patterns of abundance over the complete time course, and additionally relating these groups to blood chemistry measurements indicative of metabolic changes. This work exposed

more nuanced trends occurring over the time course, and also provided insight into how changes in the liver translate to the system as a whole.

- Lastly, we have developed a dynamic workflow to efficiently screen the entire dataset of over 30 thousand features for relation to a response variable (normalized liver mass), computationally reduce the pool of candidates based on significant association, and confirm supportive evidence for these relationships, culminating in novel and statistically significant predictions. This workflow is simple, scalable, and broadly applicable as a means of data exploration, feature prioritization, and hypothesis generation.

Characterizing the changes associated with normal post-hatch hepatic development in the broiler chicken has generated testable hypotheses about the involvement of specific genes and metabolites, clarified the importance of hypoxia to rapid organ growth, and contributed to our understanding of the molecular changes affected by decades of artificial selection.

Chapter 1

INTRODUCTION AND BACKGROUND

The modern broiler (meat) chicken is the product of more than 60 years of artificial selection for commercially desirable traits, resulting in both improved feed efficiency and breast muscle yield. The modern broiler reaches market weight in ¾ the time it took in the 1950s, yet weighs nearly twice as much as the 1950s breeds, with the breast muscle representing a greater component of the overall bird mass (1). Several studies have compared modern lines with unselected lines in terms of growth rate and feed efficiency (2,3). In one such study comparing growth of a modern broiler line (Ross 708) with a legacy line of commercial general-purpose bird unselected since the 1950s (UIUC) over the first five weeks post hatch, the breast muscle was found to comprise 18% and 9% of total body mass, respectively (4). Relative heart weight was markedly decreased in the modern broiler line, suggesting a possible difference in regulation of resource distribution between the two muscle types. The feed efficiency of the modern broiler was also significantly higher than that of the legacy line throughout the study. The legacy line is a commercial meat bird unselected since the 1950s and maintained by the University of Illinois, providing a prime comparison illustrating the effects of modern breeding efforts.

One of the physiological systems undergoing selection in the modern broiler chicken and allowing for increased feed efficiency and growth is the gastrointestinal system. The jejunal and ileal segments of the small intestine are 20% longer in the modern line compared to unselected line, presumably allowing for improved nutrient

absorption (4). Additional changes in growth pattern manifest in the liver allometry. In both lines, the relative liver mass reached a similar maximum of approximately 3.8% of body mass and then began declining. However, this peak occurred a week earlier in the modern broiler. It was hypothesized that this may signify a critical biological time point in the bird's development, the earlier onset of which gives the modern broiler a metabolic advantage.

The liver plays critical roles in macronutrient metabolism. The essential cooperation between the liver and pancreas maintains homeostatic blood glucose levels. The liver stores excess glucose as glycogen, and can mobilize these reserves when required as well as synthesize glucose from alternate sources through gluconeogenesis. In birds, the liver is the main site of lipogenesis. Broilers are fed a carbohydrate-rich diet, so they must perform fatty acid synthesis to meet most of their requirements (5). Although dietary lipids are absorbed in the small intestine, they are not transported through the lymphatic system as with mammals and instead travel via the portal blood system through the liver before circulating through the body. Hepatocytes also produce and secrete bile salts, necessary for facilitation of fat digestion, into separate bile ducts, although in birds this capability may not be fully developed until after the first week of life (6). The liver also participates in protein metabolism, catabolizing excess amino acids and performing 11% of the body's protein synthesis (7). The liver is also responsible for regulating blood volume and homeostasis, synthesizing blood plasma and immunity-related proteins. As a direct recipient of blood from the digestive tract, it is one of the first lines of defense against pathogens absorbed through the intestine. Liver-specific macrophages known as

Kupffer cells, present in the sinusoids, phagocytize foreign microorganisms and senescent blood cells.

Liver receives oxygenated blood through the hepatic artery, which accounts for only 25% of blood flow in the liver. The remaining 75% consists of deoxygenated blood delivered through the hepatic portal vein directly from the digestive tract. Mitochondria-rich hepatocytes are the main cell type in the liver, comprising 80% of the organ's volume. They are arranged in a lobular structure, surrounding sinusoids through which blood flows into the central vein, then out of the liver through the hepatic vein and to the inferior vena cava for dissemination throughout the body (8–10). In this way liver metabolites, hormones, immune factors, and blood proteins are directly secreted back into the bloodstream. Possessing both endocrine and exocrine functionality, the liver is a critical hormonal regulator of digestion, feed intake, and body growth as part of the somatotropic axis (9). Voluntary feed intake is an important factor in body growth, and modern broilers are voracious eaters compared to unselected or low-weight-selected breeds, such as egg layers. This has inspired studies directly stimulating the liver with compounds known to increase or decrease appetite. Differences have been noted when comparing response to appetite-stimulating and appetite-decreasing hormones and compounds between high-weight-selected and low-weight-selected birds, indicating there are likely perturbations in the physiology of the modern broiler's endocrine system (7). Consequently, studying the development of the liver in the modern broiler can improve our understanding of the metabolic pathways and regulatory mechanisms contributing to early post-hatch growth.

High-throughput transcriptome analyses (RNA-seq) provide snapshots of transcribed RNAs at any given time and are useful to identify differentially regulated

genes between conditions or time points. Additionally, untargeted metabolomics can identify hundreds of compounds present, indicating precise metabolic differences. Combining these two data types is a powerful means to infer hypotheses about the interactions between the transcriptome and metabolome. For example, integrating these two high throughput methods identified metabolic and signaling pathways responding to heat stress in the liver of modern broilers (11).

Previous studies have described the hepatic transcriptome of the modern broiler, either in response to different treatments, or in comparison with other bird lines. The majority of differential expression studies of broiler chickens involves the effect of treatments such as dietary supplementation (12,13) or temperature conditions (11,14,15). A recent study compared the hepatic transcriptome over six time points during the embryo to hatchling transition, from 16-day embryos to 9 day old chicks. They identified many metabolic pathways consistent with the nutrient source transition the chicks undergo in the first week post hatch, especially some affecting lipid metabolism (16). Another study examined changes in the hepatic transcriptome resulting from immediate post-hatch fasting and re-feeding, identifying genes regulated by lipogenic transcription factor THRSPA and switching between lipolytic and lipogenic states (17). One work also compared liver transcriptome in three chicken lines with differing growth potential at day 14 post hatch. They chose a high-weight-selected (broiler), low-weight-selected (layer), and a mid-weight cross, focusing on genes differentially expressed between all three lines, and identifying significantly enriched pathways. They found that genes in the FoxO signaling pathway were highly correlated with body weight, likely owing to their effects on cell cycle and metabolism (18).

Decades of genetic selection for different purposes also make the chicken an excellent model for exploring the molecular foundations of metabolic efficiency and energy partitioning. Although mechanisms conferring a metabolic advantage are well-studied in abnormal systems such as cancer, studies investigating rapid growth in functional, non-malignant tissues are lacking. Because of its central role in metabolism, its ability to regulate a number of systemic factors influencing body and organ growth, and notable differences in allometry and physiological responses previously observed between modern broilers and other lines of chicken, the liver is a prime tissue to study when investigating how the modern broiler achieves its great growth potential. Due to the allometric relationships observed by our group (4), along with what is known about concurrent physiological changes, the first 20 days post-hatch was chosen for this study as representing a crucial biological time period in the growth of the modern broiler chicken. Nine time points, from Day 4 through 20 post-hatch, are represented, with liver transcriptome, liver metabolome, and primary metabolome data for each bird, along with morphometric measurements and i-STAT blood chemistry values.

There have been no integrated high-throughput studies of the modern broiler liver under normal conditions in the critical first three weeks post-hatch. Thus, the molecular changes that are occurring during this time period – the metabolic drivers of rapid muscle growth and feed efficiency – are poorly understood. Exploring these in a data-driven fashion can elucidate new knowledge about the liver's most important functions during early post-hatch growth of the chick, and also how the liver itself is developing. The overall goal of this study is to characterize the liver transcriptome and

metabolome as it develops over the first 20 days post hatch in a modern broiler chicken.

Chapter 2

TRANSCRIPTOMIC AND METABOLOMIC CHARACTERIZATION OF POST-HATCH METABOLIC REPROGRAMMING DURING HEPATIC DEVELOPMENT IN THE CHICKEN

2.1 Background

The entire longitudinal dataset explored in this work is comprised of nine time points, bracketing the time period where relative liver mass peaks in the modern broiler, and also capturing a critical metabolic transition. Chicks undergo drastic physiological changes as a consequence of hatching. The developing embryo relies entirely on nutrients from the yolk (19–21). During late embryonic development, much of the yolk lipid is absorbed and stored in the liver, predominately as cholesteryl esters (22). At day 18 of incubation, three days prior to hatch, lipids make up 10% of the liver's mass due to absorption and storage of yolk nutrients (23). This stored lipid, along with the yolk remnant, provides the chick with a nutrients following hatch, but by day 5 post-hatch 90% of the yolk lipid has been absorbed (24). Chicks are provided with a carbohydrate-rich diet at hatch because fasting during this period stunts the early muscle growth potential of chicks (25). These early changes in nutrient source, coupled with rapid growth, mean maintaining metabolic homeorhesis is a major challenge facing the liver in the early weeks following hatch.

The first objective of this work was to provide a starting point, and focus the direction of subsequent analyses, with an emphasis on metabolism. We executed a comparison between Day 4 and Day 20 post-hatch – the two time points expected to

show the greatest phenotypic and metabolic differences. We expected to see the largest differences in carbohydrate and lipid metabolism reflecting the yolk to diet transition. Through differential expression of the liver transcriptome, and a t-test and log fold change comparison of the liver primary metabolome, we identified genes and metabolites that differed in abundance between time points. Transcriptome and metabolome were analyzed separately, then integrated at the pathway level with a focus on core metabolic pathways including carbohydrate, lipid, and amino acid metabolism. These were selected to capture the metabolic reprogramming required to support the transition from relying on stored yolk to orally consumed feed that underlies the growth rate and phenotype of the modern broiler.

2.2 Results

2.2.1 Phenotypic measurements and i-STAT blood chemistry

At D4 post-hatch, the liver was noticeably yellow in color, gradually changing to deep red by D20 (Figure 1). Mean phenotypic measurements of bird growth, liver allometry, and i-STAT blood chemistry values are shown in Table 1; Figure 2 shows hierarchical clustering of this data, which separates the two groups by age. Body mass and liver mass showed the largest difference between days and were positively correlated with bird age (Pearson Correlation Coefficient (PCC) 0.98 and 0.97, respectively). Relative liver mass was negatively correlated with bird age (PCC -0.51). The top blood chemistry values positively correlated with bird age were sodium (Na, PCC 0.89), bicarbonate (HCO_3, PCC 0.79), total carbon dioxide (TCO_2, PCC 0.77), and pH (PCC 0.75). Partial oxygen (PO_2, PCC -0.70) and oxygen saturation (sO_2, -0.56) were negatively correlated with bird age.

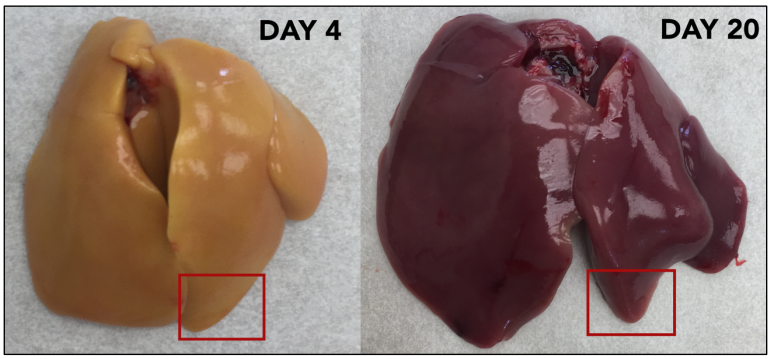

Figure 1: Contrast in liver color at D4 and D20 post-hatch. The yellow color at hatch is indicative of the absorption and storage of yolk lipid and nutrients that occurs during late embryonic development. The liver gradually changes to deep red as the chick grows, concurrent with the depletion of the liver's stores. Tissue was routinely sampled from the lower left lobe, as indicated by the red boxes. Note: Liver sizes are not on the same scale.

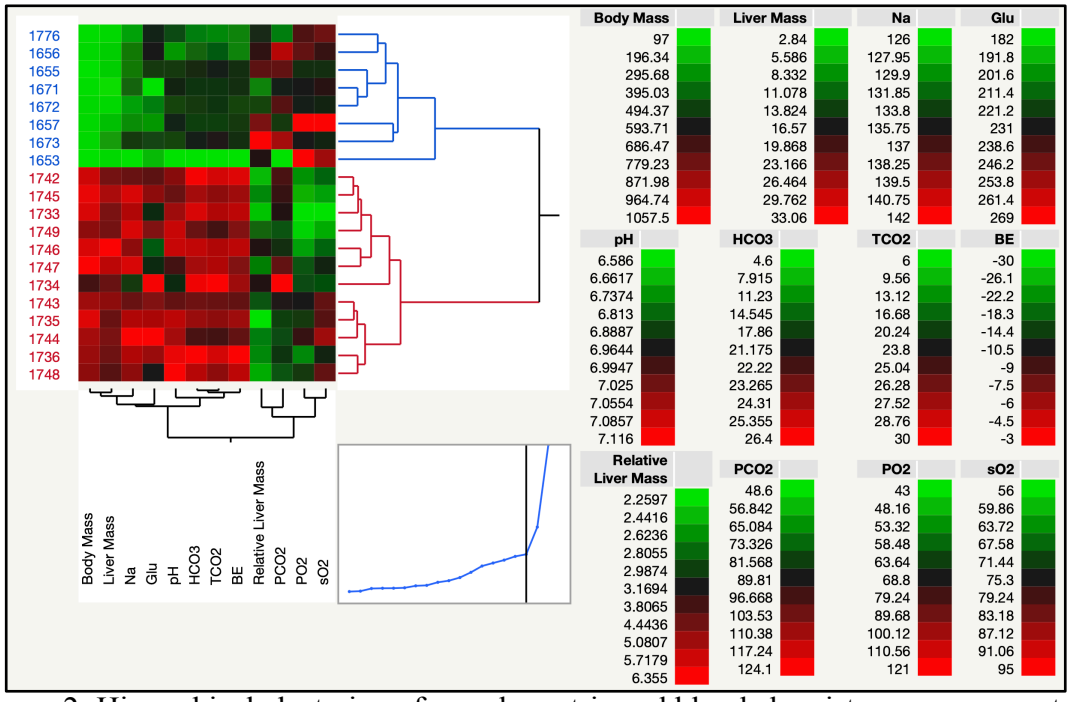

Figure 2: Hierarchical clustering of morphometric and blood chemistry measurements from all birds. There were no i-STAT readings from three D4 birds, and all D20 birds are included regardless of quality elimination from transcriptome analysis.

Table 1: Summary of phenotypic trait and blood gas values by day, along with published references for comparison. * Denotes Wilcoxon test was used instead of t-test.

	Median		**Mean ± Standard Deviation**		**p value**	**Trend with age**	**Adult Breeder Values (26)**	
	D4	**D20**	**D4**	**D20**			Range	Mean
Body Mass (g)	112.25	987.50	110.75 ± 5.54	912.64 ± 134.3	<0.0001	+	NA	NA
Liver Mass (g)	3.77	23.35	4.29 ± 1.43	25.01 ± 4.28	<0.0001	+	NA	NA
Normaliz-ed Liver Mass (%)	0.034	0.027	0.039 ± 0.012	0.028 ± 0.004	0.0214	-	NA	NA
pH	6.88	7.08	6.83 ± 0.13	7.05 ± 0.06	0.0034	+	7.28 - 7.57	7.42
PCO_2 (mm Hg)	87.70	84.40	87.47 ± 23.51	91.03 ± 16.91	0.7513	NA	25.9 - 49.5	37.7
PO_2 (mm Hg)*	82.00	61.00	88.29 ± 23.45	55.71 ± 9.16	0.0021	-	32.0 - 60.5	46.2
HCO_3 (mmol/L)	17.90	24.80	15.39 ± 5.49	24.77 ± 1.05	0.0037	+	18.9 - 30.3	24.6
Base Excess	-14.50	-6.00	-16.17 ± 3.82	-5.86 ± 0.9	0.0031	+	-6.8 - 7.2	0.2
sO_2 (%)	78.00	70.00	81.14 ± 7.73	70.29 ± 9.67	0.0398	-	70.6 - 93.3	82
Glu (mg/dL)	206.00	230.00	208.57 ± 18.79	238.86 ± 19.04	0.0112	+	207.2 - 260.7	234
TCO_2	20.00	28.00	17.71 ± 6.02	27.57 ± 1.51	0.0044	+	19.9 - 31.5	25.7
Na (mmol/L) *	130.00	140.00	130.14 ± 2.54	139.14 ± 2.41	0.0025	+	141.6 - 152.6	147.1

TCO_2, PCO_2, HCO_3, and pH are used to assess blood acid-base balance, which is maintained by the kidneys and lungs and affected by both metabolism and respiration. TCO_2 is a measure of total blood carbon dioxide while PCO_2 measures the difference between CO_2 produced by the cells and removed through respiration. HCO_3 is a blood buffer produced by the kidneys, representing the metabolic component of acid-base balance. Given a change in blood pH due to any of these values, BE can help to differentiate between respiratory or metabolic causes. It is calculated as the difference between titratable base and titratable acid, and not susceptible to respiratory factors such as changes in PCO_2. An increase in pH was observed from D4 to D20, indicating a shift in acid-base balance as the birds age. The metabolic measures of acid-base balance (buffer HCO_3 and BE) were increased from D4 to D20, while the respiratory component was unchanged (PCO_2), indicating the shift in acid-base balance is most likely due to metabolic factors.

2.2.2 Transcriptome analysis: Top 100 abundant transcripts from each day

Examination of the 100 most abundant transcripts expressed in either the D4 or D20 liver (total of 200) identify important similarities in metabolic functions at these two time points. Of these genes, 89 were common between both D4 and D20. Enriched Gene Ontology (GO) terms among these common genes included Translation, encompassing 14 ribosomal proteins and Secretory Vesicle, which included albumin along with proteins involved in lipid transport, complement and coagulation. Two other enriched GO terms shared by both days were Mitochondria and Oxidative Phosphorylation. These terms were enriched by genes encoding mitochondrial rRNAs and tRNAs along with NADH dehydrogenases, cytochrome oxidases and ATP synthase subunits. One gene product unique to D20 encodes

glucose 6-phosphatase (G6PC) an enzyme critical to gluconeogenesis. Several transcripts encoding genes affecting additional processes were found in the D4 top 100 list that were not in that D20 list. These include proteins involved in lipid metabolism and transport, amino acid catabolism, peptidase inhibitors, a sulfotransferase and hemoglobin A. These results indicate that, despite the changes undergone by the liver from D4 to D20, the major hepatic functions are preserved between time points.

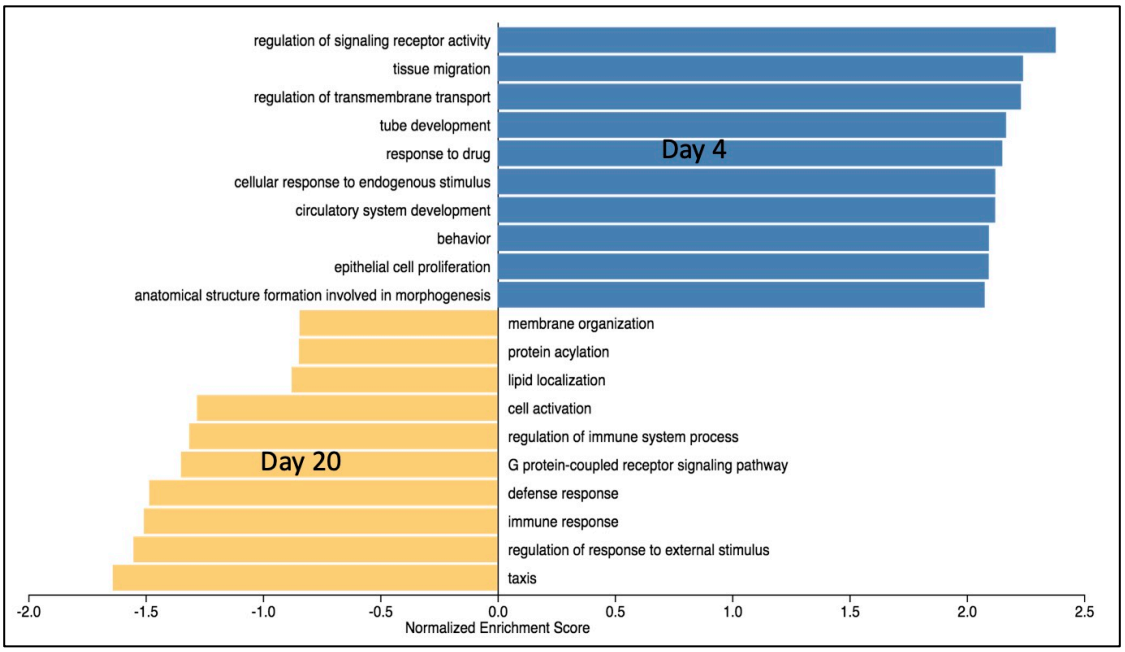

Figure 3: Gene Ontology Biological Process Terms enriched at either Day 4 (blue) or Day 20 (gold).

Ontology enrichment analysis using DAVID (27,28) showed distinct differences between time points (Figure 3). At D4, top Functional Annotation Clusters were related to a variety of cell cycle elements including mitosis, cell division,

centromeric chromosome condensation & segregation, DNA replication, and transitions between cell cycle phases. Other clusters contained terms involved in ribonucleotide binding, kinase activity, amino-acid modification, vasculature development, and migration and motility of epithelial cells. At D4, the top enriched KEGG pathway from STRING (21,22) was "Cell Cycle," with 36 out of 123 proteins represented. DNA replication and cellular senescence were also among the top ten. Purine and Pyrimidine metabolism was the only metabolic pathway enriched by the transcriptome at D4. At D20, top Functional Annotation Clusters were related to immune response, including T cell and B cell receptor signaling pathways, toll-like receptor signaling pathway, immune cell aggregation, activation, proliferation, and differentiation. One cluster contained terms related to oxidoreductase activity including heme binding and cytochrome P450. The top enriched KEGG pathway from STRING was "Metabolic Pathways," with 162 out of 1250 proteins represented. Other enriched pathways were related to carbohydrate metabolism, including fructose and mannose, and galactose, and immune-related pathway Th17 cell differentiation. Ontology and pathway analysis of the transcriptome gave the first glimpse of the major processes important to the liver at each time point: rapid organ growth and vasculature development at D4; carbohydrate metabolism and immune cell population expansion at D20.

2.2.3 Hypoxic environment at D4

Early in the process of investigating the data, it was noticed that HIF1A transcripts were elevated in the D4 liver (log2 fold change 0.56, adjusted p-value 0.03), suggesting the tissue is under hypoxic conditions. To further evaluate this

possibility, a list of human genes induced under hypoxic conditions was downloaded from the Gene Set Enrichment Analysis resource (29,30) and used to extract the orthologs from the D4 and D20 expression data. Principal component analysis revealed that 43% of the variance was associated with the day post-hatch; with the D4 samples showing elevated levels of many of the transcripts associated with hypoxia (Figure 4).

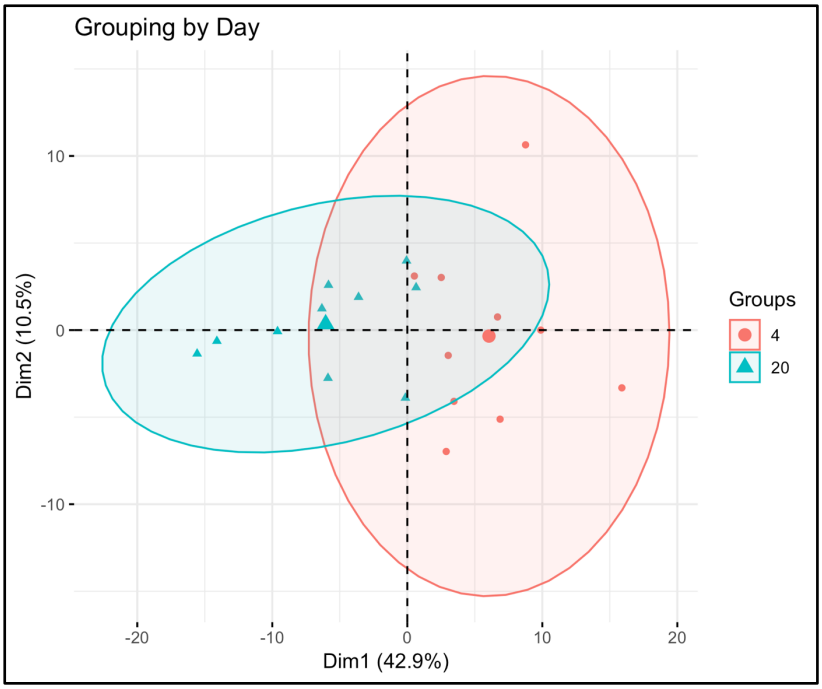

Figure 4: PCA of hypoxia genes showing clear separation by day along Dimension 1.

2.2.4 Metabolome analysis: PCA, random forest, and top significant metabolites

Principal component analysis of metabolites separated D4 birds from D20 birds (Figure 5A), and random forest also correctly classified birds by age group. The top compounds contributing to random forest classification are shown in Figure 5B. The top identified compounds contributing to random forest classification included

two more abundant at D4 (lysine, glutaric acid) and seven more abundant at D20 (CMP, fumaric acid, fructose-6-phosphate, fucose, malic acid, glucose-6-phosphate, succinic acid). Lysine is an essential amino acid important for growth, and glutaric acid is a byproduct of amino acid metabolism. Fumaric acid, malic acid, and succinic acid are TCA cycle intermediates, while fructose-6-phosphate, glucose-6-phosphate, and fucose are sugars involved in glycolysis and other carbohydrate metabolic pathways. CMP (Cytidine monophosphate), is a pyrimidine-derived nucleotide.

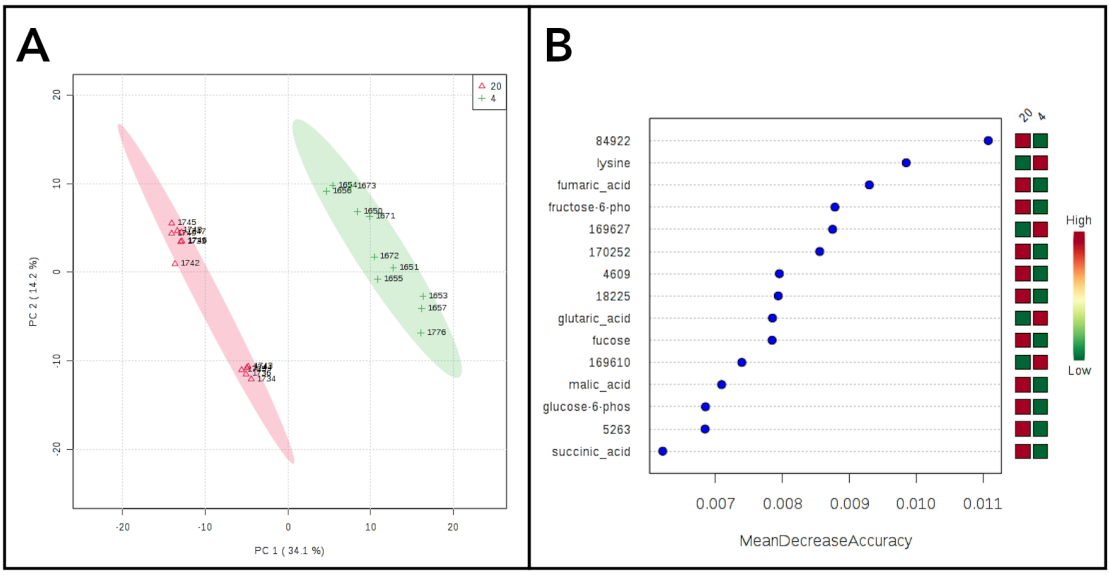

Figure 5: A: PCA showing clear separation of individuals by top metabolites. D4 = green, D20 = red. B: Top metabolites contributing to random forest classification that correctly separated D4 and D20. Compound 84922 was identified by PubChem ID as cytidylic acid (CMP).

Table 2: Top significant identified metabolites with pathway membership or role in metabolism. Lipid and amino acid metabolism-related compounds predominated in D4, while many of those present in D20 were involved in carbohydrate metabolism.

