UC Davis
UC Davis Previously Published Works

Title

Novel Combined Tissue Transcriptome Analysis After Lentogenic Newcastle Disease Virus Challenge in Inbred Chicken Lines of Differential Resistance

Permalink

https://escholarship.org/uc/item/94h3781w

Authors

Deist, Melissa S
Gallardo, Rodrigo A
Dekkers, Jack CM
et al.

Publication Date

2020

DOI

10.3389/fgene.2020.00011

Peer reviewed

Novel Combined Tissue Transcriptome Analysis After Lentogenic Newcastle Disease Virus Challenge in Inbred Chicken Lines of Differential Resistance

Melissa S. Deist[1], Rodrigo A. Gallardo[2], Jack C. M. Dekkers[1], Huaijun Zhou[3] and Susan J. Lamont[1]*

[1] Department of Animal Science, Iowa State University, Ames, IA, United States, [2] Department of Population Health and Reproduction, School of Veterinary Medicine, University of California, Davis, Davis, CA, United States, [3] Department of Animal Science, University of California, Davis, Davis, CA, United States

Disease has large negative impacts on poultry production. A more comprehensive understanding of host–pathogen interaction can lead to new and improved strategies to maintain health. In particular, host genetic factors can lead to a more effective response to pathogens, hereafter termed resistance. Fayoumi and Leghorn chicken lines have demonstrated relative resistance and susceptibility, respectively, to the Newcastle disease virus (NDV) vaccine strain and many other pathogens. This biological model was used to better understand the host response to a vaccine strain of NDV across three tissues and time points, using RNA-seq. Analyzing the Harderian gland, trachea, and lung tissues together using weighted gene co-expression network analysis (WGCNA) identified important genes that were co-expressed and associated with parameters including: genetic line, days post-infection (dpi), challenge status, sex, and tissue. Pathways and driver genes, such as *EIF2AK2*, *MPEG1*, and *TNFSF13B*, associated with challenge status, dpi, and genetic line were of particular interest as candidates for disease resistance. Overall, by jointly analyzing the three tissues, this study identified genes and gene networks that led to a more comprehensive understanding of the whole animal response to lentogenic NDV than that obtained by analyzing the tissues individually.

Keywords: chickens, RNA-seq, weighted gene co-expression network analysis, trachea, lung, Harderian gland

INTRODUCTION

Newcastle disease virus (NDV) is a single-stranded, negative-sense RNA virus. The clinical signs associated with NDV infection in chickens range from subclinical to 100% mortality depending on the virulence of the infection strain (Miller and Koch, 2013). NDV is classified into five pathotypes based on clinical signs: asymptomatic enteric, lentogenic, mesogenic, velogenic viscerotropic, and velogenic neurotropic (Afonso et al., 2012). Velogenic strains, also known in the United States as exotic NDV (END virus), have the most detrimental impacts on the poultry industry. Infection with

velogenic neurotropic NDV can result in neurological clinical signs, whereas velogenic viscerotropic strains cause intestinal lesions (Afonso et al., 2012). Both velogenic strains result in high morbidity and mortality (Afonso et al., 2012). The lentogenic strains are the least virulent, and because all NDV strains belong to the same serotype, lentogenic strains are used for vaccination (Dimitrov et al., 2017). Lentogenic NDV is limited in its replication to host cells that produce a trypsin-like protease to cleave the fusion protein (Nagai et al., 1976). This restricts lentogenic NDV replication to epithelial cells where this protease is expressed (Nagai, 1995). Hence, lentogenic strains have the largest impact near sites of infection such as the trachea, lung, and Harderian gland, all of which contain epithelial cells and are also involved in mucosal immunity. These tissues, in which the virus and host have direct contact, are ideal for examining host–pathogen interaction. Increased understanding of the host–pathogen interaction creates the potential for improving strategies to curb the negative impacts of NDV.

Two inbred chicken lines have been shown to differ in their susceptibility to lentogenic La Sota NDV (Deist et al., 2017b). The Leghorn line (GHs 6) originated from a cross of two commercial lines in the 1950s and the Fayoumi line (M 15.2) originated from the Fayoum region of Egypt (Lamont and Chen, 1992). The Fayoumi and Leghorn lines are each greater than 99.95% homozygous (Fleming et al., 2016) and have different MHC types 15.2 and 6, respectively. These lines have also shown differential responses to avian influenza (Wang et al., 2014), Marek's disease (Lakshmanan et al., 1996), *Salmonella* (Cheeseman et al., 2007), and *Eimeria* (Pinard-van der Laan et al., 2009). Inbred lines that differ in their relative susceptibility offer an excellent tool to study mechanisms of pathogen response.