Compound	Fold-change (Log2)	Adjusted p-value	Day	Pathway
Retinal	4.42	2.22E-10	D4	Vitamin A
2-Hydroxybutanoic Acid	3.34	4.25E-03	D4	Amino Acid-Glutathione Metabolism
Oleic acid	3.12	2.41E-3	D4	Lipid metabolism
Palmitoleic acid	2.87	5.98E-9	D4	Lipid metabolism
Lactobiose (lactose)	2.60	3.11E-5	D4	Carbohydrate metabolism
Phosphoserine	1.99	6.92E-5	D4	Serine metabolism
Uric Acid	1.79	1.03E-6	D4	Nitrogen metabolism
Phosphoenolpyruvate	1.77	1.11E-4	D4	Glycolysis (ATP synthesis phase)
Gamma-Tocopherol	1.73	2.73E-4	D4	Vitamin E metabolism
Uracil	1.66	1.59E-6	D4	Pyrimidine metabolism
3-Phosphoglycerate	1.63	3.67E-3	D4	Glycolysis (ATP synthesis phase)
Aspartate	-1.57	1.6E-4	D20	Amino acid metabolism
Adenosine	-1.64	5.7E-3	D20	Purine metabolism
Guanosine	-1.65	4.3E-3	D20	Purine metabolism
Hypoxanthine	-1.74	3.13E-4	D20	Purine metabolism
Creatinine	-1.82	7.65E-3	D20	Creatine metabolism
Citrate	-2.00	3.41E-5	D20	TCA cycle
Fructose-6-Phosphate	-2.15	4.86E-7	D20	Gluconeogenesis or Glycolysis (ATP-incorporating phase)
CMP	-2.42	2.53E-7	D20	Pyrimidine metabolism, TAG, lipid & sialic acid synthesis
Inosine	-2.66	1.21E-5	D20	Nucleoside metabolism
5-Methoxytryptamine	-2.81	9.43E-7	D20	Tryptophan metabolism
Hexose-6-Phosphate	-3.25	5.51E-8	D20	Carbohydrate metabolism
Succinate	-3.35	2.64E-9	D20	TCA cycle
Glucose-6-Phosphate	-3.48	1.44E-6	D20	Gluconeogenesis or Glycolysis (ATP-incorporating phase)
Fumarate	-4.87	1.9E-9	D20	TCA cycle
Malate	-5.32	9.28E-12	D20	TCA cycle

By t-test, 90 compounds were more abundant at D4 and 112 at D20. A volcano plot was used to visualize top metabolites by log2 fold change and adjusted p-value (Figure 6); some of these are detailed in Table 2. At D4, several of the top significant metabolites were yolk-derived nutrients and fatty acids including retinal, oleic acid, palmitoleic acid, and gamma-tocopherol (Vitamin E). Retinal, a retinoid derived from known egg yolk nutrient Vitamin A, is critical in numerous processes including growth regulation and lipid metabolism (31). The second most significant compound, 2-hydroxybutanoic acid, can be produced as a byproduct of threonine catabolism and glutathione synthesis, and is also part of propanoate metabolism (32). Lactobiose (lactose), a common chicken feed additive, is a disaccharide of glucose and galactose. Phosphoserine is an intermediate of amino acid metabolism, and uric acid is the major waste product of protein catabolism in birds. Phosphoenolpyruvate and 3-phosphoglycerate are intermediates of glycolysis that are also involved in several other metabolic pathways including the TCA cycle and lipid metabolism. Phosphoenolpyruvate can be generated from TCA cycle intermediate oxaloacetate and may reflect utilization of alternative carbon sources. Uracil is an RNA pyrimidine nucleobase. In the liver, as UDP-glucose, it has roles in carbohydrate metabolism where it regulates the conversion of glucose to galactose (33).

In D20, several of the most significant identified metabolites were intermediates of the TCA cycle (malic acid, fumaric acid, succinic acid, citric acid), or sugars involved in carbohydrate metabolism (glucose-6-phosphate, hexose-6-phosphate, fructose-6-phosphate). Adenosine, guanosine, and inosine are nucleosides. CMP and hypoxanthine are also part of purine and pyrimidine metabolism. 5-methoxytryptamine is derived from serotonin, a neurotransmitter derived from

tryptophan. Creatinine is a waste product of amino acid catabolism in muscle. Aspartate is a non-essential amino acid.

Metabolome results implicate lipids, vitamin A, vitamin E, carbohydrate, serine, cysteine, uric acid and uracil metabolism as metabolic characteristics of day 4 post-hatch liver. In contrast D20 metabolome data implicate the TCA, gluconeogenesis (or glycolysis) pathways along with aspartate, tryptophan, creatine, purine, pyrimidine, and inosine metabolism.

2.2.5 Metabolic Pathway-Level Integration of Transcriptome and Metabolome:

2.2.5.1 Carbohydrate metabolism

Central carbohydrate metabolism consists of glycolysis, gluconeogenesis, the tricarboxylic acid (TCA) cycle, and the pentose phosphate pathway (PPP) (Figure 7). Glycolysis consists of two stages: 1) Conversion of free glucose to two triose phosphates, 2) energy generation through production of pyruvate. The integrated data suggests that, at D4, the glycolysis pathway is enriched at the second, ATP-generating stage. The transcript encoding the platelet isoform of PFK, the rate limiting enzyme responsible for conversion of fructose-6-phosphate to fructose-1,6-bisphosphate, was more abundant at D4. This may reflect isozyme selection by HIF1A to increase efficiency of this pathway under hypoxic conditions. Furthermore, two intermediate metabolites (3-PG, PEP), and transcripts encoding two enzymes from the third stage of glycolysis (BPGM, PDHA1) were also enriched in the D4 samples. The enzyme BPGM and metabolite 3-PG represents a branching point in glycolysis. In the glycolysis pathway BPGM acts as a mutase, and regulates the entry of 3-PG into either glycolysis or serine biosynthesis through its effects on PGAM1. The product of BPGM enzymatic activity, 2,3 bisphosphoglycerate (2,3 BPG) serves as a phosphate donor to activate PGAM and promote glycolysis. LDHA, an enzyme involved in

anaerobic ATP production, was upregulated at D4, in addition to transporters responsible for both import and export of lactate (SLC16A3, SLC5A12). LDHA favors the conversion of pyruvate to lactate and regenerates the NAD+ required by the glycolytic glyceraldehyde-3-phosphate dehydrogenase (GAPDH). All of these D4 enriched molecules may be critical to supporting production of liver ATP via glycolysis under hypoxic conditions during this early stage post-hatch.

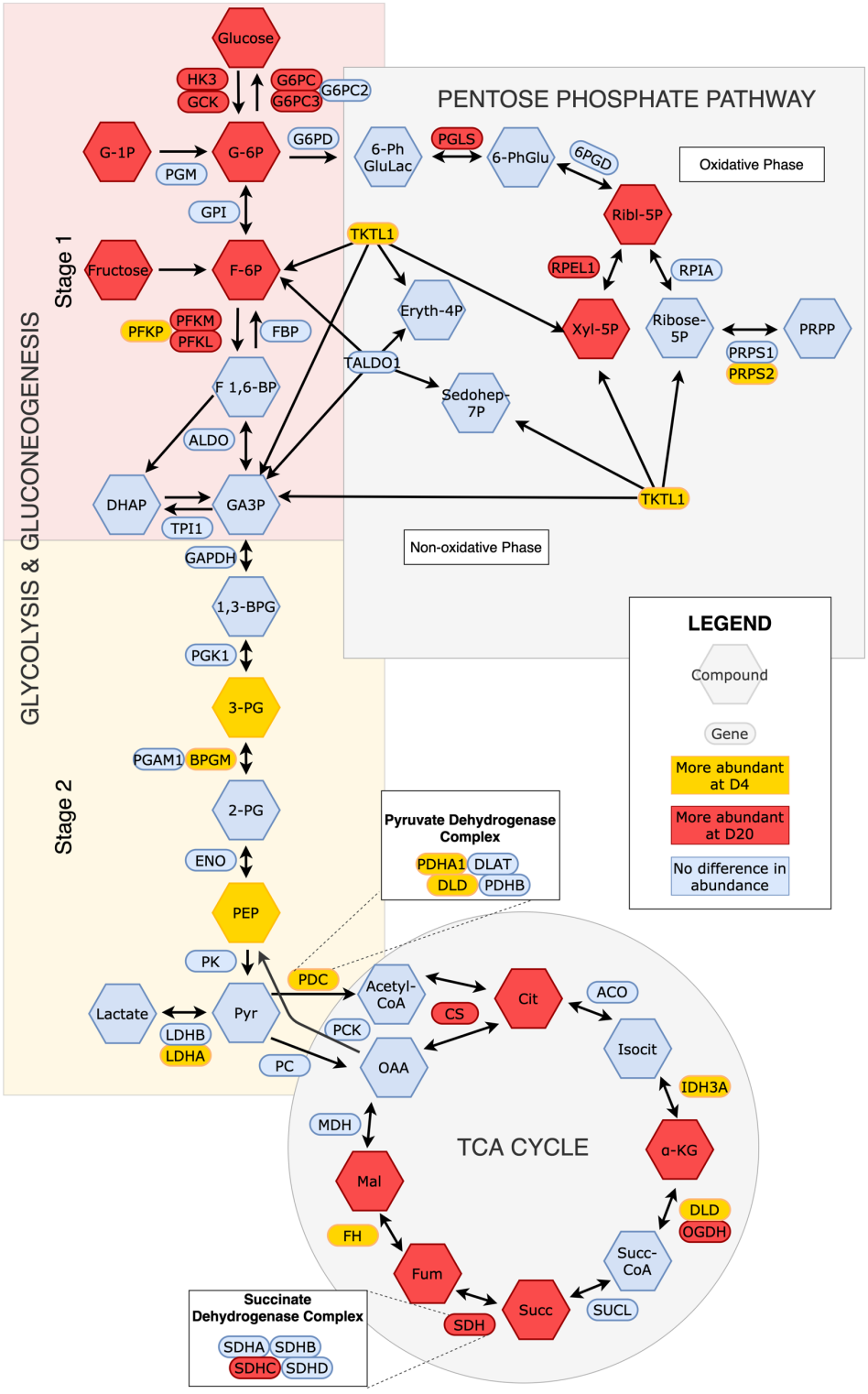
GLYCOLYSIS & GLUCONEOGENESIS
Stage 1
Stage 2
Glucose
HK3
GCK
G6PC
G6PC3
G6PC2
G-1P
PGM
G-6P
G6PD
GPI
Fructose
F-6P
PFKP
PFKM
PFKL
FBP
F 1,6-BP
ALDO
DHAP
GA3P
TPI1
GAPDH
1,3-BPG
PGK1
3-PG
PGAM1
BPGM
2-PG
ENO
PEP
PK
Lactate
LDHB
LDHA
Pyr
PDC
PC
PCK
Acetyl-CoA
CS
Cit
ACO
Isocit
IDH3A
α-KG
DLD
OGDH
Succ-CoA
SUCL
Succ
SDH
Fum
FH
Mal
MDH
OAA
TCA CYCLE
PENTOSE PHOSPHATE PATHWAY
Oxidative Phase
6-Ph GluLac
PGLS
6-PhGlu
6PGD
Ribl-5P
RPEL1
RPIA
Xyl-5P
Ribose-5P
PRPP
PRPS1
PRPS2
TKTL1
TALDO1
Eryth-4P
Sedohep-7P
Non-oxidative Phase
Pyruvate Dehydrogenase Complex
PDHA1
DLAT
DLD
PDHB
Succinate Dehydrogenase Complex
SDHA
SDHB
SDHC
SDHD
LEGEND
Compound
Gene
More abundant at D4
More abundant at D20
No difference in abundance

Figure 6: Core carbohydrate metabolism including glycolysis & gluconeogenesis, the TCA cycle, and the pentose phosphate pathway. Genes and metabolites that differed in abundance between days are highlighted, with abbreviations as follows: 1,3-BPG – 1,3-bisphosphoglycerate; 2-PG – 2-phosphoglycerate; 3-PG – 3-phosphoglycerate; 6-PhGluLac – 6-phosphogluconolactone; 6-PhGlu – 6-phosphogluconate; α-KG – α-ketoglutarate; BPGM – bisphosphoglycerate mutase; Cit – citrate; CS – citrate synthase; DHAP –dihydroxyacetone phosphate; DLD - dihydrolipoamide dehydrogenase; Eryth-4P – erythrose-4-phosphate; F-6P – fructose-6-phosphate; F 1,6-BP – fructose-1,6-bisphosphate; Fum – fumarate; G-1P – glucose-1-phosphate; GA3P – glyceraldehyde-3-phosphate; GCK – glucokinase; G6PC – glucose-6-phosphatase catalytic; G6PC3 – glucose-6-phosphatase catalytic subunit 3; G-6P – glucose-6-phosphate; HK3 – hexokinase 3; IDH3A – isocitrate dehydrogenase 3 alpha; Isocit – isocitrate; LDHA – lactate dehydrogenase A; Mal – malate; OAA – oxaloacetate; PDHA1 – pyruvate dehydrogenase E1 subunit alpha 1; PEP – phosphoenolpyruvate; PFKM – phosphofructokinase, muscle; PFKL – phosphofructokinase, liver; PFKP – phosphofructokinase, platelet; PGLS – 6-phosphogluconolactonase; PRPP – phosphoribosyl pyrophosphate; PRPS2 – phosphoribosyl pyrophosphate synthetase 2; Pyr – pyruvate; Ribl-5P – ribulose-5-phosphate; RPEL1 – ribulose-5-phosphate-3-epimerase like 1; Sedohep-7P – sedoheptulose-7-phosphate; SDHC – succinate dehydrogenase complex subunit C; Succ – succinate; Succ-CoA – succinyl-coA; TKTL1 - transketolase like 1; Xyl-5P – xylulose-5-phosphate.

The pyruvate dehydrogenase complex controls the link between glycolysis and the TCA cycle. Transcripts encoding two of the three components of pyruvate dehydrogenase, the E1 subunit (PDHA1) and Dihydrolipoyl dehydrogenase (DLD) were enriched in the D4 liver. In addition, the regulatory kinase PDK1, which inactivates pyruvate dehydrogenase, was also elevated in the D4 samples. The increased abundance of the pyruvate dehydrogenase subunit along with the negative regulatory PDK1 suggests that metabolism at D4 may be primed to respond rapidly to changes in ATP levels and oxygen availability.

Several transcripts encoding rate-limiting sugar kinases involved in the early steps of glycolysis were more abundant at D20 compared with D4 (HK3, GCK, PFKM, PFKL). Corresponding glycolytic metabolites were also more abundant in D20 (glucose, G-6P, F-6P), with G-6P having one of the highest fold changes when compared with D4 (logFC 3.48). HK3 and GCK have key differences in their regulation. GCK specifically acts on glucose, while HK will phosphorylate multiple types of hexoses. GCK also has much lower affinity for glucose than HK, and, unlike HK, GCK is not inhibited by its product, G-6P. Thus, while HK maintains basal glucose metabolism, GCK is responsible for phosphorylating excess glucose for other fates, such as glycogen synthesis or diversion to the pentose phosphate pathway. Phosphofructokinase (PFK) controls glycolytic rate and is under tight control, although there is evidence that isozymes differ in their regulation. Two isoforms of PFK were more abundant at D20 than D4, one of which (liver isoform PFKL) was upregulated in broiler chickens with high growth potential when compared to crosses and layer birds, suggesting that this isoform may contribute to rapid growth rate of maturing birds (28). The increased abundance of these enzymes and metabolites at D20 suggests surplus of free glucose that can be diverted to other metabolic fates or exported from the liver for use by other tissues.

Glycogen metabolism and gluconeogenesis are two pathways the liver uses to provide glucose to other organs during fasting. Typically, the first resource exploited is glycogen. Glycogen can be synthesized by the enzyme glycogen synthase from glucose-1-phosphate (G-1P) and broken down by glycogen phosphorylase to yield G-1P. Glycogen synthase transcripts along with two isoforms of glycogen phosphorylase (PYGL, PYGB), are enriched in the D20 liver. This, combined with the observation

that G-1P is also elevated in the D20 liver, suggests that the D20 liver is capable of rapid response to demands for either glycogen synthesis or phosphorolysis. In addition, the D20 liver is enriched for two glucose-6-phosphatase mRNAs (G6PC, G6PC3), which catalyze the last step of gluconeogenesis. As with glycogen metabolism, it appears that glucose metabolism in the D20 liver is capable of rapid responses to the demands of the body for glucose.

The TCA cycle is an aerobic pathway that continues the oxidation of pyruvate, producing electron donors NADH and $FADH_2$ which will go on to oxidative phosphorylation. Multiple components of the TCA cycle are upregulated at D20, indicating greater oxygen availability and abundance of nutrients. At D20, several intermediate metabolites in the TCA cycle were more abundant (citrate, α-ketoglutarate (α-KG), succinate, fumarate, malate), along with mRNAs encoding three enzymes (CS, ODGH, SDHC). All metabolites but α-KG were also among the top most significant compounds at D20, in terms of both fold change and significance (see Table 2, Figure 6. α-KG, fumarate, and succinate all serve as entry points for catabolized glucogenic amino acids. CS is the rate-limiting enzyme of the TCA cycle. Elevated citrate is an important regulator of metabolism, with high levels signaling abundant energy. Citrate inhibits glycolysis through its action on phosphofructokinase and stimulates fatty acid synthesis.

Components of the TCA cycle are reduced at D4 compared with D20 livers, consistent with response to hypoxic conditions. Regulation of the pyruvate dehydrogenase complex also suggests metabolic flexibility allowing for rapid response to energy and oxygen levels and utilization of alternative carbon sources for critical metabolites. At D4, four TCA-related transcripts were more abundant (PDHA1, DLD,

IDH3A, FH). The rate-limiting pyruvate dehydrogenase complex controls entry of pyruvate into the TCA cycle, and is regulated by several enzymes whose transcripts were also more abundant at D4 (PDP1, PDP2, PDK1). This could represent increased responsiveness of the pyruvate dehydrogenase complex to changes in ATP and oxygen levels. One isozyme of isocitrate dehydrogenase, which interconverts isocitrate and α-KG, was upregulated at D4 (IDH3A). IDH1 and IDH2 can catalyze in both oxidative and reductive directions and are involved in hypoxia response when downregulation of the TCA cycle requires alternate means to synthesize acetyl-CoA and citrate. IDH3A, however, is irreversible and only converts isocitrate to α-KG. IDH3A is also localized to the mitochondria, relies on NAD+ as a cofactor instead of NADP+, and is allosterically regulated by a number of factors. Although hypoxic conditions typically favor conversion of α-KG to isocitrate as an alternative way to generate acetyl-CoA and citrate (29), IDH3A still appears to have a critical role in response to hypoxia. In cancer cells, elevated levels of IDH3A ultimately lead to decreased levels of α-KG. In turn, reduced α-KG levels stabilize the HIF1A protein thereby promoting angiogenesis (30). Conceivably, the IDH3A mechanism documented in cancer cells may play an important role in the normal development of the early post-hatch liver.

The pentose phosphate pathway utilizes glycolytic intermediates to produce NADPH for reducing power and supplies pentoses for nucleotide synthesis. The non-oxidative branch of the PPP is upregulated at D4, consistent with rapid cell proliferation, while the oxidative branch is upregulated at D20, perhaps to meet increased demand for reducing power. At D4, two transcripts encoding two enzymes in the non-oxidative branch of the PPP were upregulated (TKTL1, PRPS2). TKT is the rate-limiting enzyme reversibly linking the PPP with glycolysis. Elevated levels of

TKT could indicate intermediates are being exchanged between pathways. The upregulation of PRPS2 suggests that ribose-5-phosphate generated through the non-oxidative branch is going on to purine and pyrimidine metabolism at D4. In contrast at D20, enzymes (PGLS, RPEL1) and metabolites (ribulose-5P, xylulose-5P) involved in the oxidative phase of the PPP were more abundant. Increased levels of RPEL1 suggests that ribulose-5-phosphate is also being recycled back into glycolysis, prioritizing energy production through complete oxidation of G-6P while concurrently producing NADPH to provide the reducing agent needed for lipid synthesis at D20.

2.2.5.2 Amino acid metabolism

Amino acids are the building blocks of proteins and also serve many important metabolic functions. Several amino acids, their derivatives, and waste products differed in their abundance between days, including nine more abundant at D4 (arginine, lysine, threonine, cysteine, proline, ornithine, phosphoserine, urea, uric acid) and three more abundant at D20 (aspartate, glutamine, creatinine). Of the amino acids more abundant at D4, three were essential (arginine, lysine, threonine) and three non-essential (cysteine, proline, ornithine). Metabolite data was not able to differentiate ornithine from arginine, so we assume that one or both of them were more abundant at D4. Arginine, ornithine, and proline are glucogenic, typically being converted to glutamate that is readily converted to TCA cycle intermediate α-KG. However, an alternative pathway allows glutamate to be converted to succinate. Cysteine is glucogenic and can be converted to pyruvate. Lysine was one of the top most significant metabolites more abundant at D4 and is ketogenic through acetyl-CoA. Threonine is both glucogenic, through succinyl-CoA, and ketogenic, through acetyl-CoA. Phosphoserine is an intermediate between glycolysis and serine production.

Urea and uric acid are both nitrogenous waste products. At D20, both amino acids that were more abundant were non-essential and glucogenic (glutamine, aspartate). Glutamine is converted to glutamate, while aspartate is converted to oxaloacetate. These differences in abundance may reflect increased catabolism of amino acids at D20, or differences in utilization of amino acids between days (Fig 8).

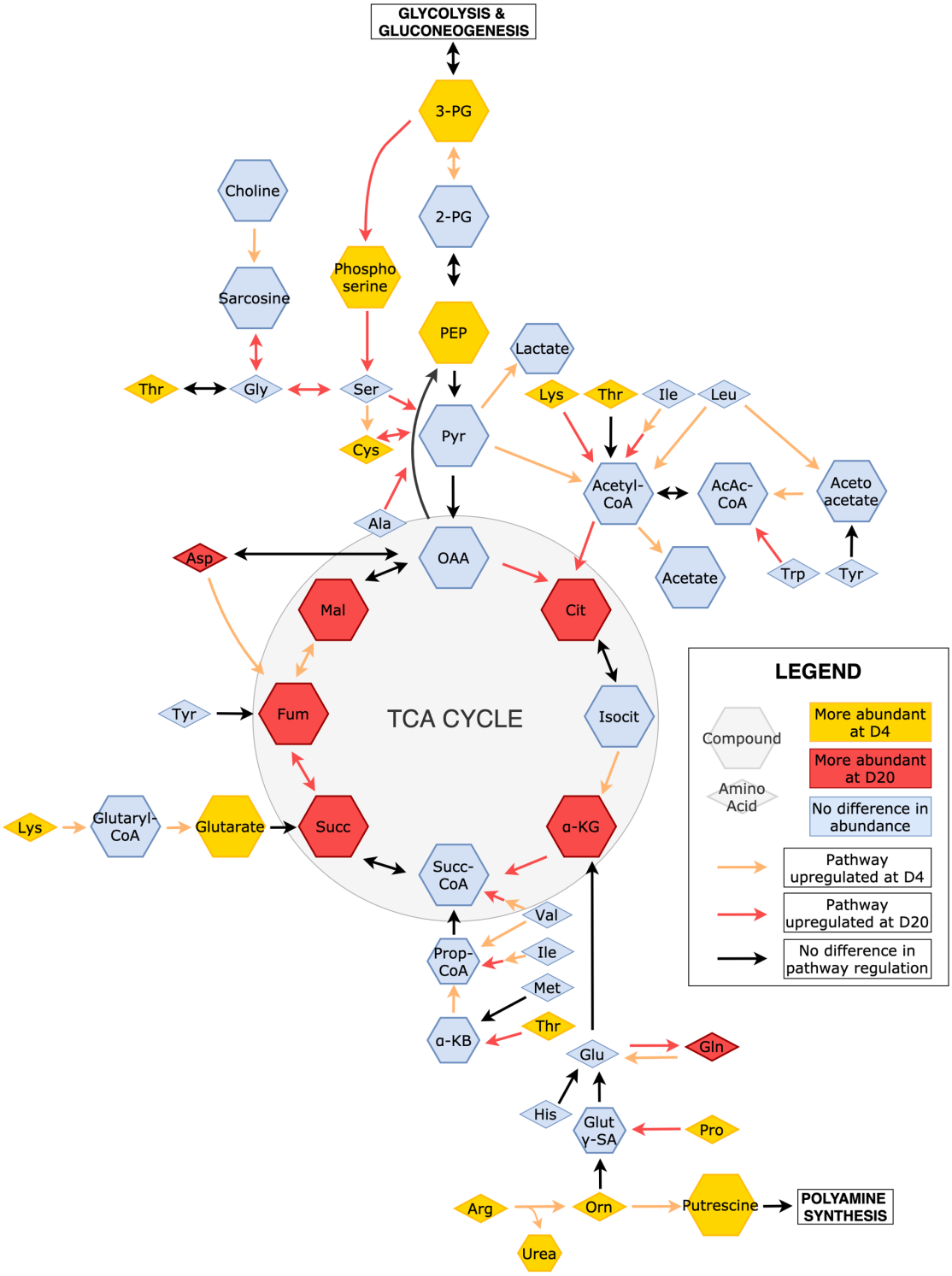
GLYCOLYSIS & GLUCONEOGENESIS
3-PG
2-PG
PEP
Pyr
Choline
Sarcosine
Phospho serine
Thr
Gly
Ser
Cys
Ala
Lactate
Lys
Thr
Ile
Leu
Acetyl-CoA
AcAc-CoA
Aceto acetate
Acetate
Trp
Tyr
Asp
OAA
Cit
Isocit
α-KG
Succ
Fum
Mal
TCA CYCLE
Tyr
Lys
Glutaryl-CoA
Glutarate
Succ-CoA
Val
Ile
Met
Thr
Prop-CoA
α-KB
Glu
Gln
His
Glut γ-SA
Pro
Arg
Orn
Putrescine
POLYAMINE SYNTHESIS
Urea
LEGEND
Compound
Amino Acid
More abundant at D4
More abundant at D20
No difference in abundance
Pathway upregulated at D4
Pathway upregulated at D20
No difference in pathway regulation

Figure 7: Amino acids as they relate to core metabolic pathways, especially the TCA cycle. Compounds that differed in abundance between days are highlighted, along with colored arrows representing upregulation. Abbreviations are as follows: 2-PG – 2-phosphoglycerate; 3-PG – 3-phosphoglycerate; α-KB – α-ketobutyrate; α-KG – α-ketoglutarate; AcAc-CoA – acetoacetyl-coA; Ala – alanine; Arg – arginine; Asp – aspartate; Cit – citrate; Cys – cysteine; Fum - fumarate; Gln – glutamine; Glu – glutamate; Glut γ-SA – glutamate γ-semialdehyde; Gly – glycine; His – histidine; Ile – isoleucine; Isocit – isocitrate; Leu – leucine; Lys – lysine; Mal – malate; Met – methionine; OAA – oxaloacetate; Orn – ornithine; PEP – phosphoenolpyruvate; Pro – proline; Prop-CoA – propionyl-CoA; Pyr – pyruvate; Ser – serine; Succ – succinate; Succ-CoA – succinyl-coA; Thr – threonine; Trp – tryptophan; Tyr – tyrosine; Val – valine.

As discussed above, at D4, the transcriptome data indicates that BPGM is shunting the intermediate 3-PG is towards glycolysis. In contrast at D20, the downregulation of BPGM suggests glycolytic intermediates are being directed towards serine biosynthesis. Two other transcripts encoding enzymes related to serine biosynthesis from glycolytic intermediates were upregulated at D20 (PHGDH, GLYCTK). PHGDH directs 3-PG towards serine biosynthesis, while GLYCTK converts glycerate to glycolytic intermediate 2-PG, a precursor of 3-PG. Several transcripts encoding enzymes involved in serine and glycine metabolism were also upregulated at D20 (SDSL, AGXT, PIPOX, SARDH, GNMT, ALAS2, GCAT, AOC3). AGXT catalyzes a number of reactions, including the interconversion of serine and glycine, interconversion of serine and hydroxypyruvate, and interconversion of glycine and glyoxylate. Both hydroxypyruvate and glyoxylate can go into glyoxylate metabolism. Although the main enzymes of the glyoxylate cycle have not been found in chickens, the liver has been observed to have glyoxylate activity (31). SARDH and PIPOX generate glycine from sarcosine, while GNMT

interconverts sarcosine and glycine. Sarcosine is an intermediate between glycine, creatine, and choline metabolism. SDSL catabolizes serine to pyruvate and also converts threonine to 2-oxobutanoate, an alpha-ketoacid intermediate of threonine catabolism to succinyl-CoA. ALAS2, GCAT, and AOC3 are all involved in generating different metabolites from glycine.

Proline and lysine metabolism may indicate increased collagen production and remodeling at D20. Although both metabolites were more abundant at D4, several enzymes facilitating their incorporation into collagen were upregulated at D20, (PYCR1, PYCRL, P4HA2, LOC425607, L3HYPDH, HYKK). PYCR1 and PYCRL are involved in the interconversion of proline, hydroxyproline, and pyrroline-5-carboxylate. P4HA2 and LOC425607 are involved in formation of collagen structural components from 4-hydroxyproline and hydroxylysine, respectively. HYKK is a kinase that phosphorylates hydroxylysine residues. One enzyme involved in collagen synthesis was upregulated at D4 (PLOD2), which is responsible for hydroxylation of lysine residues, allowing for cross-linking and stabilization of collagen.

Several transcripts upregulated at D4 encode enzymes that yield alternative TCA cycle intermediates, while several transcripts upregulated at D20 encode enzymes generating pyruvate from amino acids. In lysine degradation, two metabolites (lysine, glutarate) and two enzymes (DLD, DHTKD1) were more abundant at D4. DLD and DHTKD1 convert 2-oxoadipate to glutaryl-CoA, which can then be converted to glutarate and enter the TCA cycle through succinate. In contrast, EHHADH was upregulated at D20, supporting the canonical pathway of lysine degradation to acetyl-CoA. At D4, mRNAs encoding enzymes affecting aspartate and glutamate (ADSSL1, ALDH5A1) were enriched. ADSSL1 converts aspartate to

fumarate while ALDH5A1 metabolizes glutamate to succinate. Under normoxic conditions, aspartate is converted to oxaloacetate and glutamate is converted to α-KG. Given the TCA cycle is downregulated at D4 due to hypoxia, diverting these amino acids to different fates may allow them to be utilized more efficiently. Furthermore, this may serve a regulatory role in controlling levels of α-KG. Hence, the D4 liver may be relying on amino acids that are metabolized to intermediates suitable for anaerobic energy. This is further supported by the only TCA-specific mRNA upregulated at D4 encodes FH. Enrichment for FH would allow regeneration of oxaloacetate, which could then be converted to PEP, feeding directly into pyruvate for lactate production. At D20, mRNAs encoding enzymes producing pyruvate from various amino acids were upregulated (SDSL, CCBL1, AGXT). SDSL can convert serine to pyruvate while CCBL1 can interconvert cysteine and pyruvate. AGXT can convert alanine to pyruvate as part of the Cahill cycle, either replenishing levels of blood glucose through gluconeogenesis or oxidizing pyruvate for energy production in the TCA cycle and oxidative phosphorylation. Consequently, the liver at D20 appears to be utilizing amino acids for energy through canonical degradation pathways.

Homocysteine metabolism shows different directionality between days, possibly favoring the synthesis of cysteine at D4 and the synthesis of methionine at D20. Cysteine is synthesized from cystathionine, made from conjugating serine with methionine-derived homocysteine. The mRNA encoding the enzyme producing cystathionine was upregulated at D4 (LOC418544), along with metabolites cysteine and 2-hydroxybutanoic acid. 2-Hydroxybutanoic acid was one of the top most significant metabolites at D4 (Figure 6, Table 2), and is a byproduct of cysteine production. It is elevated when cysteine is limiting, such as when oxidative stress or

detoxification needs are high. At D4, higher levels of cysteine may be required to combat oxidative stress. In contrast at D20, two enzymes contributing to homocysteine production are upregulated (CCBL1, AHCYL1), suggesting that homocysteine is shunted to methionine salvage. Methionine is essential and one of the key amino acids important for growth in chickens (34,35).

Degradation of the branched chain amino acids valine, leucine, and isoleucine may serve different purposes between time points: generation of branched chain fatty acid precursors at D4, and complete catalysis for energy production at D20. Valine is glucogenic, entering the TCA cycle through degradation to succinyl-CoA. Leucine is ketogenic, as its metabolism results in acetyl-CoA. Isoleucine can follow either glucogenic or ketogenic routes. Seven enzymes involved in branched chain amino acid metabolism were more abundant at D4 (BCAT1, DLD, BCKDHB, ALDH6A1, AUH, AACS, HIBADH) than D20. BCAT1, DLD, and BCKDHB are involved in the early steps of degrading these branched chain amino acids. HIBADH and ALDH6A1 are involved in valine catabolism. These enzymes are responsible for the catabolism of valine to propanoyl-CoA. AUH and AACS are involved in leucine metabolism yielding acetoacetyl-CoA. As liver cannot efficiently metabolize acetoacetyl-CoA (34), it is exported for use by other organs. Four mRNAs encoding enzymes impacting branched-chain amino acids were enriched at D20 (EHHADH, HSD17B10, ACSS2, ACSS1L). HSD17B10 and EHHADH are dehydrogenases involved in beta-oxidation of a variety of compounds. HSD17B10 is mitochondrial and oxidizes steroids, while EHHADH is a bifunctional enzyme localized to the peroxisome and participates in multiple steps of branched-chain amino acid degradation as well as fatty acid oxidation. ACSS2 and ACSS1L activate short-chain fatty acids for further

metabolism, preferring acetate but also acting on propionate ultimately giving rise to citrate.

Arginine appears to contribute to polyamine synthesis at D4, consistent with increased cell proliferation. Four metabolites (urea, putrescine, ornithine, arginine) and transcripts encoding enzymes (ARG2, ODC1, GATM, AMD1, CHDH, DMGDH) involved in this were more abundant at D4. ARG2 converts arginine to ornithine, releasing urea as a byproduct. ODC1 converts ornithine to putrescine, a polyamine critical for cell proliferation (35). GATM supports the conversion of glycine to arginine through guanidinoacetate. CHDH and DMGDH support the catabolism of choline to sarcosine, a glycine precursor, which could provide additional substrates for polyamine synthesis. AMD1 is involved in methionine salvage, where is generates S-adenosyl-methionine (SAM), an important methyl group donor and component of polyamine synthesis. In addition to contributing to cell proliferation, this pathway may also divert excess arginine away from α-KG. The increased breakdown of arginine for polyamine synthesis could also explain the increased abundance of urea at D4. In contrast, one enzyme converting arginine to agmatine (AZIN2) was upregulated at D20, signifying a possible alternate route to polyamine synthesis.

Transcriptome data suggests that glutamate and glutamine metabolism is reversed between D4 and D20. Metabolism favoring glutamine production appears favored at D20, while at D4, glutamine conversion to glutamate is likely enhanced. Glutamate-ammonia ligase (GLUL, elevated at D20) synthesizes glutamine from glutamate and ammonia, while glutaminase (GLS, elevated at D4) converts glutamine to glutamate and ammonia. In adult birds, the synthesis of uric acid, the predominate nitrogen waste product, requires glutamine. In contrast, enrichment for glutamate at

D4 can support production of succinate, which has important roles in stabilizing the hypoxia response (36).

2.2.5.3 Lipid metabolism

At hatch, chicks are supplied with lipids absorbed from the yolk sac while in mature birds the liver receives dietary lipids directly from the small intestine via the hepatic portal vein. Yolk nutrients serve as the main source of lipids for D4 birds, as bile acid capabilities to facilitate dietary lipid absorption are not fully developed until after the first week post-hatch (6,36). Post-hatch, the yolk sac membrane is able to continue absorbing lipids directly into the bloodstream through endocytosis, bypassing the small intestine and the need for liver-produced bile acids; this lipid reservoir provides fuel for early post-hatch growth (36). Compared with D20 liver, the D4 liver metabolome evidences enrichment of multiple yolk-derived fatty acid metabolites including: palmitoleic, myristic, linoleic, and oleic acid. Two acyl synthetases involved in the activation of fatty acids were upregulated at D4 (ACSL3, ACSL4), Acyl synthetases activate free fatty acids in the cytoplasm via conjugation with an acyl group. ACSL3 preferentially acts on myristic acid, arachidonic acid, lauric acid, and eicosapentaenoic acid, and is involved in mediating hepatic lipogenesis. ACSL4 prefers arachidonic acid and is often expressed in steroid-producing organs. Two enzymes (CPT2 and ACAD9) involved in fatty acid beta-oxidation were upregulated at D4, and appear representative of the available substrates in yolk-derived lipids at D4. CPT2 prepares activated fatty acyl-CoAs for oxidation by removing the acyl group once they have entered the mitochondria. ACAD9 is a mitochondrial acyl-CoA dehydrogenase catalyzing the rate-limiting step of beta-oxidation, with a preference

for palmitoyl-CoA and long-chain unsaturated fatty acids. The substrate preferences of ACAD9, ACSL3 and ACSL4 are consistent with an abundance of yolk-derived fatty acids consisting of mainly palmitate and myristate.

Several enzymes involved in lipid elongation in the endoplasmic reticulum (ELOVL7, HACD2, HACD3) were upregulated at D4, suggesting the synthesis of lipids meeting different requirements such as cell membrane components. Elongation in the ER gives rise to long-chain and very-long-chain fatty acids (>C16). ELOVL7 catalyzes the rate-limiting first step, preferring long-chain and very-long-chain fatty acids (C18, C16, C20), and is involved in synthesizing membrane lipid precursors and lipid mediators. HACD2 and HACD3 are dehydratases involved in the dehydration step of lipid elongation. These long chain fatty acid products could be used within the liver or exported to other tissues to support the rapid growth seen in chicks at D4. Two enzymes associated with mitochondrial fatty acid biosynthesis were upregulated at D20 (MCAT, MECR). Enzymes acting as part of the fatty acid synthase type II (FAS2) complex catalyze mitochondrial fatty acid synthesis. Malonyl transacylase (MCAT), catalyzes the transfer of malonyl from malonyl-CoA to a scaffold protein, beginning the elongation phase. Enoyl ACP reductase (MECR) catalyzes the final step with palmitate as the main product of fatty acid synthesis in the mitochondria.

Transcripts encoding enzymes involved in beta-oxidation in the mitochondria and peroxisome were more abundant at D20 including: EHHADH, ECI1, DECR2, and ACAD11. While mitochondrial beta-oxidation generates acetyl-CoA for energy production, peroxisomal beta-oxidation breaks down very long chain fatty acids (>22C), branched fatty acids and leukotrienes to acetyl-CoA and medium chain fatty acids for biosynthesis of specific fatty acids. EHHADH is a multi-functional,

peroxisomal enzyme that acts as both an enoyl-CoA hydratase and a 3-hydroxyacyl-CoA dehydrogenase, catalyzing multiple steps in beta-oxidation. Along with EHHADH, ECI1 rearranges double bonds in unsaturated fatty acids to facilitate their oxidation. DECR2 is a peroxisomal enzyme degrading varying lengths of unsaturated fatty acids. ACAD11 encodes mitochondrial acyl-CoA dehydrogenase catalyzing beta-oxidation of fatty acids with chain lengths of 20 to 26 carbons.

Four thioesterase transcripts were enriched at D20 (ACOT4, ACOT8, ACOT11, THEM4) while only one (ACOT12) was enriched at D4. Thioesterases hydrolyze acyl-CoA to coenzyme A and free fatty acids. ACOT4 is peroxisomal and acts on succinyl-CoA, glutaryl-CoA, and long chain saturated acyl-CoAs. ACOT8 is peroxisomal and has the least specific substrate preferences, acting on medium chain (C2 to C20) saturated or unsaturated fatty acyl-CoAs and also bile acids. ACOT11 prefers long chain substrates including palmitoyl-CoA and myristoyl-CoA, which can go on to mitochondrial beta-oxidation. THEM4 is mitochondrial and prefers medium to long chain substrates (C14 to C18). ACOT12, the only thioesterase upregulated at D4, prefers acetyl-CoA as a substrate.

In glycerolipid metabolism, metabolites and gene expression patterns indicated triacylglycerol (TAG) breakdown was favored at D4 while TAG synthesis was favored at D20. Two metabolites (3-PG, 1-monopalmitin) and several enzymes (LIPC, LIPG, AKR1B10L2, PLPP1, LPIN1, GPAM) were more abundant at D4. LIPG (endothelial) and LIPC (hepatic) are lipases that hydrolyze the ester bond of TAGs to form monoacylglycerol and free fatty acids. 1-Monopalmitin was one of the top metabolites contributing to random forest classification at D4. At D20, one metabolite (glycerol-3-phosphate) and several enzymes (DGKA, DGKZ, DGAT2, AWAT1,

TKFC, GLYCTK, GPAT4, GPAT2, AGPAT4, AKR1B10L1) were more abundant. Glycerol-3-phosphate provides the glycerol backbone for TAG synthesis. Three acyltransferases (AGPAT4, GPAT2, GPAT4) are involved in different steps generating phosphatidate from glycerol and acyl-coA, and two acyltransferases (DGAT2, AWAT1) are responsible for the final step of TAG synthesis. At D4, a different acyltransferase involved in generating phosphatidate from glycerol and acyl-CoA was upregulated (GPAM). GPAM is a mitochondrial acyltransferase that prefers saturated fatty acids. In the preparation or recycling of the glycerol backbone, one aldo-keto reductase catalyzing the interconversion of glyceraldehyde and glycerol was upregulated at D4 (AKR1B10L2) while one was upregulated at D20 (AKR1B10L1), and one kinase converting glycerate to glycolytic intermediate 2-PG was upregulated at D20. The interconversion of phosphatidate and diacylglycerol shows contrasting directionality between days; two phosphatases that catalyze the conversion of phosphatidate to a diacylglycerol were upregulated at D4 (PLPP1, LPIN1), while two kinases phosphorylating diacylglycerols to phosphatidate (DGKA, DGKZ) were upregulated at D20. Given the importance of phosphatidate as an important precursor molecule for both TAG synthesis and membrane lipid synthesis, this may represent an important metabolic reprogramming between time periods.

2.3 Discussion:

Modern broilers are the product of intensive selection for rapid growth, increased skeletal muscle production, and improved feed efficiency. The liver, as the metabolic center of the body, plays an important role in supporting the emergence of

these traits. This study indicates that the significant metabolic reprogramming of the liver occurs between D4 and D20 post-hatch, summarized in Figures 9 and 10, respectively. This reprogramming is choreographed by modulation of gene expression patterns responding to changes in the proliferative needs of the liver and other tissues. In addition, the nutritional sources of the birds change during this time, shifting from stored yolk to oral nutrition. At D4, liver cells are undergoing rapid proliferation as indicated by the positive allometric growth of the organ at this time. This is supported by the transcriptome analysis that identified enrichment for multiple genes involved in cell proliferation, cell cycle, and DNA replication at D4. In addition, the integrated transcriptome and metabolome data indicates metabolic flexibility especially in carbohydrate metabolism. In glycolysis, D4 liver is manifesting metabolism reminiscent of the Warburg effect, which was first noted in rapidly dividing cancer cells (37,38) and has been shown to be common in normal proliferating cells (39). In contrast, by D20 the liver's relative growth has slowed and the transcriptome has shifted from supporting cell division to enrichment for immune function. Metabolomic data indicates D20 liver no longer displays upregulated anaerobic energy production in carbohydrate metabolism and has taken on the role of acting to store glucose and synthesize fatty acids, functions associated with the mature liver.

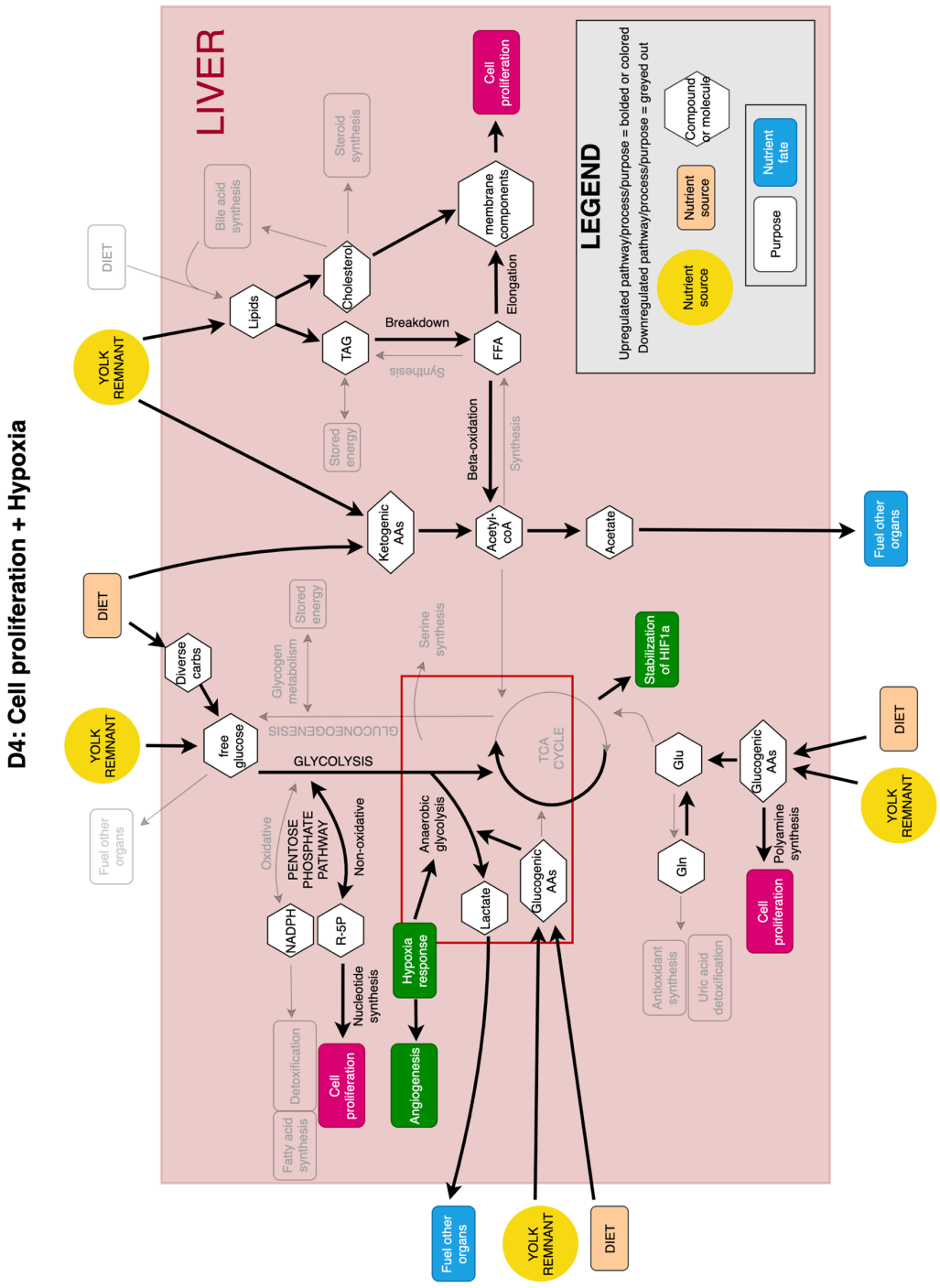

Figure 8: Summary of metabolic reprogramming at D4, in terms of core metabolic pathway regions and biological processes that are upregulated or important. Upregulated components are in color, downregulated components are greyed out. Note: Some arrows represent pathways involving multiple steps.

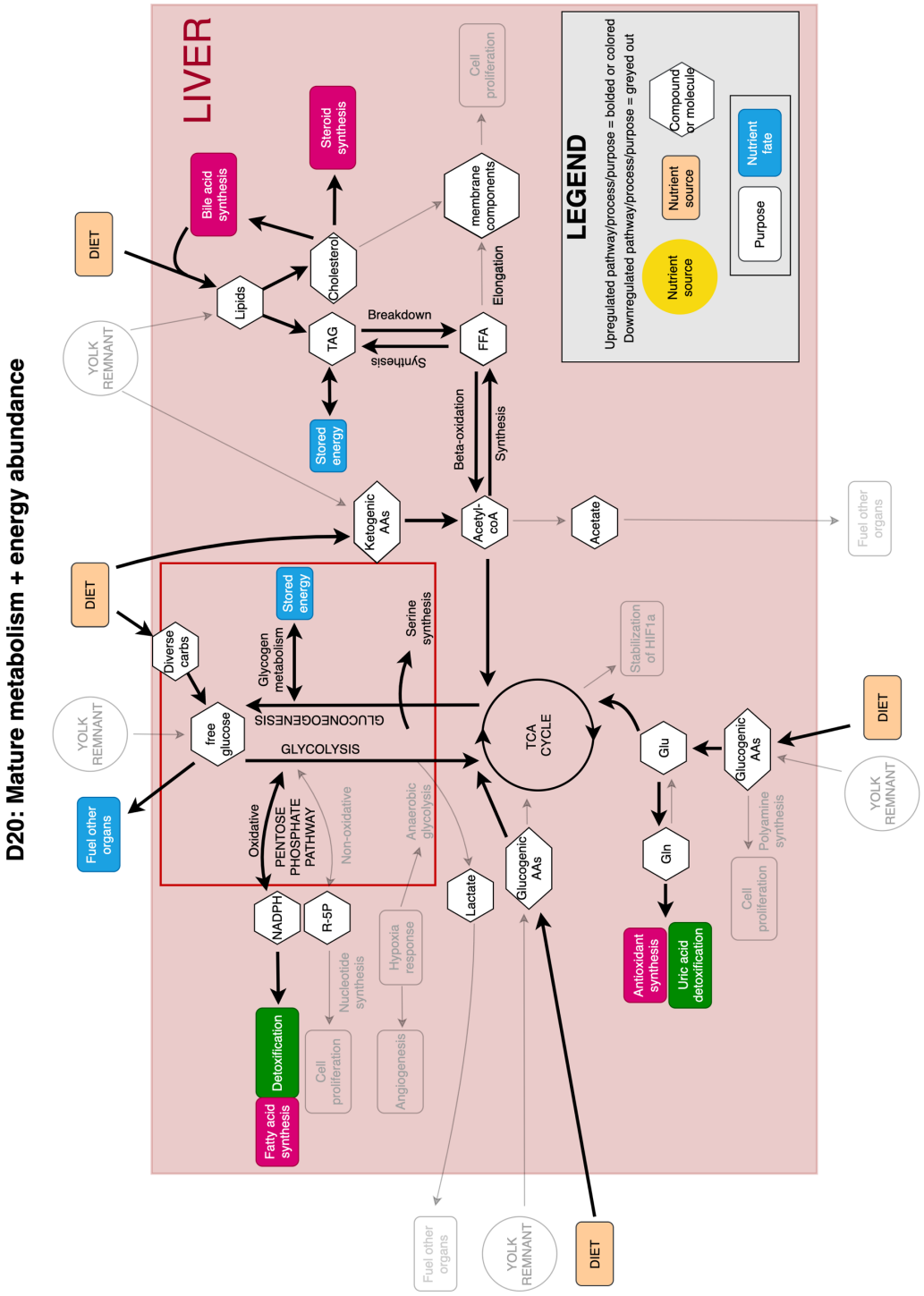

Figure 9: Summary of metabolic reprogramming at D20, in terms of core metabolic pathway regions and biological processes that are upregulated or important. Upregulated components are in color, downregulated components are greyed out. Note: Some arrows represent pathways involving multiple steps.

An important hypothesis arising from this work is that the D4 liver is hypoxic, which may arise from vascularization lagging behind the rapid growth of the liver. Enrichment of the hypoxia responsive transcription factor HIF1A transcript seen at D4 could lead to increased expression of angiogenic growth factor receptors (VEGF receptors KDR and FLT1) to improve vascular development. HIF1A also plays a central role in increasing glycolytic rate through isozyme selection, and by triggering the Warburg effect; with a central impact of increasing expression of the LDHA isoform that directs glycolysis towards lactate production (40,41). The NAD^+ produced from the reduction of pyruvate to lactate can support the GAPDH enzymatic activity required to drive the ATP production phase of glycolysis. This is necessary to rapidly dividing cells under hypoxia as it maintains ATP production. In addition, lactate secreted from the cell could reduce the blood pH at D4 compared with D20 (Table 1). This reduced pH that results from lactate may further promote matrix rearrangement and angiogenesis (42). Isozyme selection by HIF1A at D4 is also evident in upregulation of platelet-type phosphofructokinase (PFKP). Metabolic reprogramming similar to that observed in cancer, driven by HIF1A, likely allows D4 liver to continue efficient energy production under hypoxia, while increasing angiogenesis to eventually escape these conditions as metabolism matures.

The data indicates that pyruvate occupies an important metabolic fork in the D4 liver. Along with PDK, transcriptome data indicates that two components of the PDH complex, and two pyruvate dehydrogenase phosphatases (PDP) are enriched in the D4 liver. This suggests that pyruvate metabolism in the early post-hatch liver exhibits greater metabolic flexibility than is seen in a cancer cell exhibiting the Warburg effect. The growing liver tissue at D4 retains the option of reverting to

aerobic glycolysis and the TCA cycle conferred by tight regulation of this metabolic fork. In turn, this allows for a rapid shift between pyruvate being converted to either lactate in the cytoplasm or acetyl-CoA and then citrate in the mitochondria as metabolic demands change.

In the hypoxic conditions at D4, where glycolysis is the main ATP-producing pathway, metabolic flexibility also allows for differences in nutrient utilization. We hypothesize that the liver is conserving glycolytic intermediates for energy production, including glucogenic amino acids and recycled glycerol backbones, while sparing free glucose for the rest of the body. This is supported by the overall downregulation of the early stages of glycolysis and gluconeogenesis at D4, along with multiple examples of glucogenic amino acids being directed towards oxaloacetate or pyruvate precursors for anaerobic glycolysis. Transcripts encoding enzymes involved in degradation of ketogenic amino acids are upregulated at D4, along with those converting acetyl-CoA to acetate and acetoacetate to acetoacetyl-CoA. The resulting lactate, ketones, short chain fatty acids and other compounds may act as important sources of energy for other tissues. For example, free acetate produced in the liver can be exported for use by other tissues such as cardiac muscle (43), while skeletal muscle can oxidize lactate (44). The dramatically increased expression of two monocarboxylate transporters – SLC16A3 (logFC 3.29) and SLC5A12 (logFC 5.22) – at D4 support the increased import and export activity of compounds like lactate, pyruvate, and oxo acid products of branched chain amino acid catabolism, further suggesting that differences exist in the metabolic demand for these compounds between time points, in both the liver and rest of the body.

Inspection of glycolysis metabolites suggests a second metabolic fork is present in this pathway at 3-phosphoglycerate. BPGM, the enzyme controlling this fork, is upregulated at D4. Knockout of BPGM in mice increases serine and phosphoserine biosynthesis (45). Thus, elevated BPGM levels at D4 suggest that it is directing 3-PG towards glycolysis rather than serine biosynthesis. In addition to modulating this branching point, BPGM is known for its role in decreasing hemoglobin's affinity for oxygen. The metabolite 2,3-BPG which is the product of BPGM enzymatic activity, can escape cells and promote the release of oxygen from red blood cell hemoglobin, possibly ameliorating anoxic conditions (46). Both 3-phosphoglycerate and phosphoserine are enriched in the D4 liver suggesting that some of the glycolytic intermediates are exchanged with serine production at this time. Serine serves several important functions in the metabolism of rapidly growing cells (47), including its roles in the synthesis of proteins, phospholipids, cysteine, glycine and single carbon metabolism. Phospholipids are incorporated into the membranes and single carbon metabolism is essential to the formation of many different metabolites.

The origin of the 3-phosphoglycerate fork may be explained by the enrichment for phosphoenolpyruvate at D4. Pyruvate kinase (PKLR) in the chicken is orthologous to human PKM1/2. In humans, proliferating cells frequently express the embryonic form of pyruvate kinase, which is responsible for the formation of pyruvate from phosphoenolpyruvate (48). The activity of PKM2 is responsive to allosteric regulators and post-translational modifications. In proliferating cells, PKM2 activity is inhibited, yielding elevated levels of phosphoenolpyruvate (seen at D4) and other glycolytic products (49). Some of these products may be diverted from pyruvate production to anabolic processes such as serine production or supporting nucleotide production via

the non-oxidative phase of the pentose phosphate pathway (PPP). Elevated levels of the TKT1 and PRPS2 transcripts in the D4 liver support enrichment for the non-oxidative PPP phase. Alternative splicing of the PKM transcript produces PKM2. Based on what is known about PKM1/2 in humans, we hypothesize that in the chicken, PKLR is similar in both role and splicing. Our current analysis does not distinguish splice products, so investigation of the hypothesis of the potential role of PKLR in D4 glycolytic regulation awaits future work.

Consistent with a hypoxic environment, metabolomic data indicates that the TCA cycle is suppressed in the D4 liver compared with D20. However, transcripts encoding enzymes functioning in the TCA cycle including Isocitrate Dehydrogenase 3A (IDH3A), Dihydrolipoamide Dehydrogenase (DLD), and Fumarate Hydrolase (FH) are enriched at D4. Elevated IDH3A and DLD protein expression would reduce levels of α-ketoglutarate, a potent inhibitor of HIF1A activity (50). FH could function in the anaplerotic roles of glutamate, lysine, threonine and proline at D4. These amino acids would replenish the levels of oxaloacetate that is needed to react with acetyl-CoA to yield mitochondrial citrate. This mechanism is apparently seen when citrate is being actively transported from the mitochondria to support lipid synthesis (51). In addition, glutamate and threonine can be metabolized to Succinyl-CoA, an important precursor along with glycine to the synthesis of porphyrins.

Another metabolic pathway enriched at D4 is the synthesis of polyamines, which are important to sustaining proliferation. This pathway can also explain the elevated levels of urea seen in the D4 chick. While urea is the typical excretory product of nitrogen metabolism in mammals, very little is produced in the liver of chickens with most nitrogen waste disposed of as uric acid. However, urea is produced

as a byproduct of an early step in the synthesis of polyamines, the enzymatic conversion of arginine to ornithine by arginase (ARG2). Hence, we hypothesize that this is the origin of the elevated urea in the D4 compared with the D20 liver.

By D20, the liver is apparently under normoxic conditions with the expression of HIF1A reduced compared to D4. TCA enzymes, mitochondrial processes including oxidative phosphorylation, and other oxygen-dependent metabolic pathways are also more active at D20. Comparison between the D20 and D4 transcriptome and metabolome provide evidence for significant metabolic reprogramming between the two days. By D20, glucose metabolism is enriched for transcripts and metabolites associated with phase 1 of the glycolysis pathway, with glucose-6-phosphate residing at a metabolic fork. Transcripts encoding enzymes associated with both glycolysis (HK3, GCK, PFKL, PFKM) and gluconeogenesis (G6PC and G6PC3) are enriched at D20. This is consistent with the need of the mature liver to capture glucose by phosphorylation, or export glucose during fasting for use by other tissues. The G6P can be converted to glucose 1-phosphate (enriched at D20) and stored as glycogen. Alternatively, G6P can be converted to 6-phosphogluconolactone, which is the first step of the pentose phosphate pathway. In fact, by D20 the pentose phosphate pathway has become enriched for components of the oxidative phase, likely increasing the production of NADPH to support the need for fatty acid synthesis and detoxification of compounds. Upregulation of glycolytic enzyme PFKL has also been linked to growth potential in broilers, with one study suggesting that this isoform may contribute to the rapid growth rate of maturing birds (18).

In contrast to D4, TCA cycle metabolites are elevated at D20, possibly by translation of the enriched citrate synthase transcript. We hypothesize that the elevated

levels of citrate play a fundamental role in the metabolic reprogramming seen between D4 and D20. Citrate plays multiple roles as a metabolite and as a regulator of enzymatic activities. As a metabolite, it is the first component of the mitochondrial TCA cycle and can be exported to the cytoplasm and metabolized to support both nucleotide and fatty acid synthesis. As an enzyme regulator it serves as an indication of the energy state of the cell, becoming abundant at high ATP levels. Citrate inhibits the activity of phosphofructokinase (PFK) preventing the ATP synthesis phase of glycolysis, as is suggested in the D20 data. Citrate also serves as an activator of acetyl-CoA-carboxylase (ACACA) (52), the first enzyme in the cytoplasmic synthesis of fatty acids. Although there was no measurable difference between D4 and D20 ACACA levels, high levels of citrate would likely drive fatty acid synthesis in the D20 liver. Another impact of a high citrate level that results in elevated malonyl-CoA is inhibition of CPT2 that is part of the lipid transport system that brings fats into the mitochondria for beta-oxidation.

D20 transcriptome and metabolome data provide additional information relevant to fatty acid metabolism. Transcripts encoding enzymes driving mitochondrial fatty acid synthesis and those involved in beta-oxidation in both the mitochondria and peroxisome were elevated compared with D4. Peroxisomal beta-oxidation is a major source of acetate that can be released from the liver for use by other tissues following conversion to acetyl-CoA (53,54). The D20 liver is enriched in glycerol-3-phosphate, an important metabolite in triacylglycerol synthesis. The potential importance of triacylglycerol synthesis in the D20 liver is also supported by enrichment for transcripts encoding enzymes responsible for phosphatidate synthesis, which is essential for TAG production. Also enriched were transcripts encoding

phosphatidylcholine synthetic enzymes. In addition to its role in lipid membranes, phosphatidylcholine is a major component of very low-density lipoproteins, which are secreted from the liver to transport TAGs.

2.4 Conclusions

The liver of the modern broiler must meet the needs of rapid growth that has resulted from intensive genetic selection. Allometric comparisons indicate that the modern broiler liver reaches its maximum normalized size several days earlier than slower growing chickens. As slow growing and modern broilers are of similar size at hatch, this suggests that selection has compressed the positive allometric growth phase of the liver into the first week post-hatch. The transcriptome and metabolome of the D4 liver indicate the tissue is undergoing rapid proliferation under hypoxic conditions with carbohydrate, lipid and amino acid metabolism primed to support this growth. At D4, ATP is likely to be largely derived from glycolysis, the PPP is favoring production of nucleotide precursors and returning carbon backbones to glycolysis, and lipid biosynthesis appears to support fatty acid elongation, perhaps for membrane production in the proliferating cells. By D20, the liver is undergoing negative allometric growth and has transitioned from predominantly supporting its own proliferation, to supporting the metabolic needs of other tissues. Metabolic reprogramming has shifted the glycolytic pathway to the ATP investment phase allowing for rapid responses to either store or release glucose, the PPP now appears to be shifted towards producing NADPH to support anabolic reactions such as lipid production and lipid metabolism appears to have shifted to triacylglycerol production for lipid export.