The RNA sequencing (RNA-seq) approach shows the genes and pathways impacted by a treatment; however, it only shows a snapshot of response in a particular tissue and time. Because the transcriptome is not stagnant, a more comprehensive physiological picture can be obtained by an integrative analysis of the transcriptomes of multiple tissues at multiple time points after a treatment. Previously, the trachea epithelial cells, lung, and Harderian gland of Fayoumis and Leghorns at 2, 6, and 10 days post-infection (dpi) with lentogenic NDV have been individually analyzed with RNA-seq (Deist et al., 2017a; Deist et al., 2017b; Deist et al., 2018). The three tissues had very distinct responses to NDV (Deist et al., 2017a; Deist et al., 2017b; Deist et al., 2018). Taking a systems biology approach and analyzing the three tissues together will provide a more comprehensive picture of how the whole animal responded to NDV. Prior joint-tissue analyses using microarray or RNA-seq to study the impact of avian pathogenic *Escherichia coli* on broiler chickens have been performed (Sandford et al., 2012; Sun et al., 2016). Improvements in the technologies and tools for analyzing expression data now enable a more thorough examination of these complex data.
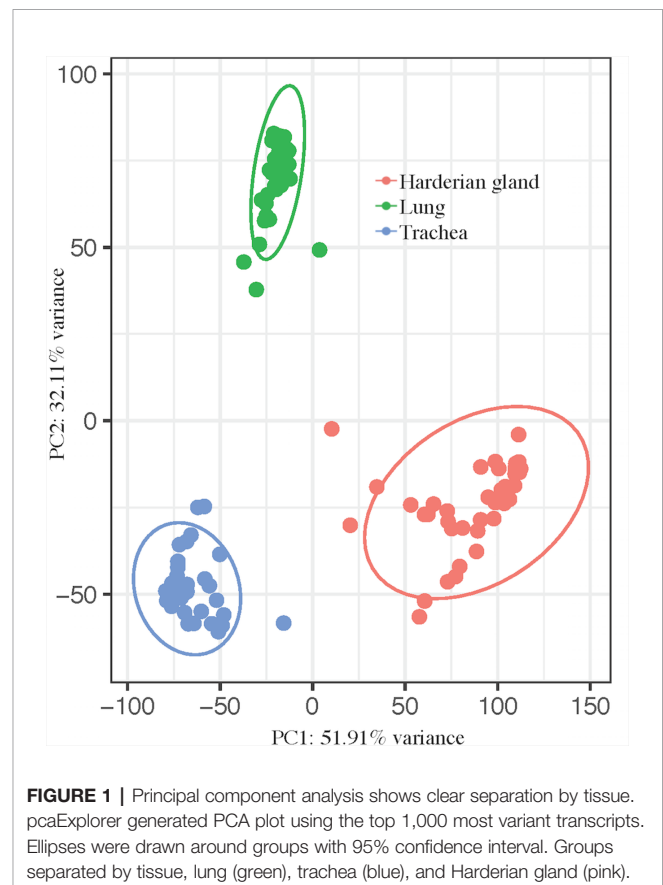
The objective of this study was to gain a more comprehensive understanding of the transcriptomic response of these two chicken lines to lentogenic NDV. Utilizing co-expression analysis across all three tissues identifies clusters of co-expressed genes that are associated with factors of interest such

as NDV challenge, line, and dpi. We expect to identify novel genes and gene clusters that were not previously identified when the tissues were analyzed individually. Clusters of genes that are associated with NDV challenge and line may help identify genes potentially associated with resistance to lentogenic NDV. These data from multiple key tissues after a lentogenic challenge lay the foundation for future comparison of networks and pathways that may differ in experimental parameters including use of velogenic NDV challenge.

RESULTS

Principal Component Analysis of the Trachea, Lung, and Harderian Gland

The transcriptomes of three tissues from birds infected with lentogenic NDV were jointly analyzed to more comprehensively characterize the host response to NDV. Principal component analysis showed a clear separation between the tissues, as expected (**Figure 1**). Principal component (PC) 1 (PC1 = 51.91%) separated all three tissues and PC2 (32.11%) separated the trachea and Harderian gland from the lung (**Figure 1**). No clear clustering was observed by any other parameter in all tissues.