Understanding the molecular and physiological changes occurring in post-hatch broiler liver by exploiting high-throughput methods has broad utility. A major research goal at the intersection of computation and biology involves understanding the strengths and limitations of different types of data, as well as the relationship between them. In any system, comparing knowledge gained from different types of high-throughput technologies can corroborate existing biology and generate new, testable hypotheses. Lastly, the chicken is a valuable, accessible, and well-supported model both in terms of research and infrastructure for studying the effects of genetic selection especially as it relates to metabolic efficiency. In this work we have developed a preliminary contrast of the liver at two time points representing critical biological differences in the modern broiler. Some of our observations recapitulate known biology and provides possible explanations for phenotypic characteristics and allometric relationships observed in the modern broiler. While hypoxia has a documented role in controlling embryonic development (55), the role of hypoxia in normal post-hatch or post-natal tissue development is not well characterized. A major future goal will be to build networks of interactions to better understand the regulation of this metabolic reprogramming as the growing chicken transitions from hypoxic to normoxic conditions.

2.5 Methods

2.5.1 Bird Husbandry, Necropsy, and Tissue Collection

Day-old male Ross 708 chicks were obtained from a commercial hatchery (Mountaire Farms, Millsboro, DE) and grown on the University of Delaware Farm (Newark, DE). Standard management and husbandry procedures were followed, as

approved by the Animal Care and Use Committee of the University of Delaware (IACUC 72R-2017-0). On each even-numbered day post-hatch from Day 4 (D4) through Day 20 (D20), 12 birds were randomly chosen and humanely euthanized by cervical dislocation. Prior to euthanasia, birds were weighed and blood was drawn from the brachial wing vein for immediate i-STAT blood chemistry analysis using CG8+ cartridges. 200uL of blood was also collected in EDTA tubes on ice and then centrifuged to separate plasma for metabolome analysis. During necropsy, multiple tissues were systematically collected and snap frozen in liquid nitrogen. Select organ masses (heart, liver, breast muscle) and intestine segment lengths were also recorded. Liver was collected from the caudal portion of the left lobe, with an additional tube of tissue saved for metabolome analysis. Frozen tissues and plasma were subsequently stored at -80^{o}F until further use.

2.5.2 Analysis of phenotypic and i-STAT measurements

Phenotypic measurements and i-STAT blood values were analyzed using JMP Statistical Software (JMP®, Version 14.0.0). Pearson correlations were calculated between all measurements. Hierarchical clustering was performed and a heatmap generated for visualization. Given the small sample sizes of each day (7 and 10), a Shapiro-Wilk test was used to test each variable distribution for normality. To assess statistical significance of the differences between the two bird age groups, tests were done to compare measures of center. A 2-sample t-test was conducted on each variable meeting normality criteria, while a non-parametric 2-sample Wilcoxon test was performed for each variable that did not (PO_2, BE, Na).

2.5.3 Metabolome

Frozen liver and plasma were shipped on dry ice to the UC Davis West Coast Metabolomics Center (Davis, CA) for untargeted metabolomics analysis. Primary metabolism analysis was done with gas chromatography-time of flight/mass spectrum (GC-TOF MS) and reported as peak heights normalized to mean total ion current (mTIC), or average total metabolome levels for each tissue. A total of 657 compounds were reported. 205 were identified by name, while 452 were assigned BinBase database identifiers. Analysis of metabolome data was done using Metaboanalyst (56–58). Statistical analysis was performed on all samples (11 D4, 12 D20). Variables with very small values were considered to be non-informative and were detected and filtered using median intensity value. Remaining data was log transformed and Pareto scaled. Principal component analysis (PCA) and random forest were performed. Compounds that differed in level between bird age groups were identified by fold change and t-test, and visualized using a volcano plot. The volcano plot was generated using a log fold-change cutoff of 1, equal variances, and an adjusted p-value cutoff of 0.1. Pathway enrichment analysis was done for each day separately using curated compound names. Compounds used were determined by t-test (adjusted $p<0.05$), resulting in 37 at D4 and 41 at D20. Fisher's Exact Test was used for overrepresentation, out-degree centrality, the *Gallus gallus* pathway library, and the default background. For KEGG Search and Color Pathways search (59), InChiKeys provided by UC Davis for each metabolite were matched with compounds with the same isotopes but varying stereochemistry using PubChem Identifier Exchange (60) to obtain KEGG Compound IDs. Thus, compounds would have the same mass to charge ratio but vary in structure, allowing for chemically similar forms of metabolite compounds (i.e. different conformations of sugars).

2.5.4 Transcriptome

Total and small RNA was isolated from each frozen liver sample using mirVana miRNA Isolation Kit with phenol (Thermo Fisher Scientific, AM1560) and DNAse treated using DNA-free DNA Removal Kit (Thermo Fisher Scientific, AM1906) according to the manufacturers' protocols. After each step, RNA concentration and purity were tested using a NanoDrop spectrophotometer (NanoDropTM 1000, Thermo Scientific). Sample quality was further assessed at the University of Delaware Sequencing & Genotyping Center (Delaware Biotechnology Institute, Newark, DE) using an AATI Fragment Analyzer. Samples meeting a threshold of RNA Integrity Number (RIN) >= 7.0 were retained for library preparation. In total, 10 samples from D4 and 7 samples from D20 met the cutoff. cDNA libraries were prepared and indexed using Illumina TruSeq Stranded mRNA Library Prep Kit (Illumina, RS-122-2101) according to the manufacturer's protocol. Final library concentration was assessed using Qubit dsDNA BR Assay Kit (Thermo Fisher Scientific, Q32850) and Qubit fluorometer 2.0 (Thermo Fisher Scientific). Libraries were pooled in groups of eight and paired-end sequenced at DBI on an Illumina HiSeq 2500 using one lane per pool and generating read lengths of 51 base pairs (bp).

Fastq files were down-sampled to 15000 reads each using seqtk (61) to dilute any batch effect. Read quality was analyzed both before and after trimming using FastQC v0.11.7 (62). Trimming was done using TrimGalore v0.4.5/CutAdapt v1.16 (63,64). Paired-end reads were mapped to Ensembl galGal6 (GRCg6a) (65) using splice-aware mapping software HiSat2 v2.1.0 (66) and converted to sorted bam files with SAMtools (67). Transcripts and gene abundance in both fragments per kilobase of transcript million mapped reads (FPKM) and transcripts per million (TPM) were

quantified with Stringtie v1.3.4d (68–70) using Ensembl reference annotation version 95 (65). Raw counts were quantified with featureCounts v1.6.0 (71).

Differential expression was performed using statistical software R DeSeq2 package (72). DeSeq2 uses Benjamini-Hochberg (FDR) correction for multiple hypothesis testing to adjust p-values (73). Genes meeting significance threshold of adjusted p-value <0.05 from each method were compared for agreement, yielding 945 genes more highly expressed in D4 and 1265 genes more highly expressed in D20.

2.5.5 Ontology & pathway enrichment

Differentially expressed genes (DEG) were analyzed for gene ontology and pathway enrichment analysis using DAVID 6.8 (27) and PANTHER (74,75). Significant DEG (adjusted p<0.05) for each day were converted to NCBI gene IDs using bioDBnet (76) and put into DAVID for ontology analysis. 1342 gene IDs were recognized for D4, and 1703 for D20. Default parameters were used, except for gene ontology categories, "FAT" was chosen rather than "Direct." Ontology was also analyzed using official gene symbols to take advantage of the improved annotation quality for human genes. 1288 gene IDs were recognized for D4 by this method, and 1495 for D20. Differentially enriched pathways were identified using KEGG (59), Reactome (77), and STRING (78).

2.5.6 Principal component analysis

Stringtie outputs gene abundance in transcripts per million (TPM). Genes were filtered, retaining those with values of at least 0.1 in 7 or more samples. Principal component analysis (PCA) was done using these 13117 genes and R packages FactoMineR (79) and factoextra (80).

2.5.7 Metabolic pathway analysis and interpretation

The list of differentially expressed genes and differentially abundant metabolites were input to KEGG Search & Color pathway tool. KEGG Search & Color allows the input of lists of genes and/or metabolites, with coloring based on groupings specified by the user. NCBI gene IDs from all significant DEG (adjusted p <0.05, total 3065 genes) and KEGG compound IDs from metabolite analysis (243 compounds) were used to search against *Gallus gallus*-specific pathways, with those more abundant at D4 colored yellow, and those more abundant at D20 colored red. Core metabolic pathways, including carbohydrate, lipid, and amino acid metabolism, were then visualized to identify specific areas of these pathways that were enriched in each day. After focusing on pathways and areas of interest, specific functions of differentially abundant genes and metabolites, in relation to metabolism and biological context, were researched and interpreted.

Chapter 3

CORRELATION NETWORK ANALYSIS INTEGRATING TRANSCRIPTOME AND METABOLOME

3.1 Background

The comparison of D4 and D20 hepatic transcriptome and primary metabolome described in Chapter 2 generated multiple hypotheses and questions and identified important biological events such as metabolic reprogramming and expansion of the immune system. However, this perspective was limited due to the exclusion of intervening days and most of the metabolomic dataset. Thus, the aim of the next analysis was to involve data from the intervening time points, including liver transcriptome, and liver and plasma primary metabolome, to decipher more complex patterns throughout the time course. In particular, we wanted to identify features that differed significantly in abundance over the time course, cluster features into groups showing correlated patterns of abundance and association with physiological traits, and begin to develop a systems-biology perspective about how changes in the liver's metabolism may be affecting the rest of the body. To accomplish this we chose an all-inclusive correlation-based approach: Weighted Gene Correlation Network Analysis (WGCNA) (81).

WGCNA is a computational method commonly used with expression data that clusters features into modules based on weighted pairwise correlations, and relates these modules to traits of interest. WGCNA is a powerful and highly interpretable method for dimension reduction, feature prioritization, and pattern identification that

lends itself to many types of downstream analysis. The clusters contain features with similar patterns of abundance, each represented by a module eigengene. Each feature is assigned to a module based on strength of module membership, or correlation with the module eigengene, and associated p-value. Highly interconnected features, both within and between modules, are considered to be important and may drive patterns in each individual module. Modules may represent features with closely-linked regulation, or in this analysis, those that are critical to liver function despite other fluctuations throughout the time course. The novelty of this analysis involved using WGCNA to integrate two types of high-throughput data, so special care was taken in data preprocessing to ensure features were as comparable as possible (see Methods). Additionally, the use of meaningful, computationally-driven clustering methods such as WGCNA can place unknown genes and unannotated metabolites in context by associating them with groups of genes and metabolites with known function.

3.2 Results

3.2.1 ANOVA

Figure 10 shows principal component analysis (PCA) of primary metabolites by each tissue. 264 liver primary metabolites and 118 plasma primary metabolites were significantly different between at least two time points. The liver metabolome is more variable than the plasma metabolome, suggesting that throughout this time course, the liver is doing its job of buffering blood compound levels. Table 3 shows the top 25 compounds differing in each tissue along with the top 5 contrasts for each.

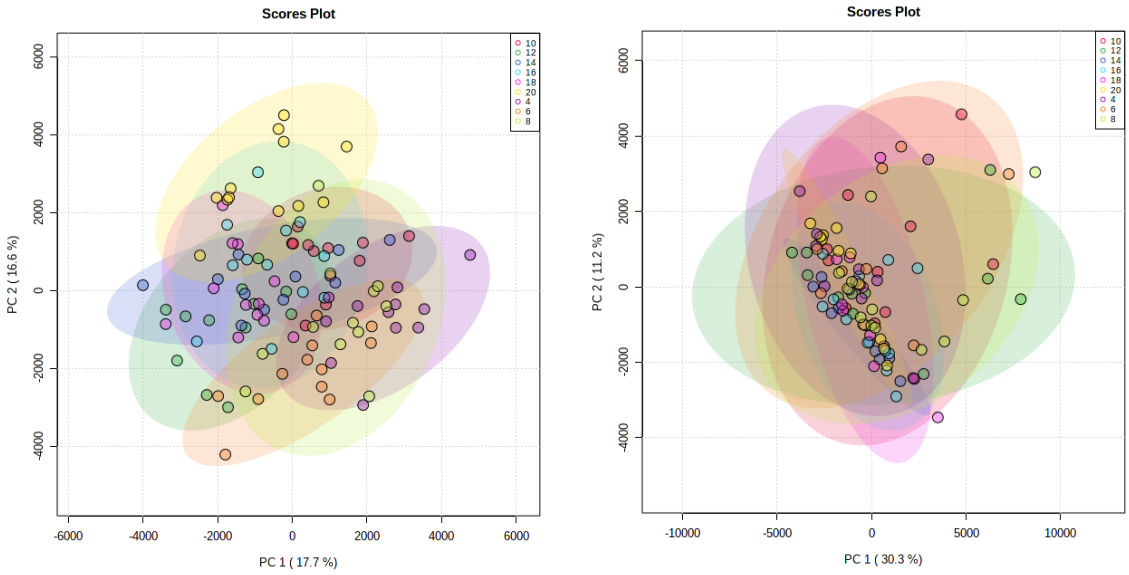

Figure 10: PCAs of Liver Primary Metabolites (left) and Plasma Primary Metabolites (right) showing clustering of birds by Day. The plasma was more tightly clustered than the liver, and there were fewer compounds that differed significantly between two or more time points, suggesting that the liver is buffering plasma compounds over the time course.

Table 3: Top 25 compounds for each tissue that were significantly different between two or more time points. The top 5 contrasts are given, in decreasing order of magnitude, with the day exhibiting higher abundance appearing first.

Liver Compound	Adjusted p-value (FDR)	Top Contrasts
retinal	<0.0001	4 - 10; 6 - 10; 4 - 12; 6 - 12; 8 - 12
palmitoleic_acid	<0.0001	10 - 20; 4 - 10; 6 - 10; 8 - 10; 4 - 12
malic_acid	<0.0001	16 - 10; 18 - 10; 20 - 10; 20 - 12
fucose	<0.0001	12 - 10; 14 - 10; 16 - 10; 18 - 10; 20 - 10
fumaric_acid	<0.0001	16 - 10; 18 - 10; 20 - 10; 20 - 12; 18 - 14
adenine	<0.0001	12 - 10; 14 - 10; 16 - 10; 18 - 10

Table 3 continued

succinic_acid	<0.0001	20 - 10; 20 - 12; 20 - 14; 20 - 16; 20 - 18
palmitic_acid	<0.0001	4 - 10; 6 - 10; 8 - 10; 4 - 12; 6 - 12
aspartic_acid	<0.0001	20 - 10; 10 - 4; 10 - 6; 12 - 14; 12 - 4
uracil	<0.0001	10 - 16; 10 - 20; 4 - 10; 6 - 10; 12 - 20
glucose-6-phosphate	<0.0001	20 - 10; 20 - 12; 12 - 4; 12 - 6; 20 - 14
threonic_acid	<0.0001	12 - 10; 14 - 10; 16 - 10; 18 - 10; 20 - 10
myristic_acid	<0.0001	4 - 10; 6 - 10; 8 - 10; 4 - 12; 6 - 12
isoleucine	<0.0001	12 - 10; 18 - 10; 6 - 10; 8 - 10; 12 - 14
glycerol-alpha-phosphate	<0.0001	20 - 10; 10 - 4; 10 - 6; 14 - 12; 16 - 12
hexose-6-phosphate	<0.0001	20 - 10; 20 - 12; 12 - 4; 12 - 6; 20 - 14
leucine	<0.0001	12 - 10; 14 - 10; 18 - 10; 6 - 10; 8 - 10
methanolphosphate	<0.0001	20 - 10; 16 - 12; 20 - 12; 20 - 14; 14 - 4
ribonic_acid	<0.0001	14 - 10; 16 - 10; 20 - 10; 14 - 12; 16 - 12
1-monopalmitin	<0.0001	10 - 18; 4 - 10; 6 - 10; 4 - 12; 6 - 12
glucose	<0.0001	18 - 10; 20 - 10; 10 - 4; 14 - 12; 18 - 12
valine	<0.0001	12 - 10; 18 - 10; 6 - 10; 8 - 10; 12 - 14
fructose-6-phosphate	<0.0001	12 - 10; 16 - 10; 20 - 10; 20 - 12; 12 - 4
tyrosine	<0.0001	12 - 10; 18 - 10; 10 - 20; 4 - 10; 6 - 10
nicotinamide	<0.0001	12 - 10; 14 - 10; 16 - 10; 18 - 10; 12 - 20
Plasma Compound	**Adjusted p-value (FDR)**	**Top Contrasts**
threonic_acid	<0.0001	14 - 10; 16 - 10; 18 - 10; 20 - 10; 14 - 12
cholesterol	<0.0001	10 - 20; 14 - 12; 16 - 12; 18 - 12; 12 - 20
phenylethylamine	<0.0001	14 - 10; 16 - 10; 18 - 10; 14 - 12; 16 - 12
retinal	<0.0001	4 - 10; 4 - 12; 4 - 14; 16 - 20; 4 - 16
4-hydroxybutyric_acid	<0.0001	20 - 10; 8 - 10; 20 - 12; 8 - 12; 20 - 14

Table 3 continued

hydroxylamine	<0.0001	20 - 10; 20 - 12; 20 - 14; 20 - 16; 20 - 18
valine	<0.0001	10 - 16; 10 - 18; 4 - 10; 6 - 10; 12 - 14
glyceric_acid	<0.0001	16 - 10; 18 - 10; 20 - 10; 16 - 12; 20 - 12
phenylalanine	<0.0001	4 - 10; 6 - 10; 6 - 12; 6 - 14; 4 - 16
mannose	<0.0001	14 - 10; 16 - 10; 18 - 10; 20 - 10; 16 - 12
inosine	<0.0001	10 - 20; 4 - 10; 4 - 12; 6 - 12; 8 - 12
isothreonic_acid	0.0001	4 - 10; 8 - 10; 4 - 12; 8 - 12; 4 - 14
putrescine	0.0001	4 - 10; 6 - 10; 8 - 10; 4 - 12; 6 - 12
tocopherol_gamma-	0.0002	14 - 10; 10 - 20; 10 - 6; 10 - 8; 14 - 12
3-hydroxybutyric_acid	0.0002	10 - 6; 10 - 8; 12 - 14; 12 - 20; 12 - 4
myo-inositol	0.0003	16 - 10; 10 - 4; 14 - 12; 16 - 12; 12 - 4
methionine	0.0003	20 - 10; 12 - 4; 12 - 8; 14 - 4; 14 - 8
alanine	0.0003	20 - 10; 4 - 10; 6 - 10; 12 - 16; 20 - 12
uric_acid	0.0003	10 - 20; 4 - 10; 8 - 10; 12 - 20; 4 - 12
oxoproline	0.0005	14 - 10; 16 - 10; 20 - 10; 14 - 12; 20 - 12
phosphate	0.0006	12 - 10; 14 - 10; 16 - 10; 18 - 10; 20 - 10
glycyl_proline	0.0008	14 - 10; 16 - 10; 10 - 6; 14 - 12; 16 - 12
conduritol-beta-epoxide	0.0009	10 - 18; 10 - 20; 4 - 10; 4 - 12; 8 - 12
beta-alanine	0.0010	20 - 10; 20 - 12; 20 - 14; 20 - 16; 20 - 18
citric_acid	0.0018	16 - 10; 18 - 10; 20 - 10; 16 - 12; 18 - 12

3.2.2 WGCNA

35 WGCNA modules were created, ranging in size from 33 to 2307 features. Of these, 11 contained both genes and metabolites, 18 contained only genes, and 5 contained only metabolites. Many modules showed high correlation with physiologic

traits (Figure 11), but we chose to focus on four blood chemistry measures related to metabolism that showed fluctuations around D6 and D16 (Figure 12). All measures (pH, BE, HCO_3, Glu) tend to increase with bird age and body mass, but show slight fluctuations in levels at one or more points during the time course, suggesting shifts in metabolism. Blood pH is a measure of acid-base balance, of which base excess (BE) and HCO_3 represent the metabolic component. These two measures were noted in the previous analysis to be significantly different between D4 and D20, with little change in respiratory components of acid-base balance. Glucose (Glu) was chosen for its direct relation to the liver's metabolism.

12 WGCNA modules, summarized in Table 4, were selected for export to Cytoscape (82), containing a total of 946 features. Molecular Complex Detection (MCODE) is an algorithm that identifies subclusters within a network based on connectivity (83). When used on protein-protein interaction networks, it can identify groups that may identify complexes or pathways, but it also commonly used on WGCNA networks (84,85). Although WGCNA networks do not explicitly represent protein-protein interactions, the correlation-based clustering of features has powerful potential to identify functionally related groups of features, relate these to physiologic traits, and quantify network-based measures of relatedness based on the Topological Overlap Measure (TOM). MCODE identified 28 clusters within the Cytoscape network. While most genes and metabolites were segregated in separate clusters, there were several clusters containing features from multiple different WGCNA modules, and one containing both genes and metabolites.

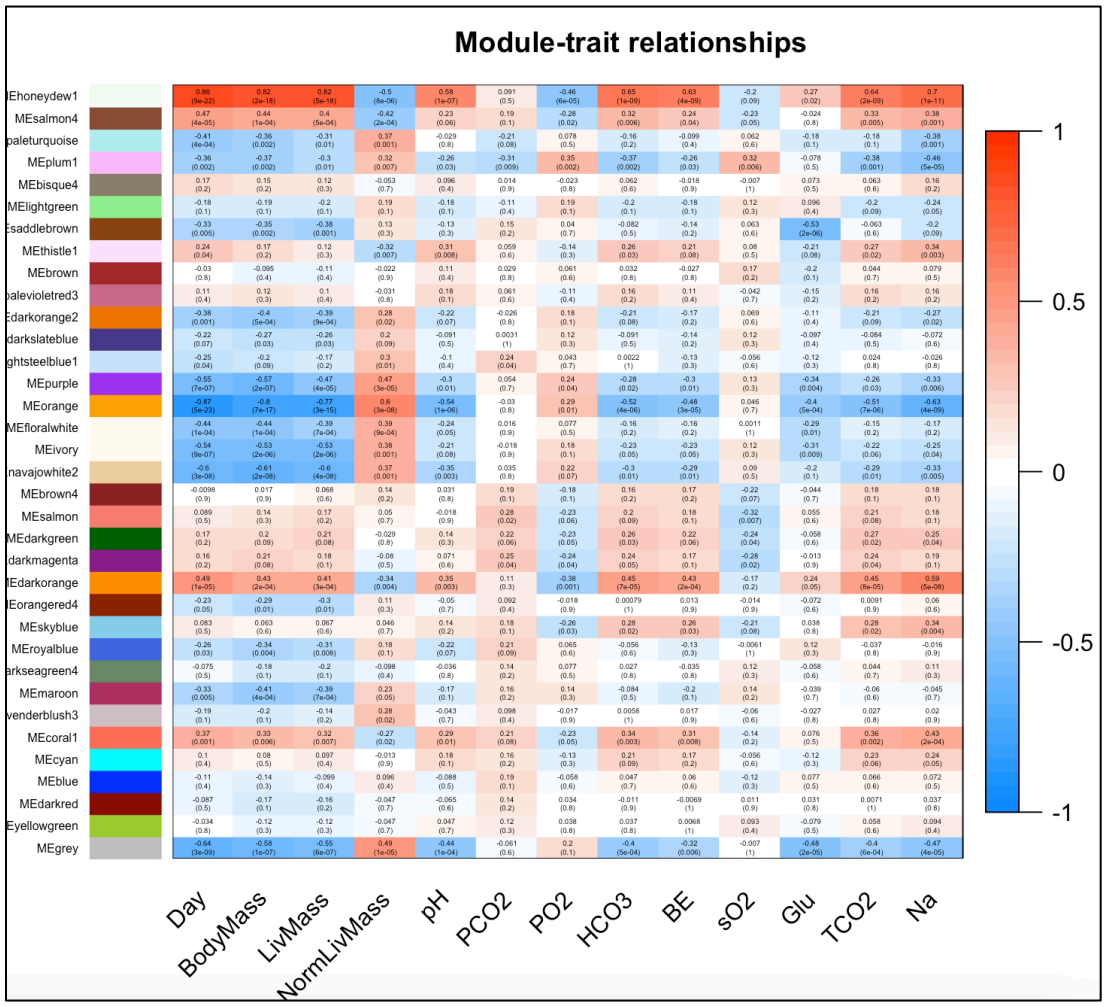

Figure 11: Correlations of each module with physiologic traits. The modules, identified by the color bars on the left, represent clusters of features that exhibit synchronized patterns of abundance throughout the time course. The heat map shows correlation of each module eigengene (ME) with physiological traits, including organ masses and blood chemistry measurements. WGCNA modules were selected based on quality measures, functional enrichment, and correlation with four traits of interest: Blood pH, HCO3, BE, and Glu.

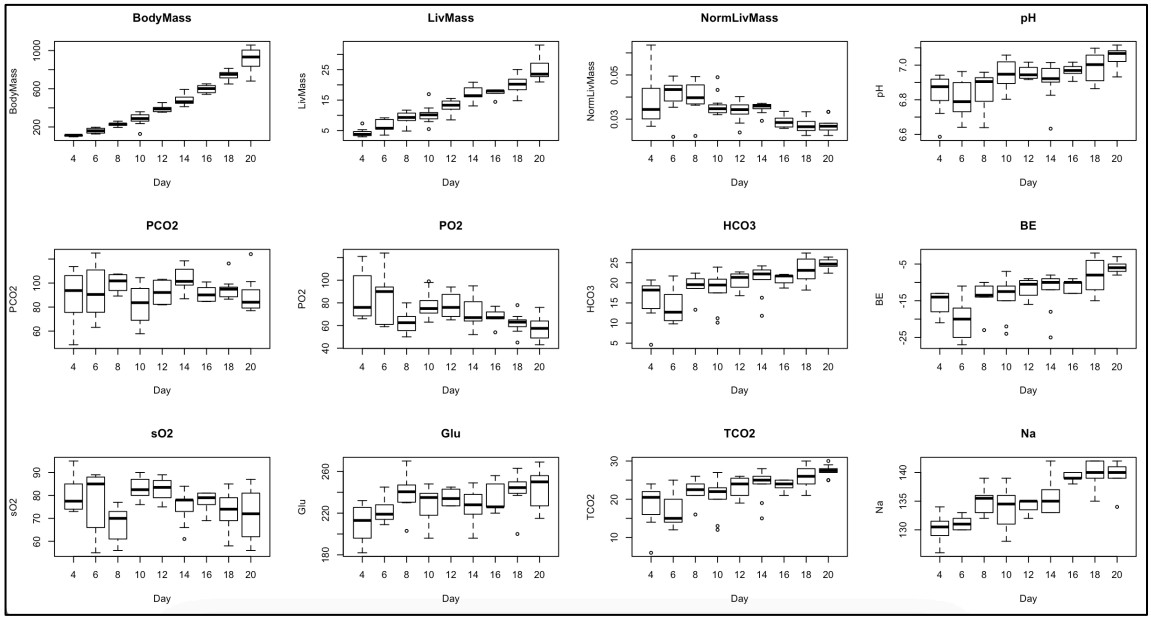

Figure 12: The metabolism-related blood chemistry measures that influenced WGCNA module choice, including pH, base excess (BE), bicarbonate (HCO3), and glucose (Glu). While many measures were strongly correlated with Body Mass and Day, these four measures showed fluctuations during the intervening days of the time course, suggesting they may be related to more subtle changes in metabolism.

Table 4: Summary of 12 WGCNA clusters chosen for export to Cytoscape, along with notable enrichment and features. Hub feature is the feature with the top Module Membership value, or correlation with the Module Eigengene.

WGCNA Cluster	Total features, type	Functional enrichment	Other notable features	Hub feature
coral1	33 genes	Metabolism	GC, ADH5, SORD, PAH	Autocrine Motility Factor Receptor (AMFR)

Table 4 continued

floral white	67 genes	Collagen biosynthesis, ECM, Protein digestion & absorption, PDGF binding, TGF	LGALS1, FSTL1, TGFB3	Serpin Family H Member 1 (SERPINH1) - collagen biosynthesis
honey dew1	45 genes and metabolites	Immunoglobulin, TCA & sugar metabolites	JCHAIN, TCA &, carbohydrate metabolites	Immunoglobulin lambda like polypeptide 1 (ENSGALG00000049450)
ivory	68 genes and metabolites	Blood coagulation, Angiogenesis, Integrin signaling, Epithelial cell proliferation	TGFBR3, VWF, HK1, aspartate	Tyrosine Kinase With Immunoglobulin Like And EGF Like Domains 1 (TIE1) - angiogenesis
navajo white	54 genes	Vasculogenesis	FABP5	Transmembrane Protein 204 (TMEM204) - cell adhesion
orange	107 genes and metabolites	AP-1 complex, gas transport/hemoglobin, PPAR signaling, immune (T cell, B cell, IL)	FOS, JUN, PPARD, CTGF, CYR61, retinal, plasma putrescine, fatty acids, vitamin E	Translocator Protein 2 (TSPO2) - cholesterol binding during erythrocyte maturation

Table 4 continued

plum1	85 metabolites	Glutamine/glutamate metabolism, Amino acid metabolism	Liver: lactamide, glutamine, erythronic acid lactone, glutamate; Plasma: phosphate	Unidentified liver metabolite (120781)
purple	287 genes and metabolites	DNA replication, Cell cycle, p53 signaling, Pyrimidine metabolism	E2F8, glucose, ethanolamine	Nucleolar And Spindle Associated Protein 1 (NUSAP1)
saddle brown	103 genes and metabolites	Phosphatase and hydrolase (ester) activity, Response to insulin stimulus	MYC, LPIN1, IGFBP1, LIPG, G6PC	WT1 Associated Protein (WTAP)
salmon4	55 metabolites	Glutamine/glutamate metabolism, Amino acid metabolism	Plasma: lactamide, oxoproline, gluconate, galactonic acid, glutamine	Unidentified plasma metabolite (208646)
skyblue	105 genes	Complement, immunity, EGFR, Acylglycerol metabolism	DGAT2, CD48	Cathepsin S (CTSS) - immune, ECM remodeling
thistle1	56 genes and metabolites	Fatty acid metabolism, PPAR signaling, Insulin resistance, AMPK signaling	GYS2, EGFR, IGFBP2, AGPAT4, Plasma fatty acids (palmitic, arachidonic)	Carnitine Palmitoyltransferase 1A (CPT1A) - mitochondrial oxidation of LCFA

3.2.3 Summary of selected metabolism-related WGCNA modules

Each WGCNA module is formed around a Module Eigengene that represents the first principal component and summarizes the abundance profile of the features in the module. Each feature has a Module Membership value and associated p-value representing its correlation with the Module Eigengene. Features with very high Module Membership (>0.8) are most representative of the Module Eigengene abundance profile, and are considered hub features that are likely to be important, highly connected within the network, and may be drivers of other features in the module. The abundance profiles of the hub features for each of the 12 selected WGCNA modules are shown in Figure 13, and a summary of each of these modules is shown in Table 4.

It is important to note that an unsigned WGCNA network was chosen for this analysis, due to the biological relevance of both positive and negative relationships between features and traits. However, this does complicate interpretation, as some features belonging to a cluster may have a negative Module Membership value, and an inverse relationship to the Module Eigengene and therefore traits. Features with weaker Module Membership correlations may also vary significantly from the abundance profile of the Module Eigengene. Without observing the Module Membership value, or the trend of the feature over time, it is impossible to tell this, and therefore it is important not to assume that a feature's membership in a particular module is reflective of its abundance over time or correlation with a physiological trait. For this reason, WGCNA with an unsigned network is best used to select groups and features of interest, prior to downstream methods that can investigate specific trends and relationships.

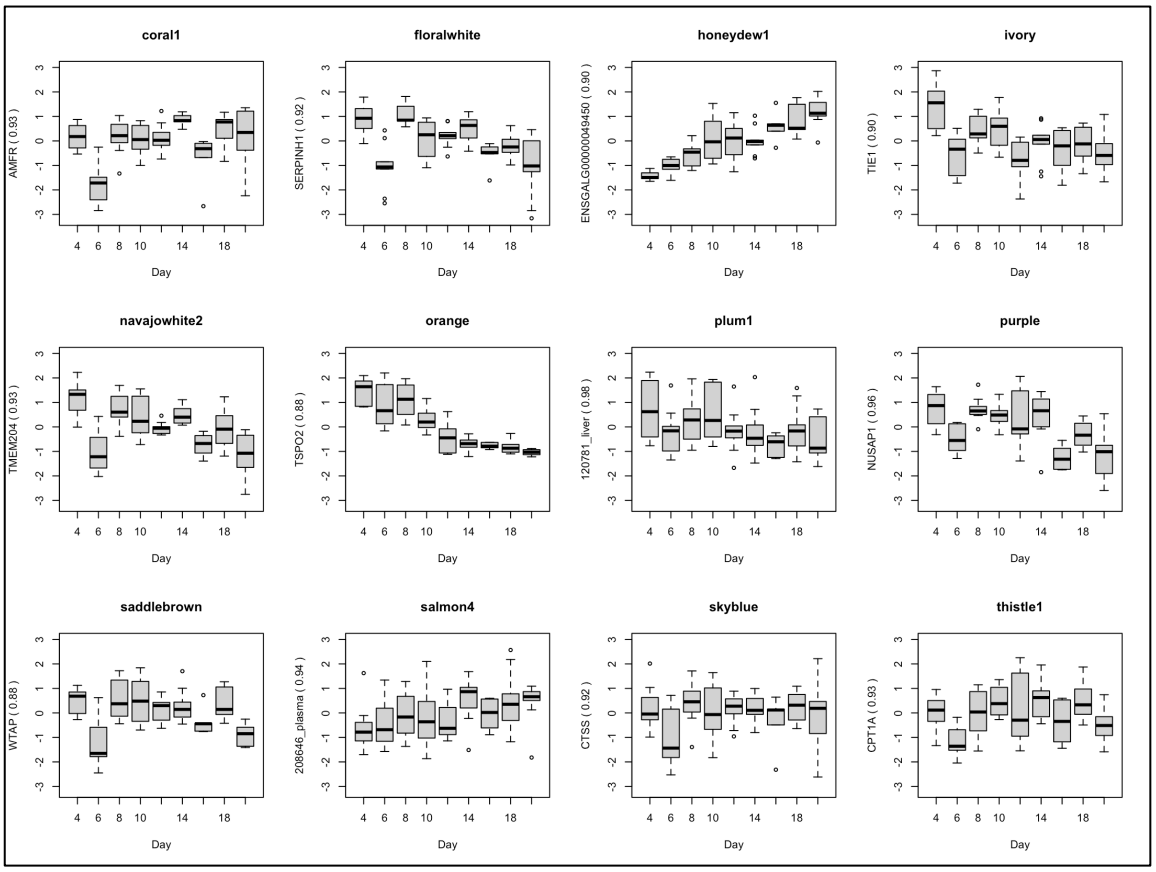

Figure 13: Hub genes for 12 selected modules over time, in logTPM and scaled by z-score. Correlations with Module Eigengene are shown next to feature name.

3.2.4 Ontology enrichment of 12 WGCNA modules chosen

The entire gene list with symbols matching human proteins (616 in total) and those matching chicken proteins (631 total) was dominated by ontology terms related to different stages of DNA replication and the cell cycle due to inclusion of the highly enriched purple module, which was also the largest containing 287 features. Other notable terms included those related to hypoxia detection, gas transport, fatty acid metabolism (especially triacylglycerol synthesis), vasculature development, T-cell differentiation, immune system development, bicarbonate transport, and collagen

biosynthesis. Top KEGG pathways were DNA replication and Cell cycle, with several signaling pathways (p53, FoxO, MAPK, PI3K-Akt), and two metabolic pathways (fatty acid degradation, protein digestion and absorption) also represented.

3.2.5 Summary of MCODE clusters:

The MCODE algorithm identifies densely connected regions within a network, that may be biologically interpreted as representing complexes or pathways (83). 28 clusters were identified within the network of 12 selected WGCNA modules; these clusters are summarized in Table 5. Cluster 1, containing genes from the purple WGCNA module that was positively correlated with normalized liver mass (0.47), and negatively correlated with bird age and metabolic measures of interest, was highly interconnected and functionally enriched for DNA replication and Cell cycle related terms. It also contained three transcription factors (ZNF367, E2F8, TFDP2). E2F8 has roles in angiogenesis and polyploidization. TFDP2 and E2F8 form a complex that regulates progression from G1 to S phase in the cell cycle. The purple module also contained glucose, although it was not included in this cluster.

Cluster 2 contained liver metabolites from the plum1 WGCNA module, which was positively correlated with NLM and negatively with metabolic measures of acid-base balance. The plum1 module was mostly isolated in the WGCNA network, sharing one connection between liver metabolite ribonic acid and two liver metabolites from the orange module (fucose and 196279). Notable metabolites also in this module included dehydroascorbic acid, lactamide, glutamine, L-gulonic acid, and erythronolactone.

Cluster 3 contained mostly genes from the thistle1 module, but also three from saddlebrown. The two modules had mostly opposite trait correlations, with

saddlebrown most negatively correlated with glucose, and thistle1 most positively correlated with pH and HCO_3. The cluster was functionally enriched for fatty acid metabolism, PPAR signaling and AMPK signaling, containing multiple important genes noted in Chapter 2 to be differentially expressed between D4 and D20, and involved in both carbohydrate and lipid metabolism. It also contained several transporters and metabolic enzymes. There were multiple interesting first-neighbor connections between members of this module and those in other modules, including coral1, purple, ivory, saddlebrown, and thistle1.

Cluster 4 contained genes from the skyblue module, which was positively correlated with HCO_3 and BE. This cluster was mostly isolated, with no first neighbors from outside modules except for FGR (ivory module), an important regulator of immune cell migration and activation. The seed for this module was SLC2A6, a glucose transporter thought to regulate glycolysis in macrophages. The cluster itself was enriched for terms related to complement cascade, macrophage differentiation, and collagen. Cluster 14 also contained genes from the skyblue module, that were enriched for terms related to immune and Th1 cells, and included transcription factors TBX21 and IRF5.

Cluster 5 contained genes from the coral1 module, which positively correlated with HCO_3, BE, and pH, and negatively correlated with NLM. The cluster was enriched for metabolism, and genes contained in this module shared connections with Cluster 6 contained genes from the ivory, navajowhite2, and floralwhite modules, which were positively correlated with NLM and negatively with pH, HCO_3, BE, and Glu. It was not strongly enriched for relevant ontology terms or pathways, according to String, but contained two transcription factors (ZNF521, ZNF423) and growth

regulator TGFBR3. Cluster 21 also contained genes from the ivory and navajowhite2 modules, and was enriched for vasculature development and HIF1 signaling. Notable features included transcription factor GATA4, two genes involved in angiogenesis and vasculature development (TEK, EGFL7), and hexose kinase HK1. Features from this cluster were also highly connected with many others from the floralwhite, saddlebrown, ivory, and orange modules. GATA4 was connected with PDGFC.

Cluster 8 contained genes from the honeydew1 WGCNA module, which was positively correlated with HCO_3, BE, pH, and Glu, and negatively correlated with NLM. It contained almost entirely immunoglobulin-related genes, including JCHAIN. This cluster was fairly isolated, with three of its members sharing a connection with liver metabolite fumaric acid, also from the honeydew1 module. Overall, the honeydew1 module was enriched for glycolysis and the TCA cycle. Cluster 17 contained multiple important TCA and glycolysis-related liver metabolites from the honeydew1 module, with malic acid as the seed. Other metabolites in this cluster were fumaric acid, H-6P, G-6P, F-6P, and GAP. Features in Cluster 17 were first neighbors with other metabolites from the honeydew1 module, including uridine, uric acid, succinic acid, two from the purple module (glucose, ethanolamine) and two from the orange module (adenine, inosine).

Cluster 9 contained both genes and liver metabolites from the orange WGCNA module, which was strongly positively correlated with NLM, and negatively correlated with pH, HCO_3, BE, and Glu. The seed for this cluster was liver metabolite retinal, and the cluster was enriched for gas transport and hemoglobin. Other features of interest were TSPO2, the hub feature for the orange module, which is involved in cholesterol binding and erythrocyte maturation, and metabolites (threonic acid, fucose,

adenine). First neighbors of these cluster members included ribonic acid (plum1 module), H-6P & G-6P (honeydew1 module), and NCAPD2 (purple module), along with others from ivory, floralwhite, and saddlebrown. Cluster 28 contained three fatty acid metabolites from the orange module (stearic acid, palmitic acid, myristic acid), which are yolk-derived fatty acids noted in Chapter 2 to decrease from D4 to D20.

Cluster 10 contained genes from the saddlebrown WGCNA module, which was the module most highly negatively correlated with Glu (-0.53). It was enriched for T-cell differentiation, and contained LPIN1, FOXP3, and PTPN2. Members of this cluster were mostly isolated, but shared connections with some features from the thistle1 cluster, including GYS2, CPT1A, and EGFR. The entire saddlebrown module was enriched for metabolism, specifically phosphatase and hydrolase activity, and also response to insulin stimulus, and contained multiple interesting genes (MYC, LPIN1, LIPG, G6PC), and the hub gene was tumor suppressor WTAP, which controls mRNA methylation, the cell cycle transition from G2/M phase, and likely also regulates splicing. Cluster 12 contained a different group of genes from the saddlebrown module, including IGFBP1, follistatin (FST), and LGALSL. One study found that in mice, knocking down hepatic FST affected glucose tolerance, regulation of hepatic glucose production by insulin, and thus may be a potential target for those controlling insulin resistance (Tao 2018). LGALSL encodes a galectin-like protein. Several other galectins were represented in the entire network, including LGALS1 (floralwhite), LGALS2 (orange), and LGALS3 (skyblue). LGALS1 in particular is involved in T-cell regulation, innate immunity, apoptosis, and cell adhesion. IGFBP1 is an insulin-like growth factor mainly expressed in the liver and responsible for binding IGF1 and IGF2, stabilizing them and regulating their interactions, and may also be related to

glucose tolerance. Cluster 24 also contained genes from the saddlebrown module, including a transcription factor involved in regulating glucose metabolism and aerobic glycolysis (FOXK1).

Cluster 11 contained genes from the navajowhite2, floralwhite, orange, and ivory modules, which were mainly positively correlated with NLM and negatively with other measures. It was enriched for Collagen related terms, and protein digestion and absorption. Features of interest included a kinase involved in angiogenesis (TIE1), an FST-related gene thought to act on growth factors that regulate endothelial cell proliferation and differentiation (FSTL1), and TGFB3.

Cluster 19 contained plasma metabolites from the salmon4 module, which was positively correlated with HCO_3 and BE, and negatively correlated with NLM. Two metabolites of note were gulonic acid and lactamide. Notably, the same compounds in the liver also grouped together (plum1 module and Cluster 2). In each case, these metabolites had opposite correlations with the module eigengenes (lactamide positive, gulonic acid negative) but were very strongly related.

Cluster 22 contained genes from the floralwhite, orange, and saddlebrown modules, which were mainly positively correlated with NLM and negatively with other measures. It was enriched for cell proliferation and MAPK signaling, and included transcription factor MYC, and PPARD.

The metabolites lactamide and oxoproline showed strong relationships to each other within each tissue type: liver and plasma. These two metabolites occurred in the same WGCNA modules for each tissue (plasma – salmon4, liver – plum1). Glutamine, gulonic acid, and galactonic acid also occurred in these same modules, but with a negative module membership. The module containing liver metabolites (plum1) was

positively correlated with the chosen blood chemistry measures of metabolic interest (HCO_3, Glu, pH, BE), while the module containing plasma metabolites (salmon4) was negatively correlated with these measures.

Table 5: Summary of the network subclusters identified by MCODE algorithm, in order of decreasing size and score. Cluster seed is the feature with the highest connectivity within each cluster, indicating possible importance or regulatory role.

MCODE Cluster	Feature Type	WGCNA modules	Cluster Seed	Functional enrichment	Features of interest
1	genes	purple	Centromere Protein Q (ENSGALG00000016692)	DNA replication, Cell cycle	E2F8, TFDP2
2	liver metabo-lites	plum1	Unidentified liver metabolite (47)	None	vitamin C, lactamide, glutamine, gluconic acid, erythronolac-tone
3	genes	thistle1, saddle-brown	SLC25A47 (Hepatocellular Carcinoma Down-Regulated Mitochondrial Carrier Protein)	Fatty acid metabolism, PPAR, AMPK	ACSL1, GYS2, G6PC, EHHADH, HADHB, CPT1A, EGFR

Table 5 continued

4	genes	skyblue	Solute Carrier Family 2 Member 6 (SLC2A6)	Complement, macrophage differentiation, collagen, VEGF signaling, immune	None
5	genes	coral1	MFS domain-containing protein (ENSGALG00000011579)	Metabolism	ADH5, PAH, SORD
6	genes	ivory, navajo-white2, floral-white	Von Willebrand Factor (VWF)	LDL, Coagulation	TGFBR3, ZNF521, ZNF423
7	plasma metabo-lites	salmon4	Unidentified metabolite (31460)	None	none, all unidentified
8	genes	honeydew	Immunoglobulin Lambda Variable 3-10 (ENSGALG00000051167)	Immune	JCHAIN, immuno-globulins

9	genes, liver metabo-lites	orange	retinal	Gas transport, Hemoglobin	TSPO2, retinal, threonate, fucose, adenine
10	genes	saddle-brown	Coenzyme Q10B (COQ10B)	T-cell differentiation	LPIN1, FOXP3, PTPN2
11	genes	navajo-white2, floral-white, orange, ivory	Aquaporin 1 (AQP1)	Collagen, Protein digestion and absorption	TIE1, FSTL1, TGFB3
12	genes	saddle-brown	Salt Inducible Kinase 1 (SIK1)	None	IGFBP1, FST, LGALSL
13	liver metabo-lites	honeydew	Unidentified liver metabolite (5990)	None	none, all unidentified
14	genes	skyblue	T-Box Transcription Factor 21 (TBX21)	Immune, Th1 cells	TBX21, IRF5
15	plasma metabo-lites	salmon4	None	none	none, all unidentified
16	genes	floral-white, ivory	Proline And Arginine Rich End Leucine Rich Repeat Protein (PRELP)	Endothelial cell proliferation and migration, ECM	None

Table 5 continued

17	liver metabo-lites	honeydew	malic acid (liver)	TCA, Glycolysis	Malate, Fumarate, H-6P, G-6P, F-6P, GAP
18	genes	skyblue	Gasdermin A (GSDMA)	Immune	LGALS3
19	plasma metabo-lites	salmon4	gulonic acid	None	Gulonic acid, Lactamide
20	genes	ivory, navajo-white2	None	None	ISYNA1
21	genes	ivory, navajo-white2	Lymphatic Vessel Endothelial Hyaluronan Receptor 1 (LYVE1)	Vasculature development, HIF1 signaling	GATA4, TEK, HK1, EGFL7
22	genes	floral-white, orange, saddle-brown	Annexin A2 (ANXA2)	Cell proliferation, MAPK signaling	MYC, PPARD
23	genes	purple	None	Cell cycle, DNA repair	USP1
24	genes	saddle-brown	Serine/ Threonine/ Tyrosine Interacting Protein (STYX)	None	FOXK1

Table 5 continued

25	genes	purple	Ankyrin Repeat And LEM Domain Containing 1 (ANKLE1)	DNA repair	None
26	genes	floral-white	None	Collagen	PDGFC
27	genes	saddle-brown	Lipase G, Endothelial Type (LIPG)	None	None
28	liver metabo-lites	orange	stearic acid	Fatty acids	stearic acid, palmitic acid, myristic acid

3.3 Discussion

3.3.1 Purple module represents strong initial cell proliferation

The hub gene for the purple module (NUSAP1, 0.96) gradually decreases with bird age, but shows peaks around D14 and D18, and was positively correlated with NLM (0.47). This module was highly enriched for transcripts involved in the cell cycle and DNA replication. Three subclusters were identified by MCODE (Clusters 1, 23, 25). Three identified liver metabolites (glucose, pantothenic acid, ethanolamine) were in this module but negatively correlated, along with one plasma metabolite (methionine). Glucose, methionine, and ethanolamine are important in cell proliferation, while pantothenic acid increases glutathione levels to protect cells from oxidative stress (86). The inverse relationship of these metabolites with bird age and

robust cell proliferation may point to metabolic utilization during proliferation, or metabolic regulation of proliferation.

Clusters of genes involved with cell cycle and DNA replication gradually decrease with bird age, but show peaks around D14 and D18, indicating possible periods of increased cell proliferation at these time points. Although this module did not show the strongest correlation with NLM, it may contain the main drivers and regulators of the liver's overall growth. In Chapter 2, we noted that cell proliferation was increased at D4 when compared with D20. Additionally, the subcluster identified by MCODE (Cluster 1) was highly interconnected and likely represents a very robust group of genes responsible for the regulation of liver growth. Metabolites (liver: lysine, conduritol-beta-epoxide, urea, threonine; plasma: 2-Glyceryl monooleate, conduritol-beta-epoxide) show more weak associations, but may be important to cell proliferation, either as building blocks or signaling molecules. Without further experiments, there is little indication of which cell types are represented by this proliferation, but we hypothesize that the majority of early proliferation is occurring in hepatocytes and epithelial cells driven by HIF1a-upregulated angiogenesis, while later peaks at D14 and D18 represent proliferating populations of immune cells.

As suggested in one study comparing liver growth in meat-type chickens with that in low-weight selected and a crossbreed (18), the liver in meat-type birds may proliferate mainly via hypotrophy from polyploidization rather than hypoplasia. The metabolic demands on meat-type birds that support increased growth potential could require this more efficient means of increasing cell volume and organ mass. Polyploidy is common in hepatocytes, and although its role is unknown, it has been proposed to confer a metabolic advantage (87,88). A critical transcription factor

stimulating polyploidization is E2F8, a member of the E2F family involved in regulation of cell proliferation and the cell cycle. They form complexes with another transcription factor, TFDP2, stimulating transcription of target genes. Both E2F8 and TFDP2 were present in the purple module and highly interconnected Cluster 1. Furthermore, RBL1, which regulates the effects of E2F on their targets, was also present in the purple module. MYC, which has also been shown to promote polyploidization in liver cells (89,90), was part of the saddlebrown module and decreased with bird age. Thus, early liver growth via polyploidization is supported in our data, and could confer a metabolic advantage to broilers during this period of increased cell proliferation. Cluster 1 may represent a tight group of genes whose function is to ensure the proliferation of cells critical to liver function and metabolism, indicating an effort to establish the liver's identity early on. Later peaks in these genes may be associated with other cells proliferating, especially immune. Glucose, also included in the purple module but with a gradually increasing trend with age, is critical in cell proliferation (91). Features within Cluster 1, those unclustered in the purple module, and those showing connections with other modules warrant further investigation to determine if their correlative relationships fluctuate at different points in the time course, suggesting possible regulatory changes.

3.3.2 Saddlebrown – linking metabolic changes with other functions

The saddlebrown module had the highest negative correlation with blood glucose (-0.53); the hub gene for this module (WTAP, 0.88) decreases very gradually with bird age, showing a slight peak at D18. This module contained genes involved in carbohydrate or lipid metabolism (LPIN1, LIPG, G6PC, HK1), cell proliferation (MYC, GADD45B) and two plasma fatty acid metabolites (palmitoleic and linoleic

acid). Several of these features were noted in Chapter 2 to be differentially expressed, an important indicator of metabolic reprogramming: G6PC was upregulated at D20, LPIN1 and LIPG were upregulated at D4. While these the patterns in abundance are likely recapitulating trends related to diet and maturing metabolism, as described in Chapter 2, the saddlebrown module goes further and attaches other features to these trends. In this way, these changes associated with metabolic reprogramming are related to more finely-tuned regulatory mechanisms or biological processes over the time course. The subclusters identified by MCODE (Clusters 3, 10, 12, 22, 24, 27), support this, as each had differing enrichment, possibly representing different functional groups contained in this module.

3.3.3 Cluster 3 - regulation of carbohydrate and lipid metabolism during metabolic reprogramming

Cluster 3 was the largest and highest-scoring group containing saddlebrown features. It was predominately comprised of features from the thistle1 module, but included three from saddlebrown (SLC25A25, G6PC, NCALD). Both modules showed slight fluctuations over the time course, but had contrasting overall trends and trait correlations. The seed for the module was a mitochondrial transporter (SLC25A47) whose substrate and mechanism is unknown, but is downregulated in hepatocellular carcinoma (92). It is thought to be a liver-specific uncoupling protein causing proton leak (or incomplete coupling of oxygen with ATP) that may help regulate the production of reactive oxygen species (93). Several other transporters were present in Cluster 3, including a mitochondrial adenine and phosphate carrier (SLC25A25), a mitochondrial coenzyme A transporter (SLC25A42), one responsible for cellular phosphate uptake (SLC20A2), a urate exporter (ABCG2), and one

catalyzing bidirectional transport on multiple types of monocarboxylates including lactate, pyruvate, ketone bodies, and derivates of branched-chain amino acids (SLC16A5). SLC16A5 was noted in Chapter 2 for its ability to export alternative energy sources for utilization by the rest of the body, ensuring the liver's own supply of glucose. Two genes were related to glucose homeostasis (GYS2, G6PC). GYS2, the rate limiting enzyme in glycogen synthesis, is transcriptionally regulated by PPARs (94). EGFR was also present, epidermal growth factor receptor with critical roles in cell proliferation. EGFR has been noted in cancer types for its regulation of metabolism, specifically glucose and glutamine utilization (95–97). Glucose-6-Phosphatase (G6PC) converts glucose-6P to free glucose for release into the bloodstream, and was noted in Chapter 2 for being upregulated at D20, a hallmark of maturing liver metabolism. Multiple genes in Cluster 3 were related to fatty acid metabolism: a carnitine palmitoyltransferase preparing fatty acids for mitochondrial uptake (CPT1A), an acyl-coA synthetase activating fatty acids (ACSL1), multifunctional enzymes involved in several steps of beta-oxidation (EHHADH, HADHB), an acyltransferase involved in lipid synthesis (AGPAT4), a peroxisomal acyl-coA oxidase (ACOX1), an acyl-coA dehydrogenase (ENSGALG00000008040), and a short-chain dehydrogenase (DHRS3).

Features in Cluster 3 that were first neighbors with those in other modules also suggest changing relationships over the time course, and warrant further investigation to determine how these correlations fluctuate in a time-dependent way. Among these were several genes encoding enzymes from the coral module and Cluster 5: one converting sorbitol to fructose (SORD), a triglyceride lipase (AADAC), a phosphatase (LHPP), a phenylalanine hydroxylase (PAH), an alcohol dehydrogenase (ADH5), and

an acyl-coA synthetase (ACSF2). First neighbors from the saddlebrown module included two plasma fatty acid metabolites (palmitoleic acid and linoleic acid), two genes from Cluster 12, discussed below, with roles in regulating metabolism and hepatic stellate cell proliferation (IGFBP1, FST), and two metabolic kinases (PDK4, PCK1). Also related was a growth regulator from the ivory module and cluster 21 (TGFBR3).

3.3.4 Cluster 12 represents control over HSC proliferation

Cluster 12 contained several notable genes (SIK1, IGFBP1, FST, LGALSL, CHAC1), that together may represent regulation of glucose metabolism, and control over proliferation and activation of different cell types. These genes were most abundant at D4, decreasing with bird age but showing small increases around D8 and D16. The seed for this cluster, SIK1, regulates hepatic lipogenesis through phosphorylating SREBP-1c (98) and also gluconeogenesis through suppressing gluconeogenic genes (99). IGFBP1 and FST may have roles in metabolically regulating the type of growth occurring. IGFBP1 regulates IGF 1 and 2, growth factors with known roles in growth and development. IGFBP1 can either inhibit or stimulate these effects, although it has been shown to promote cell migration. Follistatin (FST) has been linked to glucose tolerance in mice with metabolic abnormalities (100). Another study found that FST, through inhibiting activin, controls the proliferation of hepatic stellate cells and prevents hepatocyte apoptosis, thereby controlling hepatic fibrogenesis (101). The abundance of FST and IGFBP1 may control hepatic stellate cell proliferation and activation, to regulate ECM deposition and cell migration. LGALSL encodes a galectin-like protein (GRP). CHAC1 degrades glutathione and promotes apoptosis (102). Hepatic stellate cells have many roles, including storage of

vitamin A when quiescent (103), stimulating the proliferation of hepatocytes and other liver cells in response to energy, producing extracellular matrix (104), and hypoxia-sensing (105). When activated in response to liver injury, they release vitamin A and contribute to fibrogenesis. FST, and FSTL1 as discussed below, may be mediating hepatic stellate cell activation and proliferation to control inflammatory processes and expansion of the immune system.

Cluster 22 contained genes from floralwhite, orange, and saddlebrown modules, including two with regulatory roles in metabolism and cell proliferation (PPARD, MYC). MYC is a transcription factor involved in cell proliferation, and is involved in normal liver growth but also hepatocyte proliferation during regeneration, where it promotes the transition from G1/G0 to S phase (89). Metabolically, MYC has been shown to directly target and upregulate almost all glycolytic genes (96,106), while PPARD, widely known as a regulator of lipid metabolism through PPAR signaling, also directly regulates rate-limiting glycogen synthase GYS2 (94). GYS2 was related to the saddlebrown module through cluster 3. EGFR, also related to the saddlebrown module through Cluster 3, may affect glycolysis through its regulation of the MYC pathway; inhibition of EGFR was shown to decrease glycolysis in cancer cells (96). GADD45B, although unclustered, is also involved in hepatocyte proliferation, serving protective functions to promote cell survival and proliferation during different types of growth (107). The sustained but more gradual decrease of MYC and GADD45B, in contrast with the sharper drop in the purple module strongly linked to cell cycle and DNA replication, may represent a more steady, prolonged regulation of liver cell proliferation once the initial liver growth and establishment of the organ's functions has reached a critical threshold.

Four clusters (10, 12, 24, and 27) contained only genes from the saddlebrown module. Cluster 10 contained three genes associated with T-cells (FOXP3, PTPN2, CYLD) as well as one involved in multiple steps of fatty acid metabolism (LPIN1). This cluster may be related to lipid metabolism during development of T-cells. It also contained WTAP, a tumor suppressor and hub gene of the saddlebrown module. Cluster 24 contained endothelial lipase G (LIPG), involved in fatty acid metabolism, and also FOXK1, a transcription factor critical to metabolic reprogramming that promotes aerobic glycolysis through its upregulation of multiple glycolytic enzymes (108).

3.3.5 Regulation of oxygen homeostasis

The orange module was most strongly correlated with NLM (0.6), and negatively with other measures. Its hub gene (TSPO2, 0.88) showed a steady decrease with bird age. MCODE identified three subclusters from the orange module with differing functional enrichments including cell proliferation (Cluster 22, discussed previously), and fatty acids (Cluster 28). A number of other notable features were unclustered (PODXL, EZR, FOS, JUN). FOS and JUN make up the multi-functional AP-1 complex. PODXL and EZR are known markers for progression of some cancer types because of their associated with migration and morphogenesis (109). EZR was first neighbors with transcription factor E2F8 from the purple module.

Cluster 9 was the strongest MCODE subcluster containing orange features, and includes both genes and metabolites. The seed of this cluster was liver metabolite and antioxidant retinal, which was most abundant at D4, decreased steadily until D12 and then remained fairly constant until D20. The presence of retinal may be related to mediating oxidative stress, however one study found that while vitamin C and E may

have protective roles, vitamin A in particular increased erythropoietin production, and also may increase levels of HIF1A (110). Cluster 9 contained genes encoding for multiple components of hemoglobin (HBE1, HBM, HBA1, HBBA) and also erythrocyte-related genes (TSPO2, RHAG, EPB42, SLC4A1, ANK1). Four of these (SLC4A1, RHAG, EPB42, ANK1) are erythrocyte membrane components, with SLC4A1 involved in anion transport and RHAG involved in transport of ammonia. TSPO2 and is involved in cholesterol metabolism during erythrocyte differentiation (111). Three of the metabolites (threonic acid, fucose, adenine) were inversely related to the module eigengene, increasing gradually with bird age. This cluster may represent an increased population of erythrocytes in the liver during early post-hatch growth, that decreases with bird age. This could point to mechanisms regulating oxygen homeostasis that are especially critical during the first week post-hatch, when an increased rate of cellular proliferation is occurring. Controlling oxygen levels may serve in stabilizing the hypoxia response, or may be related to mediating oxidative stress through increased binding of reactive oxygen species. The biological importance of this is discussed further in Chapter 4, which corroborates the relationship of hemoglobin, erythrocytes, and oxidative stress to normalized liver mass.