FIGURE 1 | Principal component analysis shows clear separation by tissue. pcaExplorer generated PCA plot using the top 1,000 most variant transcripts. Ellipses were drawn around groups with 95% confidence interval. Groups separated by tissue, lung (green), trachea (blue), and Harderian gland (pink).

Differential Expression Analysis

A differential expression analysis was performed to determine if there was a difference in transcript expression levels between tissues after challenge with lentogenic NDV. The numbers of differentially expressed genes (DEG) for the challenge by tissue interaction, within each line and dpi, are shown in **Table 1**. Large numbers of DEG were identified by the interaction between the

TABLE 1 | Comparing the tissues' response to lentogenic challenge within each line and time.

Line	Time	Interaction contrast[a]	DEG[b]
Fayoumi	2 dpi	Lung vs. Harderian gland	10
		Trachea vs. Harderian gland	151
		Trachea vs. lung	135
	6 dpi	Lung vs. Harderian gland	0
		Trachea vs. Harderian gland	1
		Trachea vs. lung	3
	10 dpi	Lung vs. Harderian gland	1,523
		Trachea vs. Harderian gland	54
		Trachea vs. lung	20
Leghorn	2 dpi	Lung vs. Harderian gland	0
		Trachea vs. Harderian gland	141
		Trachea vs. lung	115
	6 dpi	Lung vs. Harderian gland	743
		Trachea vs. Harderian gland	877
		Trachea vs. lung	226
	10 dpi	Lung vs. Harderian gland	0
		Trachea vs. Harderian gland	0
		Trachea vs. lung	0

[a]Within each line and dpi: (Tissue1, Challenged–Tissue1, Non-challenged)–(Tissue2, Challenged–Tissue2, Non-challenged).
[b]Differentially Expressed Genes (FDR < 0.05).

lung and Harderian gland in the Fayoumis at 10 dpi (**Table 1**), and separately the large numbers of DEG in the Leghorns at 6 dpi (**Table 1**) may have resulted from large numbers of DEG due to challenge within an individual tissue (Deist et al., 2017a; Deist et al., 2018). With the exception of the Fayoumis at 10 dpi and the Leghorns at 6 dpi, there were few DEG between the lung and Harderian gland. Interestingly, there were no DEG at 10 dpi in the Leghorns and few DEG in the Fayoumis at 6 dpi (**Table 1**). Overall, there were differences in the number of DEG from each interaction contrast, depending on genetic line and time. Of the 18 contrasts examined, *CD8B* and *LY6D* were differentially expressed most frequently in seven and six contrasts, respectively. Full lists of DEG (FDR < 0.05) for each contrast are provided in **Table S1**.

Gene Co-Expression Analysis in the Trachea, Lung, and Harderian Gland After Lentogenic Challenge

The module–trait relationship analysis resulted in several associations between the modules generated by weighted gene co-expression network analysis (WGCNA) based on co-expression and factors of the study: line, dpi, challenge status, sex, and tissue (**Figure 2**). The lightyellow module strongly positively correlated with line (0.99 for Fayoumi = 1 versus Leghorn = 0, $p = 1e-120$), but did not significantly correlate with any other factor, suggesting that for the 161 transcripts within this module, there was no association between line and any other factor. The lightyellow module could help better characterize the inherent differences between these two inbred genetic lines. The orange module was strongly positively correlated with sex (0.97
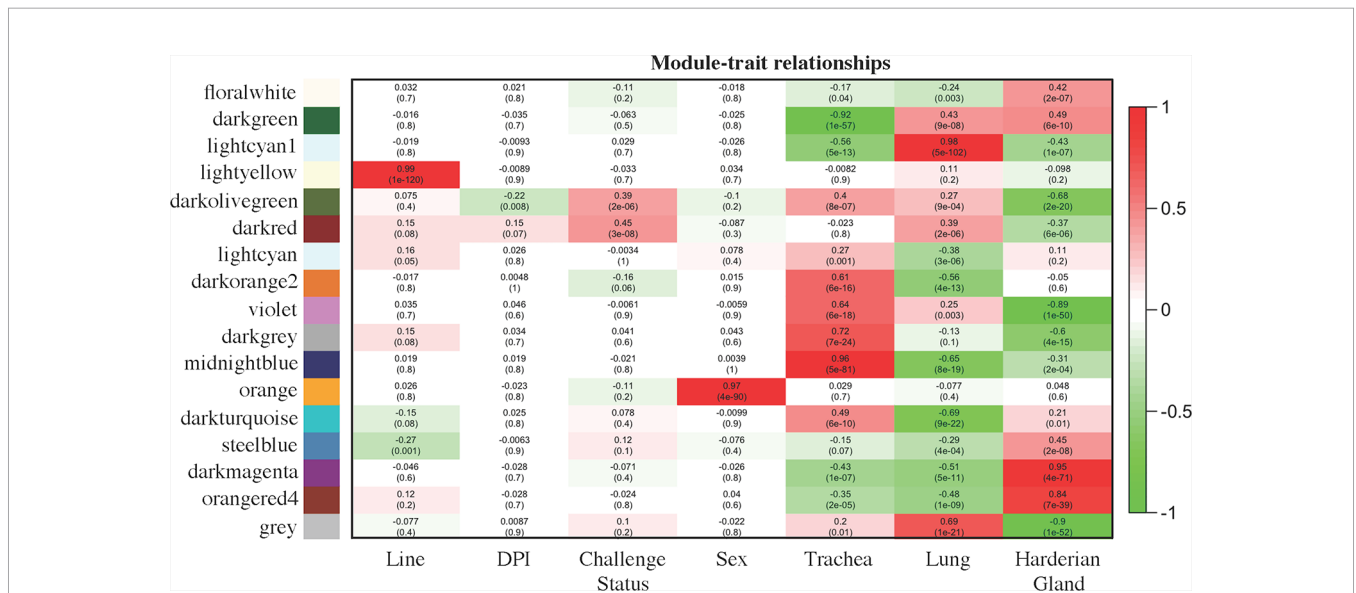