3.3.6 Clustering of metabolites identify possible relationships

Most metabolites clustered separately in MCODE analysis, likely due to unavoidable differences in feature scale. However, even metabolite-only clusters show biological relationships, recapitulating known function and also suggesting the power of this method to infer unknown relationships or classify unidentified compounds by relation to known metabolites. For example, gulonic acid and lactamide were two metabolites that were strongly but oppositely related in both liver and plasma; they

occurred together in both WGCNA module and MCODE cluster. Gulonic acid is an intermediate of pentose and glucuronate interconversions and ascorbate metabolism, while lactamide is derived from lactic acid. Lactamide also occurred in the same module as oxoproline, both for liver and plasma. Oxoproline is an amino acid derivative associated with glutamine and glutathione metabolism.

Another example is the plum1 module, made up of all liver metabolites, and positively correlated with NLM. The hub metabolite for this module (Unidentified – 120781, 0.98) decreases gradually with bird age, with several small peaks. About half of plum1 features grouped into MCODE Cluster 2. Most of these were unidentified, but several were identified including dehydroascorbic acid, lactamide, glutamine, L-gulonic acid, and D-erythronolactone. Dehydroascorbic acid and L-gulonic acid are both involved in ascorbate metabolism. Dehydroascorbic acid is the oxidized form of Vitamin C (ascorbic acid), and is transported by glucose transporters where it is reduced in normal cells (112,113). One study suggested that the liver's role in recycling dehydroascorbic acid and releasing could be important in maintaining plasma levels of ascorbic acid, and dealing with oxidative stress (114). The levels of dehydroascorbic acid and ascorbic acid increased gradually with bird age, while levels of these metabolites in plasma did not change appreciably, possibly demonstrating the liver's role in buffering plasma levels. The juxtaposition of these known metabolites alongside unknowns may help with the prioritization of these compounds for subsequent identification and research into function, and may inevitably represent biomarkers.

3.3.7 Rise of B-cells is concurrent with increased carbohydrate availability

The honeydew module was characterized by gradual increase with bird age, and contained many metabolites, including a group of carbohydrates and TCA intermediates (Cluster 17) and a group of immunoglobulin-related genes (Cluster 8). Cluster 8 likely represents an increase in antibody-producing B-cells in the liver. Mature liver harbors many types of immune cells, which serve a variety of functions including protection against pathogens, triggering local or systemic immune response, and tissue-sensing roles such as identifying inflammation or injury (115,116). Young broilers still benefit from maternal antibodies contained in the yolk remnant, but this is gradually depleted within the first week post-hatch, so the liver must begin developing its own immune system, both recruiting non-resident cells and allowing resident cells to proliferate. The gradual increase in this cluster of immunoglobulin-related genes may be stimulated both by the depletion of maternal antibodies and the consumption of oral nutrition.

The metabolites present in the honeydew module may be related to immune cell metabolism, but are likely a coincidence of similar trends. The increase in carbohydrates and TCA metabolites as the birds age, first noted in Chapter 2, is representative of both the liver's maturing metabolism and increased reliance on a carbohydrate-based diet. Cluster 17 contains phosphorylated sugars (H-6P, G-6P, F-6P) glycolysis intermediates (glycerol-alpha-phosphate) and TCA intermediates (fumarate, malate). Immune cells do undergo metabolic reprogramming upon activation (117,118). Although liver metabolite succinate was also contained in the honeydew module, and is associated with this metabolic reprogramming (119), there are no other signs indicating robust activation of immune cells. However, immune cells depend on glucose for proliferation, and activation of immune cells increases

glucose metabolism (120,121), suggesting a possible metabolic control over the proliferation of immune cells in the liver. HIF1A controls a number of different functions in immune cells through its effects on metabolic genes (122), and is also known as the main regulator of metabolism in B-cells, while MYC is the main regulator of transcription (123).

3.3.8 Clusters facilitating and indicating hepatic expansion of immune system

The skyblue module hub gene showed only subtle fluctuations throughout the time course (CTSS, 0.92), yet this module contained three highly connected regions identified by MCODE and enriched for immune-related terms including differentiation and migration. CTSS is a protease expressed by a number of immune cells, but has roles in regulating inflammation and also contributing to extracellular matrix (ECM) remodeling (124). CTSS was part of Cluster 4, along with genes related to macrophages (SLC2A6, NRROS), T-cells (BTK, HCLS1), cell migration (MERTK, FUT7, FES, CD48, PLCG1, CCL24, ANG, BIN2), and collagen (C1QC, C1QA, C1QB). The cluster seed was SLC2A6, a lysosomal glucose transporter affecting glycolysis in macrophages (125). Cluster 14 contained a chemokine (ENSGALG00000014585), an interferon regulatory factor (IRF5) and was driven by transcription factor TBX21. TBX21 is a T helper cell-specific gene involved in development. Cluster 21 also contained macrophage-related genes (CSF1R), an interleukin receptor (IL21R). Overall, the skyblue module may represent regulation of immune cell expansion in the liver that takes over once the liver cells have completed their initial rapid burst of proliferation.

Clusters 6, 11, 21, and 26 contained features from mixed WGCNA clusters, mainly navajowhite2, floralwhite, ivory, and orange. Except for the orange module,

which showed a strong decreasing trend with bird age, these modules exhibited subtle fluctuations. Together, these clusters were enriched for collagen-related terms, coagulation, PDGFC binding, and angiogenesis, suggesting roles in liver structural formation and core functions. In Cluster 11, two genes were related to vascularization: TIE1 plays an important role in angiogenesis through inhibiting angiopoetin 1, while FSTL1 promotes endothelial development (126). FSTL1, similar to FST in Cluster 8, is also known to suppress proliferation of hepatic stellate cells (127)s. One metabolic gene present in the ivory module and Cluster 21 was hexokinase HK1, which showed decreasing expression with bird age, and is predominately expressed in erythrocytes (128). These clusters of features, along with those in skyblue, likely represent a plethora of functions that occur once the main burst of liver cell proliferation is slowing. The enrichment of collagen and ECM-related genes, along with those related to immune cell migration, proliferation, and differentiation, suggest that the liver is being structurally prepared for immune cell expansion and colonization.

3.3.9 Regulation of glucose availability plays a greater role in liver development

In Chapter 2, we described metabolic reprogramming between D4 and D20 post-hatch. We hypothesized that HIF1A is regulating carbohydrate metabolism to allow increased cell proliferation under oxygen-limited conditions, causing a Warbug-like effect at D4, and contrasting with upregulated TCA cycle and mature liver metabolism at D20. We also proposed the metabolic conditions at D4 may lead the liver to conserve glucose and glycolytic intermediates for its own needs, releasing alternative energy sources such as lactate and short-chain-fatty-acids for use by the rest of the body. Glucose uptake increases from hatch in young birds, as the gastrointestinal system develops, absorptive capabilities increase, and the birds

consume a carbohydrate-rich diet (7). The maturation of liver metabolism, increased availability of carbohydrate nutrients, and release from hypoxia is evident in metabolism at D20. The metabolite and blood chemistry data across the time course support this as well: both blood and liver glucose gradually increase with bird age. The increases in glycolysis, gluconeogenesis, and glycogen metabolism noted at D20 also support the abundance of carbohydrate nutrients, and the liver's enhanced ability to release or store excess nutrients. However, regulation of glucose availability may play an even more critical role in growth of the liver than previously thought.

Glucose is crucial in cell proliferation, and as an energy source in anaerobic glycolysis observed at D4. HIF1A promotes this, allowing for ATP production and also stimulating angiogenesis to prepare the liver for eventual aerobic metabolism. The shift in metabolism, however, may not indicate reaching a threshold of vascularization and oxygen availability, but instead a coordinated mechanism of metabolic regulation that frees up carbohydrate resources to allow another critical process to occur: proliferation of the immune system. We hypothesize that HIF1A and other factors regulating metabolism in the liver play a multifaceted role in keeping the immune system, especially inflammatory processes, under control until the liver's identity is established.

3.3.10 Benefits and challenges of WGCNA as a method for data integration

The identification of modules of correlated features, and relation to measured traits with WGCNA is a powerful tool for dimension reduction and feature prioritization, but can still yield modules that are unwieldy or difficult to interpret. Visualizing the network, with nodes representing features and edges representing the strength of the relationship between their abundance profiles (based on TOM),

exposed closely related features both within and across modules. The use of the MCODE algorithm in conjunction with this provides a way to reduce these groups even further by analyzing the connectivity of the network itself, identifying highly-connected subclusters that may represent biological pathways or complexes. Computationally predicted relationships do not always hold true to biological meaning, however in this analysis there were many examples where they did. For example, Cluster 8 contained JCHAIN, and almost entirely immunoglobulins. Similar effects were observed with metabolites, such as Cluster 17 which contained TCA metabolites and sugars. The ability of these algorithms to identify related biology adds confidence when extrapolating to genes and metabolites of unknown function, and hints at possibly novel relationships. WGCNA modules, the WGCNA network edges, and MCODE clusters provide multiple ways to locate unknown genes and metabolites in context of potential biological functions, by relating them to those with known function.

WGCNA is typically used on one type of data, as it can be highly sensitive to variation including sample outliers and differences in feature scale. Even when using with purely transcriptome data, great care must be taken to normalize, remove batch effect and other unwanted sources of variation. Therefore, a major challenge in this analysis was appropriate normalization and scaling of the disparate metabolome and transcriptome datasets, in preparation for integration. Although differences in feature scale could not be entirely removed, the fact that many WGCNA modules contained both genes and metabolites was indicative of the success of this scaling. Additionally, functional enrichment and biological interpretation provided further support for the validity of the modules. As the ability of the methods to identify biologically

meaningful relationships has been established, future analyses will explore different types of networks such as signed networks. Additionally, WGCNA will be done separately on transcriptome and metabolome to examine relationships that may have been missed by shortcomings in the feature scaling methods or noise due to integration of different data types. As WGCNA was done mainly for data integration and a means of feature selection based on biological knowledge, it was considered an exploratory analysis in this work.

3.4 Conclusions

WGCNA integrating liver transcriptome, and liver and plasma primary metabolome yielded a plethora of information about clusters of features with coordinated abundance profiles across the time course, and additionally linked some of these with interesting physiologic traits, including body mass, normalized liver mass, and blood chemistry measurements. WGCNA, along with MCODE, also demonstrated the ability to identify biologically related groups of features from correlations. Many other clusters and relationships were not investigated, but there is great potential in these methods and results generated by the current analysis to explore the dataset even further, and many opportunities for downstream methods that can help to characterize the nature of these relationships in the context of regulatory networks or modeling.

This analysis built on hypotheses from Chapter 2, the comparison of D4 and D20 liver transcriptome and liver primary metabolome, and contributed to our understanding of the complex fluctuations occurring as the liver develops. In particular, it provided further evidence that control of metabolism by hypoxia at D4 not only ensures efficient energy production during rapid tissue growth, but plays a

critical regulatory role in proliferation of the immune system. Maintaining hypoxic regulation of metabolism in early post-hatch growth allows liver cells to proliferate and sequester glucose. After the first week post-hatch, the liver's metabolism shifts to a more mature state marked by upregulation of the TCA cycle, gluconeogenesis, and glycogen metabolism. This transition, likely driven by the liver reaching a threshold of mass or metabolic functionality, along with increased availability of glucose due to dietary carbohydrates and improved absorption, is a metabolic trigger for the rise of the immune system in the liver. Proliferation and tissue health during the first week post-hatch is protected by several mechanisms: regulation of glucose supply and utilization, regulation of oxygen availability, and increased factors combatting oxidative stress. Hypoxia and other mechanisms suppress inflammation and proliferation of cell types such as hepatic stellate cells.

3.5 Methods

3.5.1 Transcriptome and metabolome preprocessing

Preprocessing and normalization of datasets, in preparation for integration, was done according to recommendations by WGCNA package authors (81), in particular to minimize noise and sample-sample variability, stabilize variance, and scale features. This was considered highly important due to integration of different data types (transcriptome and metabolome) that were on different scales.

Transcriptome samples were rigorously filtered for quality, including correlation within day groups and rest of dataset; 99 samples were retained. Transcriptome data was filtered to include genes with expression of at least 0.1 TPM in at least 10 samples, and then by interquartile range (IQR) to retain genes in the top

75% of variance, retaining 11261 genes. Batch effect correction using sva/combat was not elected for transcriptome, as samples were not appropriately randomized, and to avoid introducing confounding bias.

Primary metabolome samples from liver and plasma were filtered separately to exclude those present in less than 4% of samples for each tissue, retaining 532 liver compounds and 521 plasma compounds. Filtering was then done separately for each organ using Metaboanalyst (56–58): by median to exclude metabolites with very small values, and IQR to exclude those with near-constant values. 399 liver and 391 plasma metabolites were retained. Transcriptome, primary liver metabolome, and primary plasma metabolome datasets were merged by overlapping samples, for a total of 83 samples and 12051 features. Combined dataset was variance-stabilized by log2(x+1) transformation, and features scaled and centered by conversion to z-scores. Transcriptome and metabolome data was combined in a meta-analysis, as there were not enough replicates per time point to generate individual networks for comparison.

3.5.2 ANOVA of liver and plasma metabolome

Statistical analysis was done using Metaboanalyst on each tissue (liver and plasma) separately after filtering described previously: Pareto normalization (mean-centering, then dividing by the square root of standard deviation for each metabolite), followed by PCA. Parametric ANOVA was done using adjusted p-value cutoff (FDR) of 0.05, and Fisher's Least Significant Difference (LSD) for post-hoc analysis, to identify metabolites that were significantly different in abundance between at least two time points.

3.5.3 Integrated WGCNA and network analysis using MCODE

WGCNA analysis was done in R using WGCNA package (81). 12 outlier samples were removed based on hierarchical clustering of samples, after visual inspection and choosing a cutoff, which led to greatly improved scale-free topology. Scale-free topology (>=0.8) was met and used to choose a soft threshold of 9. The network was generated step by step, beginning with calculating a similarity matrix using median-based biweight midcorrelation for robustness to outliers. The adjacency matrix and topological overlap matrix (TOM) were calculated from this using the soft threshold power and unsigned option, due to interest in both positively and negatively correlated features. Module assignment parameters were as follows: minimum module size of 30, and deepSplit of 3. Module merging was then performed using a module eigengene (ME) dissimilarity threshold of 0.15, corresponding to correlation of 0.85, and yielding 35 modules ranging in size from 33 to 2307 features. Quality of modules was evaluated by module size, biological enrichment, quality of feature module membership. Modules were selected for further analysis by overall quality, biological interest, and containing both gene and metabolite features. 12 modules were selected based on correlation with pH, HCO_3, Glu, and BE and connectivity (from TOM) exported for Cytoscape (82) using a cutoff threshold of 0.02, for a total of 946 features. Network analysis, including identification of densely connected regions representing subnetworks, was done using Cytoscape MCODE plugin with default settings (83). WGCNA networks and MCODE subclusters were visualized in Cytoscape, especially in reference to crosstalk between different modules and clusters.

3.5.4 Functional enrichment and interpretation

Ontology and pathway enrichment was done in STRING (78) on all genes present in the 12 selected WGCNA modules, and also separately on each WGCNA module and MCODE cluster. Gene symbols and *Homo sapiens* terms were used, as these are more strongly annotated than *Gallus gallus*. Pathways were also visualized using KEGG search and color tool (59), with *Gallus gallus*-specific pathways, by inputting the list of genes (identified by NCBI ID) and metabolites (identified by KEGG compound ID) and coloring each by WGCNA module to which it belonged. Genes were also annotated using PANTHER (74,75).

Chapter 4

MODELING FEATURES ASSOCIATED WITH A RESPONSE VARIABLE: NORMALIZED LIVER MASS

4.1 Background

The final aim of this dissertation was to develop a workflow that could be used to take any high-throughput dataset through feature screening and prioritization, model development, and generation of computationally validated hypotheses based on statistically significant associations between a set of strong predictors and a response variable. High-throughput biological datasets are unwieldy by nature; even with sophisticated feature reduction and clustering methods they can be difficult to reduce to biologically meaningful groups that are small enough for modeling, without relying on prior knowledge. Additionally, complex biological relationships may not be fully explained by what we already know. While incorporating prior knowledge is necessary at some point during analysis, doing so too early can introduce bias and also overlook one of the main strengths of high-throughput data: exposing potentially unrecognized biological relationships.

Heterogeneity is inevitable in biological populations, even when working with genetically identical individuals. Being able to capture this heterogeneity and associate it with meaningful biology can be powerful and informative. In this work, one measure displaying heterogeneity was Normalized Liver Mass (NLM). The trend of this measure over time was part of the inspiration for this study, and rationale for choosing the first three weeks post-hatch. Specifically, a previous study comparing

allometric growth of the liver between a modern broiler line (Ross 708) and an unselected line (UIUC), noted that while a similar peak in normalized liver mass occurs in both lines, this peak is shifted one week earlier in the modern broiler line (4). It was then hypothesized that this peak in normalized liver mass may represent some biological event that occurs earlier in the modern broiler, and may be related to its metabolic edge or growth potential. Allometric data collected during the current study provided a higher-definition look at NLM, showing a peak around D6, and then a second one around D14 (Figure 14). The preliminary differential expression of D4 and D20 transcriptome, as described in Chapter 2, revealed an increase in cell proliferation at D4, occurring prior to the first peak, compared with an increase in immune-related genes at D20, occurring after the second peak. A subsequent study on the same line of birds using RNA Fluorescence in situ Hybridization (FISH) to visualize mRNA expression in the liver over the first three weeks post-hatch also confirmed that there is a higher number of cells expressing immune cell marker CD3D at D20 post hatch when compared with D4. This led to a refined hypothesis: the peaks in NLM may represent proliferation of different cell types - hepatocytes and endothelial cells to increase liver function and vascularization at D6, and immune system expansion at D16, concurrent with the depletion of maternal yolk-derived antibodies. Thus, we expected to see strong correlation of markers of cell proliferation associated with NLM.

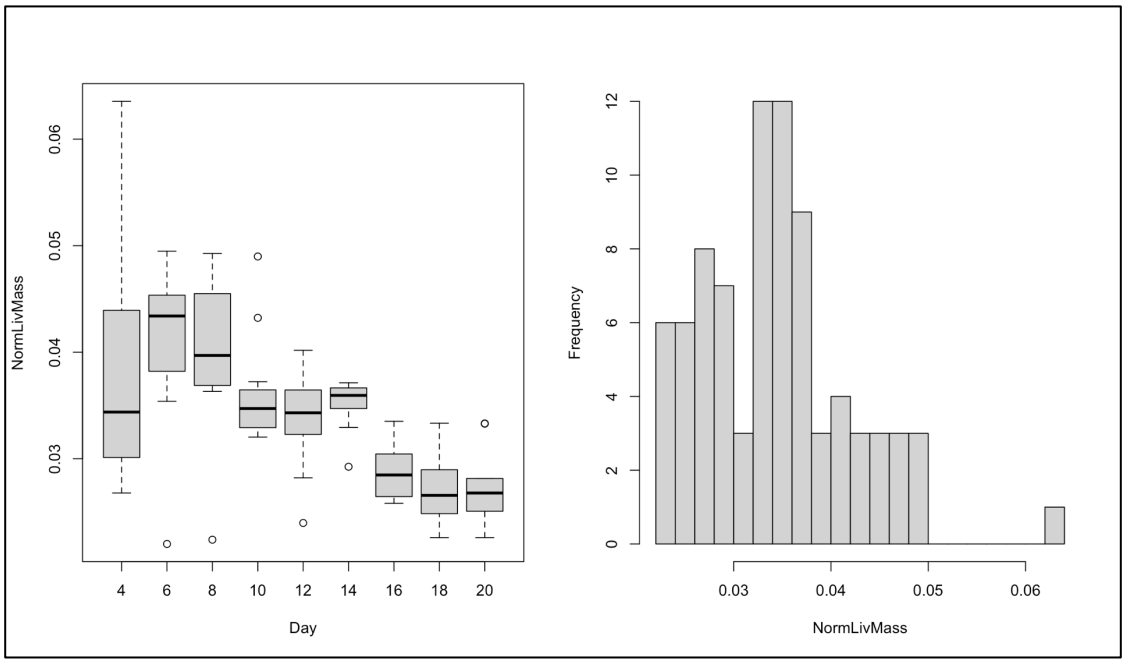

Figure 14: Normalized Liver Mass over time, by age group (left) and frequency of Normalized Liver Mass values (right). Although the total liver mass increases steadily with time, its mass relative to body mass decreases to approach adult values, showing two distinct peaks at Day 6 and Day 14. The histogram shows that Normalized Liver Mass may be comprised of three different distributions, possibly driven by latent sources of variation in the birds.

4.1.1 Workflow overview

The workflow, summarized in Figure 15, can be split into five parts: Exploratory analysis, feature prioritization, model development and refinement, model validation, and generation of hypotheses based on statistically significant associations. Exploratory analysis can begin with investigating unknown heterogeneity in a population. One way to capture this is through Finite Mixture Modeling (FMM). FMM can be used to statistically confirm the presence of different distributions within a variable. In this way, it could serve as a simple clustering method for use with

downstream modeling, or could also be used purely for exploratory analysis. These different distributions may represent "latent classes" or groups of individuals driven by unknown sources of variation. Determining which features contribute to these groups can distill the biological differences between them, prioritizing features or identifying biomarkers. In modeling, the groups can also be used to investigate if the relationships between features and the response variable change between the groups of individuals. For this analysis, we used FMM to confirm heterogeneity present within NLM, further justifying our interest in this measure.

The second step in the workflow is to prioritize features based on their relation to a response variable. This is done using Spearman correlation, which is simple, efficient, and distribution-free, and provides a ranked list of features. Spearman correlation is rank-based and robust, requiring little pre-processing of data. It can also screen tens of thousands of features within minutes, expediting the most computationally-intensive step of the workflow. Features are ranked using rho value and p-value, applying generous cutoffs ($rho>0.2$, $p<0.2$). Features meeting these typically show good association with the response variable, but if the group of potential predictors is still too large the user can select a top percentage of features based on the statistical power required. Here, to validate the workflow and also aid in interpretation of the model, we chose to incorporate broad biological knowledge and focus on the top 20 features with identifying information.

Before building linear models, it is important to shrink the pool of possible features, to help find the simplest group of predictors and also eliminate collinear features that would invalidate the assumption that predictor variables are independent. Least absolute shrinkage and selection operator (Lasso) regression, is one way to do

this that is also easy to incorporate in an automated workflow; while other methods may fail, lasso regression will automatically reduce the coefficients of collinear or insignificant variables to 0 (129). Lasso relies on simulations, yielding coefficients for each variable that give the minimum mean cross-validated error for the model. However, when considering two collinear variables, it is possible to accidentally eliminate the one that actually describes more variation in the response variable. To avoid this problem in our workflow, we run 1000 Lasso simulations and retain the variables that survive 60% of the time.

After eliminating collinear features and reducing the pool of possible predictors using Lasso, our workflow proceeds to linear modeling. Importantly, this step begins to move beyond correlation to quantify the association of predictors with the response variable. The goal of this is not to build a predictive model, but rather to identify features that are strongly associated with the response. The first linear model is built using the surviving features from lasso regression, and model fit evaluated based on adjusted R^2. This also helps to further reduce the pool of features through eliminating variables with insignificant regression coefficients. The second, refined linear model is then tested to see if it is still a good fit for the response variable. In our model, we also incorporated bird age in the model, both as a continuous measure and categorical measure, to see which model performed better.

The final step in the workflow is to validate the model parameters to check for overfitting, ensuring the model is not dependent on the data that generated it. For this, we use bootstrap resampling (130). With small sample sizes, bootstrap is a good method to simulate sampling from a larger population. Resampling is done with replacement, generating a distribution for the parameter of interest. The model statistic

should lie within the 95% confidence interval for the simulated parameter. Although bootstrap can be used with any model statistic, we chose to validate the regression coefficients from the refined linear model, confirming that the predictors are likely to be associated with the response variable, even in other datasets.

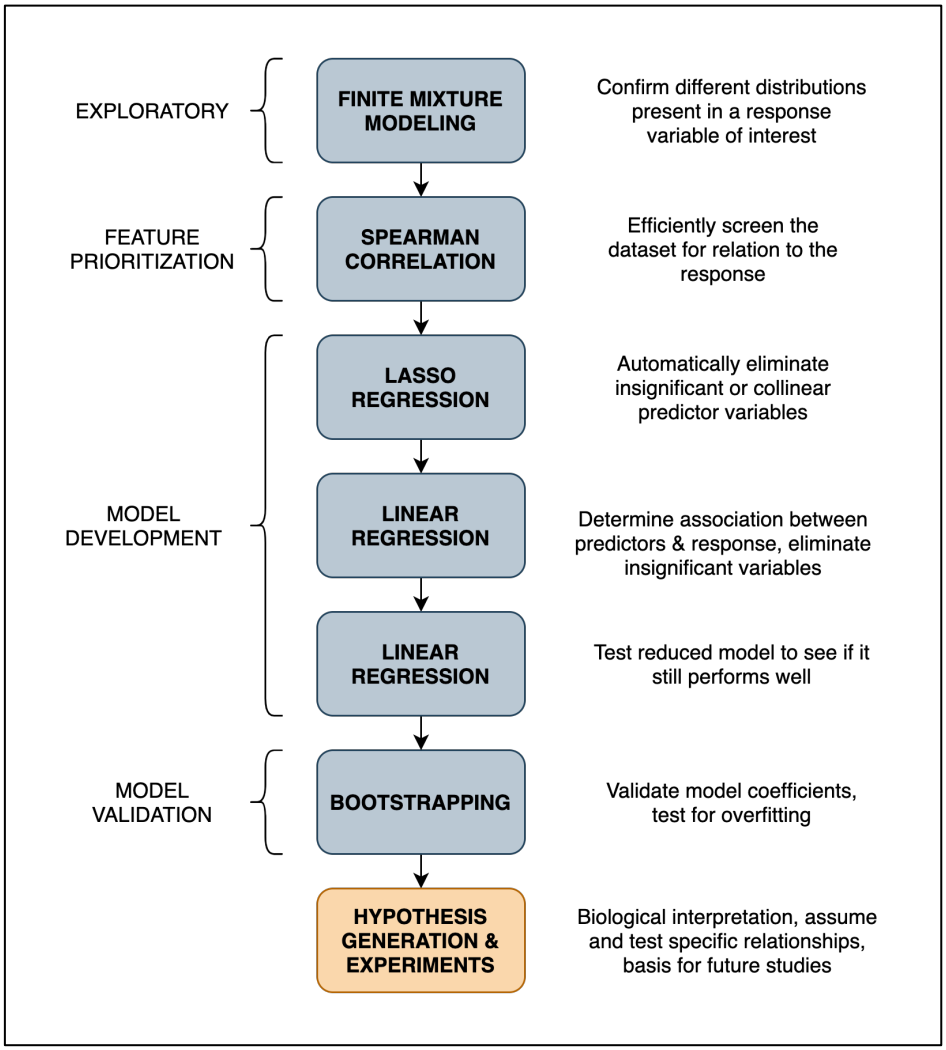

Figure 15: Automated workflow computationally screens datasets and develops models describing associations between predictor variables and a response variable; statistically significant associations provide the basis for hypothesis generation.

4.2 Results

4.2.1 Evaluation of response variable – Normalized Liver Mass

NLM was chosen as the response variable based on biological interest and visual inspection of the frequency histogram (Figure 14) which appears to show three Gaussian distributions. Three distinct clusters were identified by FMM (Figure 16). Cluster 1, made up of the lowest values of NLM (mean = 0.0263), contained mostly older birds, but also a few younger ones. Cluster 2 (mean = 0.0348), with mid-range NLM values, contained birds in every age group, but had a higher proportion of birds D8-D14. Cluster 3, with the highest values of NLM (mean = 0.0414), were represented by the youngest birds. This exploration confirmed our interest in NLM, but NLM clusters were not used in further modeling for this analysis.

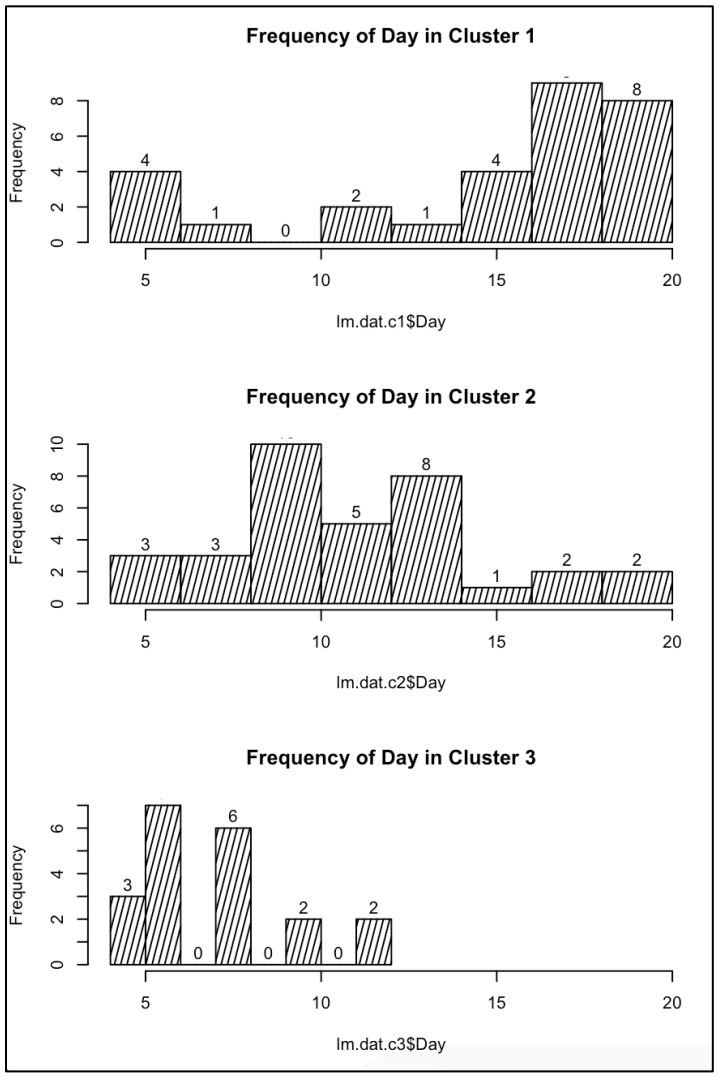

Figure 16: Distribution of bird ages within Normalized Liver Mass clusters identified by FMM.

4.2.2 Top features from Spearman correlation

3668 features met the generous cutoffs of rho>0.2 and p-value<0.2. Many of the top features correlated with NLM (rho >0.6) were unidentified metabolites, mostly complex lipids detected in the liver, but also some from plasma, along with several

biogenic amines. For ease of interpretation, unidentified compounds and plasma metabolites were excluded from the current workflow. Out of the top 100 features with identifying information, 82 were genes, and 67 of these possessed gene symbols associated with human proteins in String. These were enriched for gas transport, but also terms related to cell proliferation. Two genes were related to retinal metabolism (RPE65, ALDH1A3), and two were fatty acid binding proteins (FABP1, FABP2). The remaining 18 were mostly primary metabolites, including several fatty acids (palmitoleic acid, palmitic acid, oleic acid), amino acids (threonine, lysine), amino acid derivatives (N-Methyllysine), antioxidants (vitamin E, retinal), waste products of amino acid metabolism (uric acid, urea), and intermediates of carbohydrate and lipid metabolism (3-phosphoglycerate, 1-monopalmitin).

Table 6: Top 20 identified features associated with NLM by Spearman correlation, showing which survived after each step in the workflow. The final four significant features are shown along with regression coefficients and p-values. Note: all p-values associated with rho correlation were <0.0001.

Feature	rho	Lasso	Linear Model 1	Final Coef, p-value
10.91_920.82_10.92_944.87 (liver)	0.638	X		
Hemoglobin beta, subunit A (HBBA)	0.607			
Galectin 2 (LGALS2)	0.599	X		
Hemoglobin subunit mu (HBM)	0.595			
Erythrocyte Membrane Protein Band 4.2 (EPB42)	0.581			
Hemoglobin subunit alpha 1 (HBA1)	0.574	X	X	0.0037287, 0.0006

Table 6 continued

Translocator Protein 2 (TSPO2)	0.572	X	X	-0.0050558, 0.0001
Vitamin E (liver metabolite)	0.564	X	X	0.0024936, 0.0004
Glial Fibrillary Acidic Protein (GFAP)	0.559	X		
Neurotrophic Receptor Tyrosine Kinase 3 (NTRK3)	0.552	X		
palmitoleic_acid (liver metabolite)	0.550	X		
Nestin (ENSGALG00000013239)	0.548	X	X	
11.32_922.84_11.33_946.88 (liver metabolite)	0.545	X		
Hemoglobin Subunit Epsilon 1 (HBE1)	0.542			
PRC1 - protein regulator of cytokinesis 1 (ENSGALG 00000012836)	0.533	X		
Prolactin-releasing peptide (ENSGALG00000043381)	0.525	X	X	
Solute Carrier Family 47 Member 1 (SLC47A1)	0.523	X	X	0.0020024, 0.004
retinal (liver metabolite)	0.519			
Rh Associated Glycoprotein (RHAG)	0.515	X		
DNA Polymerase Theta (POLQ)	0.512	X		

4.2.3 Features significantly associated with Normalized Liver Mass

Table 6 shows the top 20 genes and liver metabolites with identifying information selected by Spearman correlation with NLM and used in lasso regression. Although selection due to prior knowledge is not necessary in this workflow, we chose to focus on known features to aid in model validation and interpretation. 1000 lasso simulations were repeated, and 15 features survived 60% of the time. Of the first

linear models, that using bird age as a continuous measure performed the best. This model contained 6 features with significant regression coefficients. Three genes (HBA1, TSPO2, SLC47A1) and one liver metabolite (Vitamin E) remained significant in the refined linear model, which predicted NLM with an adjusted R^2 of 0.6232. Figure 17 shows the trends of these features over time, and Figure 18 shows their associations with NLM.

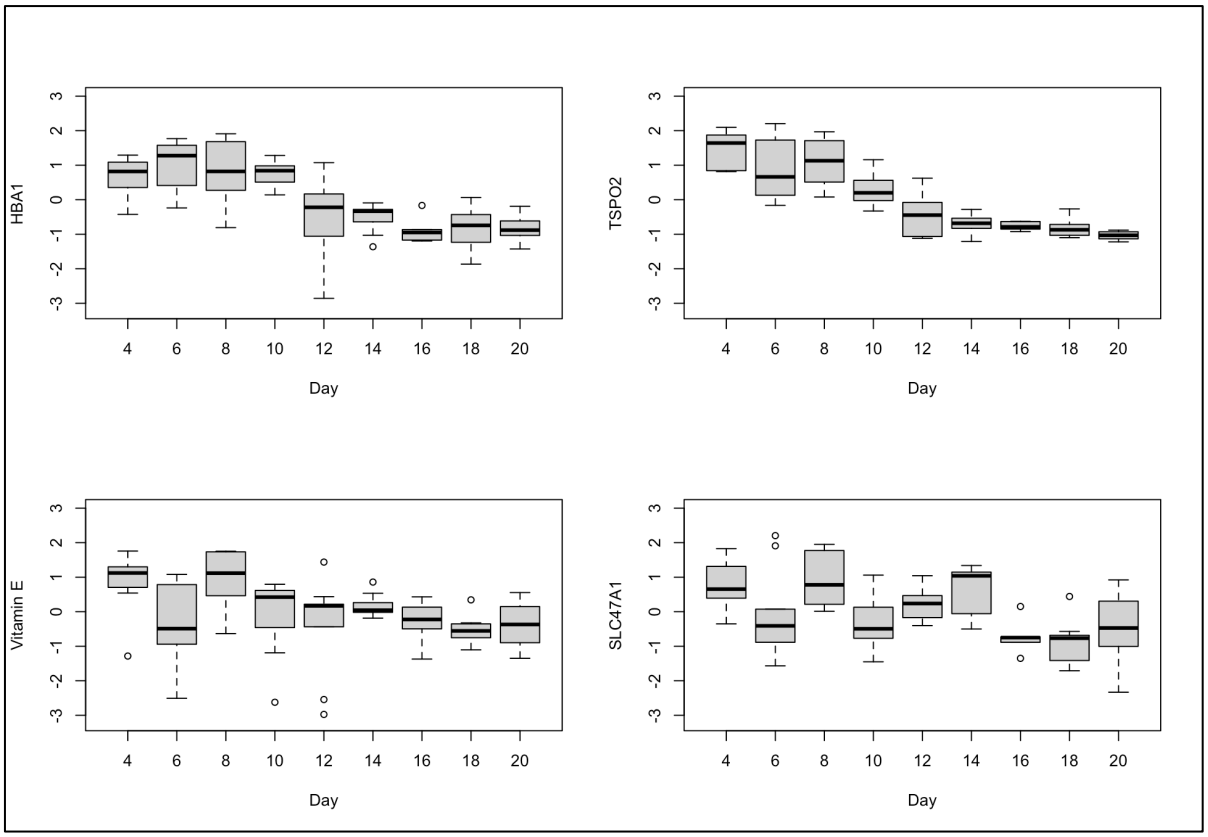

Figure 17: Trends of the four features significantly predicting Normalized Liver Mass.

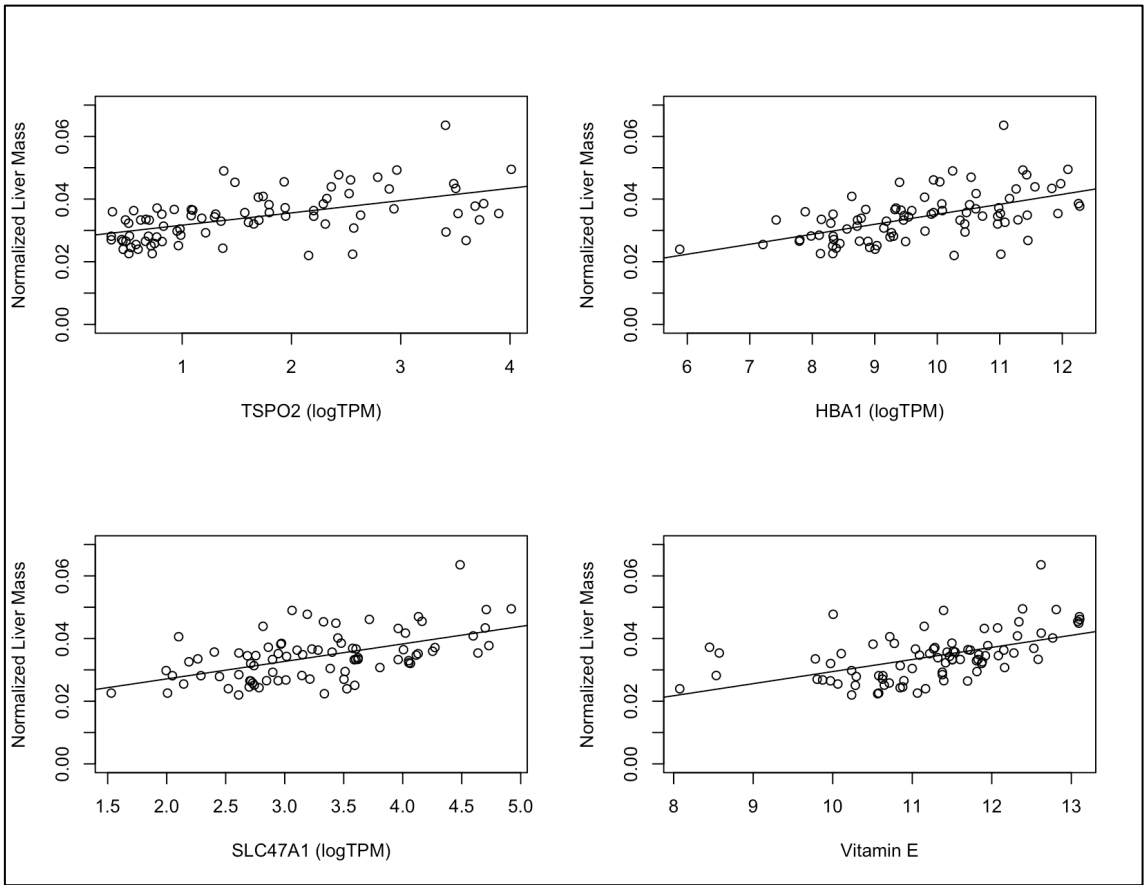

Figure 18: Scatterplots of each variable showing association with NLM.

4.3 Discussion

4.3.1 Normalized Liver Mass as an indicator of biological events

After comparing D4 and D20 transcriptome and metabolome, revealing differences in cell proliferation and immune cell genes, and confirming the increased number of cells expressing immune cell marker CD3D after the second week post-hatch, we hypothesized that the peaks in NLM correspond to different populations of cells proliferating. We expected to see strong correlation of markers of cell proliferation, such as PCNA, with NLM. This relationship was investigated and found

to have no significant association. This indicates that the real model is likely more complex and would benefit from computational prioritization of features.

Liver mass itself, though reliably stabilized in adult birds (7) may be affected by a number of unknown factors throughout development. For example, the storage of yolk lipid evident at D4 may contribute to its mass until its eventual depletion. Normalized liver mass is also convoluted, depending not only on actual liver mass, but on the combined mass of multiple other organ systems with their own allometric trends, and musculoskeletal growth, which selection for production characteristics has both increased and accelerated in the modern broiler. Given this, it is possible the features driving NLM may not be specifically related to the biological processes we hypothesize they represent, namely cell proliferation. First, the influence of other organ masses on this measure may offset any measures representing actual liver growth, enough that these may not show a strong association with NLM. Second, other mechanisms increasing organ mass may be at work, such as cell hypertrophy, autophagy, or migration of different cell types. Association of these proliferative events may also be offset, resulting in weaker correlations. For example, proliferation occurring at D4 could lead to the peak seen at D6. Furthermore, if the peaks do not represent proliferation, or are driven by different biology, it may be harder to find associations that hold true across the across the entire time course. In spite of these caveats, NLM was still considered an interesting variable to investigate. Due to the likely complexity of associations describing NLM, this is an excellent opportunity to apply our workflow, which relies on computational feature selection rather than prior biological knowledge.

4.3.2 Summary of features in final model, overlap with Chapter 3

Many of the top genes correlated with NLM were related to erythrocytes and hemoglobin, and were also present in MCODE Cluster 9 (TSPO2, HBBA, HBA1, HBM, HBE1, RHAG) within the orange WGCNA module, as described in Chapter 3. This module showed the highest positive correlation with NLM (0.6). As both of these approaches begin with correlation-based strategies, the overlap makes sense, and the findings corroborate each other. Other features that showed high Spearman correlation with NLM were present in different clusters in the orange module (LGALS2, palmitoleic acid, retinal, vitamin E, ENSGALG00000013239), or in different modules (SLC47A1 – blue; ENSGALG00000012836, ENSGALG00000043381 - purple).

Of the four remaining features, which together described NLM with an adjusted R^2 of 0.6232, all had a positive association with NLM except for TSPO2. HBA1 is a subunit of hemoglobin; tocopherol gamma (vitamin E) is an antioxidant metabolite; TSPO2 is involved in cholesterol binding in erythrocytes during differentiation; SLC47A1, or MATE1, is a solute transporter most well-studied for its ability to eliminate glucose-lowering Type 2 diabetes drug metformin (131,132). The four features identified by this workflow as being strongly associated with NLM were initially surprising, as we expected to see features strongly related to cell proliferation. However, after further investigation and comparison with results from Chapter 3, these features are likely related to cell proliferation but in a different way than previously thought.

4.3.3 Evidence for increased erythrocyte presence in the liver

Hemoglobin is typically associated with erythrocytes, but can be expressed in other cell types where maintaining oxygen homeostasis is crucial (133–135). There is

also evidence that the non-erythroid expression of hemoglobin can reduce oxidative stress (135,136). One study found that hepatocytes express HBA1 and HBB during non-alcoholic steatohepatitis to combat oxidative stress (136), suggesting a role for these proteins in dealing with oxidative stress brought on by rapid cell proliferation. Although this also indicates that expression of hemoglobin by other non-erythroid cells present in the liver is possible, many other erythroid-specific genes, including those involved in membranes, transport, and metabolism were closely related both here and in Chapter 3. We investigated the possibility that erythrocytes could be proliferating in the liver in the first week post-hatch, and thus represent part of the increase in normalized liver mass. In neonatal mammals, erythropoiesis takes place in erythroid centers in the liver, and the expression of TSPO2 decreases when these centers disappear and bone marrow becomes the main site for erythropoiesis (124). However, this does not appear to be the case in young birds, where the yolk sac is the main site of erythropoiesis through embryonic development and early post-hatch development until it transitions to the bone marrow, and any hematopoiesis in the liver takes place in portal spaces (137,138). We also considered whether the tissue hypoxia hypothesized in the liver at D4 could actually be more widespread, stimulating an overall increase in erythrocyte proliferation. Blood oxygen saturation (sO_2) does decrease from D4-8, then shows a spike at D10 and gradual decrease. Erythropoiesis is increased under systemic hypoxia due to the stimulating effects of HIF1A on erythropoietin (EPO) production in the liver and erythropoietin receptors (EPOR) receptors in bone marrow (111). Although EPO production is likely stimulated by hypoxia observed at D4, and EPO mRNA has been found in the chicken, the gene encoding EPO in chickens has not been annotated so its abundance cannot be

quantified in our data. The strong presence of hemoglobin-related genes, along with many erythroid-specific genes, led us to hypothesize that there is in fact an increase in erythrocyte population in the liver during the first week post-hatch. However, given the unlikelihood that erythropoiesis occurs in avian liver, this is likely due to another reason.

Erythrocytes in birds differ from those in mammals both morphologically and metabolically. In contrast to mammalian erythrocytes, avian erythrocytes are nucleated and contain functional mitochondria (139). Although mammalian erythrocytes depend on glucose for energy, they must produce ATP anaerobically, relying on the PPP and lactate production (140). Mammalian erythrocytes also respond to hypoxia by increasing glucose consumption (141,142). Avian erythrocytes, however, can produce energy aerobically. One study found that, in spite of this, birds had lower levels of oxidative stress in the blood compared to a size-similar mouse model (139). This suggests two important hypotheses about the increased presence of erythrocytes: 1) the metabolic reprogramming observed at D4, specifically the increase in lower glycolysis and Warburg-like effect, is unlikely to be due to an increased presence of erythrocytes at this time point, since avian erythrocytes can produce ATP aerobically; 2) Avian erythrocytes may have increased ability to mediate oxidative stress, suggesting a possibly critical role during increased cell proliferation.

4.3.4 Features describing NLM reflect critical functions during cell proliferation

Regulating oxygen homeostasis in liver tissue is likely especially important in the early stages of post-hatch development, due not only to increased cell proliferation in the liver, but also to maintain the critical hypoxia response and metabolically

regulate the expansion of the immune system. Also paramount is the liver's ability to mediate the harmful effects of oxidative stress. Two of the features describing NLM in our model are directly related to combatting oxidative stress in metabolism-related liver disease (Vitamin E, HBA1). HBA1 was found to be expressed by hepatocytes to help control oxidative stress during non-alcoholic steatohepatitis (136), while vitamin E has been shown to help reduce oxidative stress in non-alcoholic fatty liver disease (NAFLD) (143). This suggests that mechanisms of dealing with oxidative stress are strongly linked to processes resulting in rapid liver growth. Furthermore, the increased erythrocyte presence in liver, as indicated by hemoglobin-related genes, TSPO2, and other erythrocyte-specific genes, when taken with the possibility that avian erythrocytes are especially efficient at binding reactive oxygen species (ROS), suggest a critical role for these cells related to cell proliferation over the first week post-hatch, and thus NLM. Since erythropoiesis does not occur in avian liver, we hypothesize that the liver instead recruits and maintains an increased population of erythrocytes to help regulate oxygen homeostasis, stabilizing the hypoxia response and promoting a hypoxic environment while the liver increases its mass, and concurrently protecting against oxidative stress brought on by rapid cell proliferation.

The role of SLC47A1 in driving NLM is somewhat perplexing, given most of our understanding concerns its ability to transport and eliminate diabetes drug metformin. However, one study investigating its ability to transport the hypolipidemic flavonoid quercetin also noted that its overexpression enhanced cellular glucose uptake (144). Although this study did not investigate mechanisms, it hints at a more widespread metabolic role for SLC47A1 beyond toxin extrusion, hinging on its abilities to increase intracellular concentrations of a lipid-lowering compound, and

increase the availability of intracellular glucose. When this is evaluated along with its role in decreasing levels of a glucose-lowering drug, it appears this transporter may play a part in ensuring glucose supply, and thus could be critical during times of increased glucose demand as observed in early post-hatch liver. The relationship between SLC47A1 and NLM may thus be under control of a higher regulatory mechanism responsible for metabolic reprogramming associated with post-hatch hepatic development.

4.4 Conclusions:

In this analysis, four features (HBA1, Vitamin E, TSPO2, SLC47A1) showed strong ability to predict NLM. While none of these were directly linked to cell proliferation, they indicate an interesting finding: the predictors describing NLM are more related to biological processes that facilitate and mediate the effects of proliferation, rather than to proliferation itself. These features provide further support building on the hypotheses from Chapters 2 and 3 and linking the importance of oxygen homeostasis, controlling oxidative stress, and regulating glucose availability to cell proliferation.

In addition to the statistically significant, validated associations discussed in this analysis, there are multiple other ways to examine the results and extend the model to more complex situations. For example, NLM was evaluated with FMM to confirm the presence of different distributions representing latent classes of variation in our population, but these clusters were not utilized in the downstream linear models. Given that NLM may not be as closely related to cell proliferation as previously thought, or the possibility that different mechanisms affecting organ mass may be

occurring, it may be advantageous to explore the association between our predictors within each NLM cluster separately.

This approach has several advantages, making it broadly applicable and adaptable to many situations beyond the context of this work. First, it is highly scalable since the first step – prioritization of features – relies on a very simple and efficient correlation measure. High-throughput datasets even larger than the one evaluated here can be screened within minutes for relation to any response variable. Secondly, the workflow is customizable, both in terms of biological question and modeling approach. Features can be prioritized based on any physiological measure, such as NLM or blood glucose, or even based on association with another variable, such as levels of the metabolite citrate. By the time linear models are developed, the pool of predictors is small enough to allow flexible manipulation of the models themselves, to explore the effects of different interactions and covariates. Third, this workflow can rely almost entirely on computation, eliminating bias. Aside from choosing the measure by which to screen the features, driven by biological question of interest and experimental design, it is not necessary to bring in biological knowledge until the very last step, when developing the assumptions to test. Statistically significant assumptions based on these models leave us with solid, computationally validated associations, which can be interpreted and used to generate hypotheses, or tested experimentally.

4.5 Methods

4.5.1 Data preprocessing

The dataset screened in this analysis consisted of all transcriptome and metabolome data. Rigorous sample-wise quality control was conducted as described in Chapter 3, then matching of samples to include those containing both transcriptome and metabolome data. Gentle filtering was done on the larger datasets to speed up the screening. The transcriptome, quantified by relative expression (TPM) values was filtered to remove any rows consisting entirely of 0s. The primary plasma and liver metabolome was the smallest dataset and included without filtering. The liver and plasma biogenic amines and complex lipid metabolome datasets were filtered using Metaboanalyst: 1) for low repeatability using relative standard deviation (RSD = standard deviation/mean) excluding those that deviated from QC samples >25%, 2) using IQR to robustly exclude those with near-constant values, 3) using median to detect compounds with very small values close to the detection limit. The final filtered dataset consisted of 83 birds and 31475 features.

4.5.2 Workflow

All analysis was done in statistical software R unless otherwise noted. Finite mixture modeling (mixtools package) (142) was used to confirm the presence of three Gaussian distributions within the measure of Normalized Liver Mass (NLM). Spearman correlation was used to screen this dataset for correlation with NLM. 3688 features met the generous cutoffs (rho >=0.2, p <=0.2), so the top few candidates were chosen based on statistical power for the number of observations (83 birds total). For ease of downstream biological interpretation and development of models, the top 20 liver genes and live metabolites that possessed meaningful annotation or identification

were selected. Lasso regression (caret and glmnet packages) (143,144) was performed 1000 times on this group, allowing simulation results to vary, and first cross-validating and selecting the value of lambda that gave the minimum mean cross-validated error. Features that survived in 60% or more of the simulations were retained for linear modeling (15 total features). Two linear models were tested with these 15 features, one using bird age (Day) as a continuous measure, and one using it as a categorical measure or factor. The model using bird age as continuous performed the best (adjusted R^2 0.6882), with six features significant at $p<0.05$. The linear model was rerun using these six features, and still performing well (adjusted R^2 0.6232). Bootstrap resampling was performed 1000 times (boot and modelr packages) to validate the regression coefficients, which were all within the 95% confidence intervals (145-147).

Chapter 5

CONCLUSIONS

5.1 Hypotheses generated

This work generated multiple testable hypotheses highlighted below:

- A hypoxic environment exists in the liver at D4, regulated by HIF1A, and evident in metabolic reprogramming. This metabolic flexibility allows for ATP production during a period of rapid cell proliferation.
- At D4, the liver is sequestering glucose and glycolytic intermediates for its own use, to support cell proliferation, the PPP, and to feed lower glycolysis. It releases alternative energy sources, including lactate and acetate, to support the rest of the body.
- The reservoir of yolk-derived fatty acids abundant in liver at D4 are modified and elongated for utilization in cell membrane components during proliferation rather than energy through beta-oxidation. This also stabilizes the hypoxia response and limits oxidative stress that would come from oxidizing lipids for energy.
- Hypertrophy through polyploidization, controlled by E2F8, contributes to organ growth during the initial burst of cell proliferation, offering a metabolic advantage by leading to larger cells that may also contain increased numbers of mitochondria.
- Metabolic flexibility is exhibited in amino acid catabolism at D4, with glucogenic products supporting the liver's requirements for proliferation and energy production, and ketogenic products exported for use by other tissues.
- At D4, the liver is poised to respond to oxygen availability, while at D20, maintaining energy balance between storage and release of nutrients is prioritized.
- Hypoxia response controls metabolism and glucose availability, suppressing the proliferation of other cell types over the first week post-hatch, and allowing a burst of hepatocyte proliferation and vascularization to occur as the liver establishes its identity. Changes in oxygen and glucose availability then trigger expansion of the immune system in the liver.

- There is an increased erythrocyte presence in the liver in the first week post-hatch, that helps to regulate oxygen homeostasis and combat oxidative stress during rapid cell proliferation.

In addition to the hypotheses generated, this work also implemented several methods with demonstrated ability to computationally capture biological relationships, even when integrating different types of high-throughput data. WGCNA and MCODE grouped genes and metabolites into biologically related groups and correlated these with physiological traits. These clusters also linked metabolic reprogramming to other biological events such as regulation of cell proliferation and immune system expansion. The modeling workflow described in Chapter 4 identified predictors strongly associated with NLM, providing a statistical foundation for hypothesis generation. This workflow is also broadly applicable and flexible outside of the context of this work.

5.2 Future work

There are multiple opportunities to extend and enhance the analyses in each section of this work. The comparison of D4 versus D20 in Chapter 2 exposed many differences in signaling, for example, that were not evaluated due to the focus on core metabolism. There were also several clusters identified by WGCNA in Chapter 3 that contained interesting features, such as those related to hypoxia response. Time-dependent correlations, and crosstalk between modules and clusters, can also be explored computationally to clarify the relationships present and lead to observations about how regulation changes over the time course. For example, two genes involved in cell cycle, may have grouped in separate modules: gene A in purple and gene B in saddlebrown. These genes, exhibiting large enough differences in abundance over the time course to be classified in separate modules, may actually be very closely linked

during D4-D8, but show a decoupling in trends after D10. Targeting these pre-identified relationships for a closer look at how correlations change from day to day could expose regulatory changes associated with different types of cells proliferating. Strong relationships can even be investigated using ordinary differential equations to model the interactions of multiple features over time. There is also the potential to construct regulatory networks based on this dataset, identifying hierarchies of features, both within and across modules.

The workflow described and implemented in Chapter 4 can be applied as-is to this dataset to examine any number of relationships, such as plasma metabolites strongly associated with blood glucose levels. The models developed can also be extended to investigate associations in more complex, biologically accurate contexts. The linear models developed were fairly simple, but exploring interactions between predictor variables, or employing other types of models such as structural equation modeling could capture these relationships with more precision. This knowledge can then be taken further and tested experimentally, for example by knocking out genes in the chicken.

This work also included many unknown genes and metabolites that were identified computationally as important contributors to differences during liver development. The prioritization of these features, and ability of these methods to locate them within the context of well-known features and pathways, identifies them as possible targets for further studies. Although investigation of these unknowns was beyond the scope of this work, they clearly warrant further exploration to determine their identify and specific function, as they may represent important novel biomarkers or regulatory targets.

In addition to experimental testing of the hypotheses generated in this work, future work should also address the shortcomings of this study. For example, a major limitation of both RNA-seq and tissue metabolomics is that they each provide an averaged snapshot of biological processes occurring. While this is beneficial for comparing tissues that are fairly similar in terms of cell type and composition, it is something that must be taken into consideration when analyzing tissues where these factors may differ, as it will have broad impact on the overall picture. A growing liver almost certainly exhibits large shifts in cell populations as a function of development, and it is nearly impossible to quantify how much variation is due to differences in cell composition or changes in the metabolic reprogramming of actual hepatocytes, for example. Future analyses could explore the relationships noted in this analysis through techniques that allow the quantification of transcription in specific cell types, such as single-cell RNA-seq, or visualization of RNA expression in the context of the tissue, such as RNA FISH.

5.3 Broad-reaching utility

Characterizing the stages of liver development in the modern broiler in terms of metabolic regulation, oxidative stress, timing of biological processes, and awareness of different cell populations can lead to improvements in the meat bird industry – specifically management techniques, feed supplement studies, or breed enhancement strategies. However, the value of this work extends beyond the chicken, to knowledge of general gene-metabolite relationships, but also to implications in human medicine. For example, although rapid cell proliferation and metabolism under hypoxia is well-characterized in cancer cells, less is known about these processes under normal growth conditions. Understanding metabolism during normal,

development-programmed rapid cell proliferation can provide a critical contrast to metabolism during abnormal, malignant cell proliferation. This could help to differentiate between gene or metabolite targets for therapies, identifying those more likely to specifically inhibit metabolism in cancer cells, without crippling metabolism or proliferation in healthy cells. Understanding hypoxia's role as a protective mechanism during rapid organ growth and clarifying its regulatory involvement during the transition to normal metabolism could also contribute to general scientific knowledge about metabolism during organ development.

The broiler is also an interesting model for investigating processes contributing to metabolic disease in human medicine. Broilers are especially well-equipped to mobilize excess nutrients, converting this to muscle but also preventing the buildup of lipids in the liver. Due to decreased appetite regulation, the liver meets a near-constant carbohydrate burden that would cause overnutrition-related disease in many other species. Broilers display improved lipogenesis and lipid transport, even when compared with layer chickens or migratory avian breeds, and are far less prone to issues like hepatic steatosis. Understanding how broiler liver efficiently manages an abundance of nutrients could shed light on mechanisms contributing to a number of human diseases related to obesity and accumulation of fat in the liver, including Non-alcoholic Fatty Liver Disease (NAFLD). Also, although birds show major differences in insulin resistance and glucose tolerance when compared with mammals, exploring the molecular basis for this could provide a valuable comparison with diseases relating to human metabolic dysregulation, such as diabetes. Computationally, this work has demonstrated and established the value of several methods for hypotheses generation from high-throughput data, while biologically, it has led to observations warranting

further exploration and consideration of the chicken as a valuable model for organ development and metabolism.

REFERENCES

1. U.S. Broiler Performance [Internet]. The National Chicken Council. [cited 2020 Jul 16]. Available from: https://www.nationalchickencouncil.org/about-the-industry/statistics/u-s-broiler-performance/

2. Zuidhof MJ, Schneider BL, Carney VL, Korver DR, Robinson FE. Growth, efficiency, and yield of commercial broilers from 1957, 1978, and 20051 1This is an Open Access article distributed under the terms of the Creative Commons Attribution-Noncommercial License (http://creativecommons.org/licenses/by-nc/3.0/), which permits noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited. Poult Sci. 2014 Dec 1;93(12):2970–82.

3. Havenstein GB, Ferket PR, Qureshi MA. Growth, livability, and feed conversion of 1957 versus 2001 broilers when fed representative 1957 and 2001 broiler diets. Poult Sci. 2003 Oct;82(10):1500–8.

4. Schmidt CJ, Persia ME, Feierstein E, Kingham B, Saylor WW. Comparison of a modern broiler line and a heritage line unselected since the 1950s. Poult Sci. 2009 Dec;88(12):2610–9.

5. Hermier D. Lipoprotein metabolism and fattening in poultry. J Nutr. 1997;127(5 Suppl):805S-808S.

6. Krogdahl A. Digestion and absorption of lipids in poultry. J Nutr. 1985 May;115(5):675–85.

7. Denbow DM, Cline MA. Chapter 21 - Food Intake Regulation. In: Scanes CG, editor. Sturkie's Avian Physiology (Sixth Edition) [Internet]. San Diego: Academic Press; 2015 [cited 2020 Jul 16]. p. 469–85. Available from: http://www.sciencedirect.com/science/article/pii/B978012407160500021X

8. Trefts E, Gannon M, Wasserman DH. The liver. Curr Biol CB. 2017 Nov 6;27(21):R1147–51.

9. Zaefarian F, Abdollahi MR, Cowieson A, Ravindran V. Avian Liver: The Forgotten Organ. Anim Open Access J MDPI. 2019 Feb 15;9(2).

10. Gordillo M, Evans T, Gouon-Evans V. Orchestrating liver development. Dev Camb Engl. 2015 Jun 15;142(12):2094–108.

11. Jastrebski SF, Lamont SJ, Schmidt CJ. Chicken hepatic response to chronic heat stress using integrated transcriptome and metabolome analysis. PLoS ONE [Internet]. 2017 Jul 31 [cited 2020 Jul 16];12(7). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5536301/

12. Sabino M, Capomaccio S, Cappelli K, Verini-Supplizi A, Bomba L, Ajmone-Marsan P, et al. Oregano dietary supplementation modifies the liver transcriptome profile in broilers: RNASeq analysis. Res Vet Sci. 2018 Apr 1;117:85–91.

13. Korošec T, Tomažin U, Horvat S, Keber R, Salobir J. The diverse effects of α- and γ-tocopherol on chicken liver transcriptome. Poult Sci. 2017 Mar 1;96(3):667–80.

14. Lan X, Hsieh JCF, Schmidt CJ, Zhu Q, Lamont SJ. Liver transcriptome response to hyperthermic stress in three distinct chicken lines. BMC Genomics. 2016 22;17(1):955.

15. Coble DJ, Fleming D, Persia ME, Ashwell CM, Rothschild MF, Schmidt CJ, et al. RNA-seq analysis of broiler liver transcriptome reveals novel responses to high ambient temperature. BMC Genomics. 2014 Dec 10;15:1084.

16. Cogburn LA, Trakooljul N, Chen C, Huang H, Wu CH, Carré W, et al. Transcriptional profiling of liver during the critical embryo-to-hatchling transition period in the chicken (Gallus gallus). BMC Genomics. 2018 Sep 21;19(1):695.

17. Cogburn LA, Trakooljul N, Wang X, Ellestad LE, Porter TE. Transcriptome analyses of liver in newly-hatched chicks during the metabolic perturbation of fasting and re-feeding reveals THRSPA as the key lipogenic transcription factor. BMC Genomics. 2020 Jan 31;21(1):109.

18. Willson N-L, Forder REA, Tearle R, Williams JL, Hughes RJ, Nattrass GS, et al. Transcriptional analysis of liver from chickens with fast (meat bird), moderate (F1 layer x meat bird cross) and low (layer bird) growth potential. BMC Genomics [Internet]. 2018 May 2 [cited 2020 Jul 16];19. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5930858/

19. Friedmann H. The Avian Embryo. Structural and functional development. Alexis L. Romanoff. Macmillan, New York, 1960. xvi + 1305 pp. Illus. $35. Science. 1960 Apr 22;131(3408):1219–1219.

20. Speake BK, Murray AM, Noble RC. Transport and transformations of yolk lipids during development of the avian embryo. Prog Lipid Res. 1998 May;37(1):1–32.

21. Noy Y, Sklan D. Energy utilization in newly hatched chicks. Poult Sci. 1999 Dec 1;78(12):1750–6.

22. Sahan U, Ipek A, Sozcu A. Yolk sac fatty acid composition, yolk absorption, embryo development, and chick quality during incubation in eggs from young and old broiler breeders. Poult Sci. 2014 Aug;93(8):2069–77.

23. FEAST M, NOBLE RC, SPEAKE BK, FERGUSON MWJ. The effect of temporary reductions in incubation temperature on growth characteristics and lipid utilisation in the chick embryo. J Anat. 1998 Oct;193(Pt 3):383–90.

24. Noble RC, Ogunyemi D. Lipid changes in the residual yolk and liver of the chick immediately after hatching. Biol Neonate. 1989;56(4):228–36.

25. Bigot K, Mignon-Grasteau S, Picard M, Tesseraud S. Effects of delayed feed intake on body, intestine, and muscle development in neonate broilers. Poult Sci. 2003 May;82(5):781–8.

26. Martin MP, Wineland M, Barnes HJ. Selected blood chemistry and gas reference ranges for broiler breeders using the i-STAT handheld clinical analyzer. Avian Dis. 2010 Sep;54(3):1016–20.

27. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009 Jan;4(1):44–57.

28. Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 2009 Jan;37(1):1–13.

29. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci. 2005 Oct 25;102(43):15545–50.

30. Mootha VK, Lindgren CM, Eriksson K-F, Subramanian A, Sihag S, Lehar J, et al. PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet. 2003 Jul;34(3):267–73.

31. Grenier E, Maupas FS, Beaulieu J-F, Seidman E, Delvin E, Sane A, et al. Effect of retinoic acid on cell proliferation and differentiation as well as on lipid synthesis, lipoprotein secretion, and apolipoprotein biogenesis. Am J Physiol-Gastrointest Liver Physiol. 2007 Dec 1;293(6):G1178–89.

32. PubChem. 2-Hydroxybutyric acid [Internet]. [cited 2020 Jul 16]. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/11266

33. PubChem. Uracil [Internet]. [cited 2020 Jul 16]. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/1174

34. Scanes CG. Chapter 20 - Protein Metabolism. In: Scanes CG, editor. Sturkie's Avian Physiology (Sixth Edition) [Internet]. San Diego: Academic Press; 2015 [cited 2020 Jul 16]. p. 455–67. Available from: http://www.sciencedirect.com/science/article/pii/B9780124071605000208

35. Chen YP, Chen X, Zhang H, Zhou YM. Effects of dietary concentrations of methionine on growth performance and oxidative status of broiler chickens with different hatching weight. Br Poult Sci. 2013;54(4):531–7.

36. Dukes' Physiology of Domestic Animals, 13th Edition | Wiley [Internet]. Wiley.com. [cited 2020 Jul 16]. Available from: https://www.wiley.com/en-us/Dukes%27+Physiology+of+Domestic+Animals%2C+13th+Edition-p-9781118501399

37. Warburg O. The Metabolism of Carcinoma Cells. J Cancer Res. 1925 Mar 1;9(1):148–63.

38. Liberti MV, Locasale JW. The Warburg Effect: How Does it Benefit Cancer Cells? Trends Biochem Sci. 2016 Mar;41(3):211–8.

39. Vander Heiden MG, Cantley LC, Thompson CB. Understanding the Warburg Effect: The Metabolic Requirements of Cell Proliferation. Science. 2009 May 22;324(5930):1029–33.

40. Kim HS, Lee HE, Yang H-K, Kim WH. High Lactate Dehydrogenase 5 Expression Correlates with High Tumoral and Stromal Vascular Endothelial Growth Factor Expression in Gastric Cancer. Pathobiology. 2014;81(2):78–85.

41. Valvona CJ, Fillmore HL, Nunn PB, Pilkington GJ. The Regulation and Function of Lactate Dehydrogenase A: Therapeutic Potential in Brain Tumor. Brain Pathol Zurich Switz. 2016 Jan;26(1):3–17.

42. HUNT TK, ASLAM RS, BECKERT S, WAGNER S, GHANI QP, HUSSAIN MZ, et al. Aerobically-Derived Lactate Stimulates Revascularization and Tissue Repair via Redox Mechanisms. Antioxid Redox Signal. 2007 Aug;9(8):1115–24.

43. Yamashita H, Kaneyuki T, Tagawa K. Production of acetate in the liver and its utilization in peripheral tissues. Biochim Biophys Acta. 2001 May 31;1532(1–2):79–87.

44. Lund J, Aas V, Tingstad R, Hees A, Nikolić N. Utilization of lactic acid in human myotubes and interplay with glucose and fatty acid metabolism. Sci Rep. 2018 Dec 1;8.

45. Oslund RC, Su X, Haugbro M, Kee J-M, Esposito M, David Y, et al. Bisphosphoglycerate mutase controls serine pathway flux via 3-phosphoglycerate. Nat Chem Biol. 2017 Oct;13(10):1081–7.

46. Isaacks RE, Lai LL, Goldman PH, Kim CY. Studies on avian erythrocyte metabolism. XVI. Accumulation of 2,3-bisphosphoglycerate with shifts in oxygen affinity of chicken erythrocytes. Arch Biochem Biophys. 1987 Aug 15;257(1):177–85.

47. Mattaini KR, Sullivan MR, Vander Heiden MG. The importance of serine metabolism in cancer. J Cell Biol. 2016 Aug 1;214(3):249–57.

48. Israelsen WJ, Vander Heiden MG. Pyruvate kinase: function, regulation and role in cancer. Semin Cell Dev Biol. 2015 Jul;43:43–51.

49. Iqbal MA, Gupta V, Gopinath P, Mazurek S, Bamezai RNK. Pyruvate kinase M2 and cancer: an updated assessment. FEBS Lett. 2014 Aug 19;588(16):2685–92.

50. Zeng L, Morinibu A, Kobayashi M, Zhu Y, Wang X, Goto Y, et al. Aberrant IDH3α expression promotes malignant tumor growth by inducing HIF-1-mediated metabolic reprogramming and angiogenesis. Oncogene. 2015 Sep;34(36):4758–66.

51. Martínez-Reyes I, Chandel NS. Mitochondrial TCA cycle metabolites control physiology and disease. Nat Commun. 2020 Jan 3;11(1):102.

52. Hunkeler M, Hagmann A, Stuttfeld E, Chami M, Guri Y, Stahlberg H, et al. Structural basis for regulation of human acetyl-CoA carboxylase. Nature. 2018 Jun;558(7710):470–4.

53. Leighton F, Bergseth S, Rørtveit T, Christiansen EN, Bremer J. Free acetate production by rat hepatocytes during peroxisomal fatty acid and dicarboxylic acid oxidation. J Biol Chem. 1989 Jun 25;264(18):10347–50.

54. Bose S, Ramesh V, Locasale JW. Acetate Metabolism in Physiology, Cancer, and Beyond. Trends Cell Biol. 2019 Sep 1;29(9):695–703.

55. Dunwoodie SL. The role of hypoxia in development of the Mammalian embryo. Dev Cell. 2009 Dec;17(6):755–73.

56. Chong J, Yamamoto M, Xia J. MetaboAnalystR 2.0: From Raw Spectra to Biological Insights. Metabolites. 2019 Mar;9(3):57.

57. Xia J, Wishart DS. Using MetaboAnalyst 3.0 for Comprehensive Metabolomics Data Analysis. Curr Protoc Bioinforma. 2016;55(1):14.10.1-14.10.91.

58. MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis | Nucleic Acids Research | Oxford Academic [Internet]. [cited 2020 Jul 16]. Available from: https://academic.oup.com/nar/article/46/W1/W486/4995686

59. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. 2000 Jan 1;28(1):27–30.

60. PubChem Identifier Exchange Service [Internet]. [cited 2020 Jul 16]. Available from: https://pubchem.ncbi.nlm.nih.gov/idexchange/idexchange.cgi

61. Li H. lh3/seqtk [Internet]. 2020 [cited 2020 Jul 16]. Available from: https://github.com/lh3/seqtk

62. Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data [Internet]. [cited 2020 Jul 16]. Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

63. Babraham Bioinformatics - Trim Galore! [Internet]. [cited 2020 Jul 16]. Available from: http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/

64. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal. 2011 May 2;17(1):10–2.

65. Zerbino DR, Achuthan P, Akanni W, Amode MR, Barrell D, Bhai J, et al. Ensembl 2018. Nucleic Acids Res. 2018 Jan 4;46(D1):D754–61.

66. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol. 2019 Aug;37(8):907–15.

67. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009 Aug 15;25(16):2078–9.

68. Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. Nat Protoc. 2016 Sep;11(9):1650–67.

69. Pertea M, Pertea GM, Antonescu CM, Chang T-C, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol. 2015 Mar;33(3):290–5.

70. Kovaka S, Zimin AV, Pertea GM, Razaghi R, Salzberg SL, Pertea M. Transcriptome assembly from long-read RNA-seq alignments with StringTie2. Genome Biol. 2019 Dec 16;20(1):278.

71. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinforma Oxf Engl. 2014 Apr 1;30(7):923–30.

72. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014 Dec 5;15(12):550.

73. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B Methodol. 1995;57(1):289–300.

74. Thomas PD, Campbell MJ, Kejariwal A, Mi H, Karlak B, Daverman R, et al. PANTHER: A Library of Protein Families and Subfamilies Indexed by Function. Genome Res. 2003 Sep 1;13(9):2129–41.

75. PANTHER version 7: improved phylogenetic trees, orthologs and collaboration with the Gene Ontology Consortium | Nucleic Acids Research | Oxford Academic [Internet]. [cited 2020 Jul 16]. Available from: https://academic.oup.com/nar/article/38/suppl_1/D204/3112164

76. bioDBnet - Biological Database Network [Internet]. [cited 2020 Jul 16]. Available from: https://biodbnet-abcc.ncifcrf.gov/

77. Jassal B, Matthews L, Viteri G, Gong C, Lorente P, Fabregat A, et al. The reactome pathway knowledgebase. Nucleic Acids Res. 2020 Jan 8;48(D1):D498–503.

78. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 2019 Jan 8;47(D1):D607–13.

79. Lê S, Josse J, Husson F. FactoMineR: An R Package for Multivariate Analysis. J Stat Softw. 2008 Mar 18;25(1):1–18.

80. Releases · kassambara/factoextra [Internet]. GitHub. [cited 2021 Jan 8]. Available from: /kassambara/factoextra/releases

81. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008 Dec 29;9(1):559.

82. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. Genome Res. 2003 Nov;13(11):2498–504.

83. Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics. 2003 Jan 13;4(1):2.

84. Xu B-F, Liu R, Huang C-X, He B-S, Li G-Y, Sun H-S, et al. Identification of key genes in ruptured atherosclerotic plaques by weighted gene correlation network analysis. Sci Rep. 2020 Jul 2;10(1):10847.

85. Frontiers | Weighted Gene Co-Expression Network Analysis Identifies Critical Genes in the Development of Heart Failure After Acute Myocardial Infarction | Genetics [Internet]. [cited 2021 Apr 6]. Available from: https://www.frontiersin.org/articles/10.3389/fgene.2019.01214/full

86. PubChem. Pantothenic acid [Internet]. [cited 2021 Apr 6]. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/6613

87. Lee HO, Davidson JM, Duronio RJ. Endoreplication: polyploidy with purpose. Genes Dev. 2009 Nov 1;23(21):2461–77.

88. Orr-Weaver TL. When bigger is better: the role of polyploidy in organogenesis. Trends Genet TIG. 2015 Jun;31(6):307–15.

89. Zheng K, Cubero FJ, Nevzorova YA. c-MYC—Making Liver Sick: Role of c-MYC in Hepatic Cell Function, Homeostasis and Disease. Genes [Internet]. 2017 Apr 19 [cited 2021 Mar 29];8(4). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5406870/

90. Conner EA, Lemmer ER, Sánchez A, Factor VM, Thorgeirsson SS. E2F1 blocks and c-Myc accelerates hepatic ploidy in transgenic mouse models. Biochem Biophys Res Commun. 2003 Feb 28;302(1):114–20.

91. Han J, Zhang L, Guo H, Wysham WZ, Roque DR, Willson AK, et al. Glucose promotes cell proliferation, glucose uptake and invasion in endometrial cancer cells via AMPK/mTOR/S6 and MAPK signaling. Gynecol Oncol. 2015 Sep;138(3):668–75.

92. Lytovchenko O, Kunji ERS. Expression and putative role of mitochondrial transport proteins in cancer. Biochim Biophys Acta BBA - Bioenerg. 2017 Aug 1;1858(8):641–54.

93. Gutiérrez-Aguilar M, Baines CP. Physiological and pathological roles of mitochondrial SLC25 carriers. Biochem J. 2013 Sep 15;454(3):371–86.

94. Mandard S, Stienstra R, Escher P, Tan NS, Kim I, Gonzalez FJ, et al. Glycogen synthase 2 is a novel target gene of peroxisome proliferator-activated receptors. Cell Mol Life Sci. 2007 May;64(9):1145–57.

95. Momcilovic M, Bailey ST, Lee JT, Fishbein MC, Magyar C, Braas D, et al. Targeted Inhibition of EGFR and Glutaminase Induces Metabolic Crisis in EGFR Mutant Lung Cancer. Cell Rep. 2017 Jan 17;18(3):601–10.

96. Makinoshima H, Takita M, Matsumoto S, Yagishita A, Owada S, Esumi H, et al. Epidermal Growth Factor Receptor (EGFR) Signaling Regulates Global Metabolic Pathways in EGFR-mutated Lung Adenocarcinoma. J Biol Chem. 2014 Jul 25;289(30):20813–23.

97. Jin N, Bi A, Lan X, Xu J, Wang X, Liu Y, et al. Identification of metabolic vulnerabilities of receptor tyrosine kinases-driven cancer. Nat Commun. 2019 Jun 20;10(1):2701.

98. Yoon Y-S, Seo W-Y, Lee M-W, Kim S-T, Koo S-H. Salt-inducible Kinase Regulates Hepatic Lipogenesis by Controlling SREBP-1c Phosphorylation. J Biol Chem. 2009 Apr 17;284(16):10446–52.

99. Gao W-W, Tang H-MV, Cheng Y, Chan C-P, Chan C-P, Jin D-Y. Suppression of gluconeogenic gene transcription by SIK1-induced ubiquitination and

degradation of CRTC1. Biochim Biophys Acta Gene Regul Mech. 2018 Mar;1861(3):211–23.

100. Tao R, Wang C, Stöhr O, Qiu W, Hu Y, Miao J, et al. Inactivating hepatic follistatin alleviates hyperglycemia. Nat Med. 2018 Jul;24(7):1058–69.

101. Patella S, Phillips DJ, Tchongue J, de Kretser DM, Sievert W. Follistatin attenuates early liver fibrosis: effects on hepatic stellate cell activation and hepatocyte apoptosis. Am J Physiol Gastrointest Liver Physiol. 2006 Jan;290(1):G137-144.

102. Crawford R, Prescott E, Sylvester C, Higdon A, Shan J, Kilberg M, et al. Human CHAC1 protein degrades glutathione, and mRNA induction is regulated by the transcription factors ATF4 and ATF3 and a bipartite ATF/CRE regulatory element. J Biol Chem. 2015 Apr 30;290.

103. Asahina K. Hepatic stellate cell progenitor cells. J Gastroenterol Hepatol. 2012 Mar;27(Suppl 2):80–4.

104. Yin C, Evason KJ, Asahina K, Stainier DYR. Hepatic stellate cells in liver development, regeneration, and cancer. J Clin Invest. 2013 May;123(5):1902–10.

105. Dirscherl K, Schläpfer M, Roth Z'graggen B, Wenger RH, Booy C, Flury-Frei R, et al. Hypoxia sensing by hepatic stellate cells leads to VEGF-dependent angiogenesis and may contribute to accelerated liver regeneration. Sci Rep. 2020 Mar 9;10(1):4392.

106. Goetzman ES, Prochownik EV. The Role for Myc in Coordinating Glycolysis, Oxidative Phosphorylation, Glutaminolysis, and Fatty Acid Metabolism in Normal and Neoplastic Tissues. Front Endocrinol [Internet]. 2018 Apr 12 [cited 2021 Apr 5];9. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5907532/

107. Tian J, Locker J. Gadd45 in the liver: signal transduction and transcriptional mechanisms. Adv Exp Med Biol. 2013;793:69–80.

108. Sukonina V, Ma H, Zhang W, Bartesaghi S, Subhash S, Heglind M, et al. FOXK1 and FOXK2 regulate aerobic glycolysis. Nature. 2019 Feb;566(7743):279–83.

109. Flores-Téllez TNJ, Lopez TV, Vásquez Garzón VR, Villa-Treviño S. Co-Expression of Ezrin-CLIC5-Podocalyxin Is Associated with Migration and Invasiveness in Hepatocellular Carcinoma. PLoS ONE [Internet]. 2015 Jul 2

[cited 2021 Mar 26];10(7). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4489913/

110. Jelkmann W, Pagel H, Hellwig T, Fandrey J. Effects of antioxidant vitamins on renal and hepatic erythropoietin production. Kidney Int. 1997 Feb;51(2):497–501.

111. Fan J, Rone MB, Papadopoulos V. Translocator Protein 2 Is Involved in Cholesterol Redistribution during Erythropoiesis. J Biol Chem. 2009 Oct 30;284(44):30484–97.

112. Vera JC, Rivas CI, Zhang RH, Farber CM, Golde DW. Human HL-60 myeloid leukemia cells transport dehydroascorbic acid via the glucose transporters and accumulate reduced ascorbic acid. Blood. 1994 Sep 1;84(5):1628–34.

113. Goldenberg H, Schweinzer E. Transport of vitamin C in animal and human cells. J Bioenerg Biomembr. 1994 Aug;26(4):359–67.

114. Upston JM, Karjalainen A, Bygrave FL, Stocker R. Efflux of hepatic ascorbate: a potential contributor to the maintenance of plasma vitamin C. Biochem J. 1999 Aug 15;342 (Pt 1):49–56.

115. Wang Y, Zhang C. The Roles of Liver-Resident Lymphocytes in Liver Diseases. Front Immunol [Internet]. 2019 [cited 2021 Mar 31];10. Available from: https://www.frontiersin.org/articles/10.3389/fimmu.2019.01582/full

116. Li S, Tan H-Y, Wang N, Feng Y, Wang X, Feng Y. Recent Insights Into the Role of Immune Cells in Alcoholic Liver Disease. Front Immunol [Internet]. 2019 [cited 2021 Apr 6];10. Available from: https://www.frontiersin.org/articles/10.3389/fimmu.2019.01328/full

117. O'Neill LAJ, Artyomov MN. Itaconate: the poster child of metabolic reprogramming in macrophage function. Nat Rev Immunol. 2019 May;19(5):273–81.

118. Williams NC, O'Neill LAJ. A Role for the Krebs Cycle Intermediate Citrate in Metabolic Reprogramming in Innate Immunity and Inflammation. Front Immunol [Internet]. 2018 [cited 2021 Mar 31];9. Available from: https://www.frontiersin.org/articles/10.3389/fimmu.2018.00141/full

119. Mills E, O'Neill LAJ. Succinate: a metabolic signal in inflammation. Trends Cell Biol. 2014 May 1;24(5):313–20.

120. Shomali N, Mahmoudi J, Mahmoodpoor A, Zamiri RE, Akbari M, Xu H, et al. Harmful effects of high amounts of glucose on the immune system: An updated review. Biotechnol Appl Biochem. 2020 May 12;

121. Wolowczuk I, Verwaerde C, Viltart O, Delanoye A, Delacre M, Pot B, et al. Feeding our immune system: impact on metabolism. Clin Dev Immunol. 2008;2008:639803.

122. Palazon A, Goldrath A, Nizet V, Johnson RS. HIF Transcription Factors, Inflammation, and Immunity. Immunity. 2014 Oct 16;41(4):518–28.

123. Franchina DG, Grusdat M, Brenner D. B-Cell Metabolic Remodeling and Cancer. Trends Cancer. 2018 Feb 1;4(2):138–50.

124. de Mingo Pulido Á, de Gregorio E, Chandra S, Colell A, Morales A, Kronenberg M, et al. Differential Role of Cathepsins S and B In Hepatic APC-Mediated NKT Cell Activation and Cytokine Secretion. Front Immunol [Internet]. 2018 Feb 28 [cited 2021 Apr 3];9. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5836516/

125. Maedera S, Mizuno T, Ishiguro H, Ito T, Soga T, Kusuhara H. GLUT6 is a lysosomal transporter that is regulated by inflammatory stimuli and modulates glycolysis in macrophages. FEBS Lett. 2019 Jan;593(2):195–208.

126. Mattiotti A, Prakash S, Barnett P, van den Hoff MJB. Follistatin-like 1 in development and human diseases. Cell Mol Life Sci. 2018;75(13):2339–54.

127. Shang H, Liu X, Guo H. Knockdown of Fstl1 attenuates hepatic stellate cell activation through the TGF-β1/Smad3 signaling pathway. Mol Med Rep. 2017 Nov 1;16(5):7119–23.

128. Vooght KMK de, Solinge WW van, Wesel AC van, Kersting S, Wijk R van. First mutation in the red blood cell-specific promoter of hexokinase combined with a novel missense mutation causes hexokinase deficiency and mild chronic hemolysis. Haematologica. 2009 Sep 1;94(9):1203–10.

129. Tibshirani R. Regression Shrinkage and Selection Via the Lasso. J R Stat Soc Ser B. 1994;58:267–88.

130. Efron B. Bootstrap Methods: Another Look at the Jackknife. Ann Stat. 1979 Jan;7(1):1–26.

131. Chowdhury S, Yung E, Pintilie M, Muaddi H, Chaib S, Yeung M, et al. MATE2 Expression Is Associated with Cancer Cell Response to Metformin.

PLoS ONE [Internet]. 2016 Dec 13 [cited 2021 Apr 1];11(12). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5154501/

132. Wittwer MB, Zur AA, Khuri N, Kido Y, Kosaka A, Zhang X, et al. Discovery of Potent, Selective Multidrug And Toxin Extrusion Transporter 1 (MATE1, SLC47A1) Inhibitors Through Prescription Drug Profiling and Computational Modeling. J Med Chem. 2013 Feb 14;56(3):781–95.

133. Saha D, Patgaonkar M, Shroff A, Ayyar K, Bashir T, Reddy KVR. Hemoglobin Expression in Nonerythroid Cells: Novel or Ubiquitous? Int J Inflamm. 2014 Nov 5;2014:e803237.

134. Newton DA, Rao KMK, Dluhy RA, Baatz JE. Hemoglobin Is Expressed by Alveolar Epithelial Cells *. J Biol Chem. 2006 Mar 3;281(9):5668–76.

135. Nishi H, Inagi R, Kato H, Tanemoto M, Kojima I, Son D, et al. Hemoglobin Is Expressed by Mesangial Cells and Reduces Oxidant Stress. J Am Soc Nephrol. 2008 Aug 1;19(8):1500–8.

136. Liu W, Baker SS, Baker RD, Nowak NJ, Zhu L. Upregulation of Hemoglobin Expression by Oxidative Stress in Hepatocytes and Its Implication in Nonalcoholic Steatohepatitis. PLoS ONE [Internet]. 2011 Sep 12 [cited 2021 Mar 25];6(9). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3171444/

137. Tsiftsoglou AS, Vizirianakis IS, Strouboulis J. Erythropoiesis: Model systems, molecular regulators, and developmental programs. IUBMB Life. 2009;61(8):800–30.

138. Guedes PT, Oliveira BCEPD de, Manso PP de A, Caputo LFG, Cotta-Pereira G, Pelajo-Machado M. Histological Analyses Demonstrate the Temporary Contribution of Yolk Sac, Liver, and Bone Marrow to Hematopoiesis during Chicken Development. PLOS ONE. 2014 Mar 12;9(3):e90975.

139. Stier A, Bize P, Schull Q, Zoll J, Singh F, Geny B, et al. Avian erythrocytes have functional mitochondria, opening novel perspectives for birds as animal models in the study of ageing. Front Zool. 2013 Jun 8;10:33.

140. Brown KA. Erythrocyte Metabolism and Enzyme Defects. Lab Med. 1996 May 1;27(5):329–33.

141. Murphy JR. Erythrocyte metabolism. II. Glucose metabolism and pathways. J Lab Clin Med. 1960 Feb;55:286–302.

142. Kinoshita A. Simulation of Human Erythrocyte Metabolism [Internet]. Madame Curie Bioscience Database [Internet]. Landes Bioscience; 2013 [cited 2021 Mar 31]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK6563/

143. Hadi HE, Vettor R, Rossato M. Vitamin E as a Treatment for Nonalcoholic Fatty Liver Disease: Reality or Myth? Antioxidants [Internet]. 2018 Jan 16 [cited 2021 Apr 1];7(1). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5789322/

144. Lee J, Lee J, Kim Y, Lee H, Jun H-J, Lee S-J. Multidrug and Toxic Compound Extrusion Protein-1 (MATE1/SLC47A1) is a Novel Flavonoid Transporter. J Agric Food Chem. 2014 Sep 19;62.

Appendix A

IACUC APPROVAL

University of Delaware
Institutional Animal Care and Use Committee
Application to Use Animals in Application to use animals in Research
(New and 3-Yr submission)

Title of Protocol: Transcriptome, proteome and metabolome studies of organ growth in the post hatch chicken	
AUP Number: 72R-2017-0	**← (4 digits only — if new, leave blank)**
Carl J. Schmidt	
Common Name (Strain/Breed if Appropriate): Chicken **Genus Species:** Gallus gallus	
Date of Submission: 01/06/2017	

Official Use Only

IACUC Approval Signature: Jan Talbert, DVM

Date of Approval: 2/17/2017

Appendix B

PERMISSIONS

Chapter 2 is accepted for publication to BMC Genomics, an open-access journal. The licensing information regarding permission to publish elsewhere can be found at the following link: https://www.biomedcentral.com/about/policies/license-agreement