FIGURE 2 | Module–trait relationships calculated by WGCNA. Factors of interest (x-axis) were correlated to each module (y-axis). The first value within each square is the correlation and the second value in parenthesis is the p value of the association. The more positively correlated the module and the factor, the more red the square; the more negatively correlated, the more green. Coding for the factors was as follows: Line [Fayoumi (1), Leghorn (0)], days post-infection [dpi; 2 dpi (0), 6 dpi (1), 10 dpi (2)], NDV [challenged (1), non-challenged (0)], sex [male (1), female (0)], Trachea [Trachea (1), Lung (0), Harderian gland (0)], Lung [Trachea (0), Lung (1), Harderian gland (0)], and Harderian gland [Trachea (0), Lung (0), Harderian gland (1)].

for male = 1 versus female = 0, p = 1e−90). Within this module, 94 are on the Z chromosome, 7 are on the W chromosome, and 17 transcripts have not been assigned to a chromosome. Lack of correlations with other modules indicates that these transcripts are only impacted by sex.

Three modules, midnightblue, lightcyan1, and darkmagenta, may be related to the specific functions of each tissue. The 1,308 transcripts in the midnightblue module were strongly positively correlated and more highly expressed in the trachea (0.96 for trachea = 1 versus lung and Harderian gland = 0, p = 5e−81). The 1,720 transcripts in the lightcyan1 module were strongly positively correlated with and more highly expressed in the lung (0.98 for lung = 1 versus trachea and Harderian gland = 0, p = 5e−102). The large darkmagenta module (3,154 transcripts) strongly positively correlated with and more highly expressed in the Harderian gland (0.95 for Harderian gland = 1 versus trachea and lung = 0, p = 4e−71).

The darkolivegreen module was negatively correlated with dpi (−0.22 for 2 dpi = 0, 6 dpi = 1, versus 10 dpi = 2, p = 0.0008), positively correlated with challenge status (0.39 for challenged = 1 versus non-challenged = 0, p = 2e−06), and negatively correlated with the Harderian gland (−0.68 for Harderian gland = 1 versus trachea and lung = 0, p = 2e−20). The 80 transcripts in this module were likely important for the chicken's early response to NDV because the negative correlation with dpi indicates they were more highly expressed at early time points. The darkred module was positively correlated with challenge status (0.45 for challenged = 1 versus non-challenged = 0, p = 3e−08), positively correlated with the lung (0.39 for lung = 1 versus trachea and Harderian gland = 0, p = 2e−0.6), and negatively correlated with the Harderian gland (−0.37 for Harderian gland = 1 versus trachea and lung = 0, p = 6e−06). Although not significant, this module showed a weak correlation with line (0.15 for Fayoumi = 1 versus Leghorn = 0, p = 0.08). The darkred module potentially included genes related to response to NDV because of its associations with challenge status and line. The 333 transcripts in the steelblue module negatively correlated with line (−0.27 for Fayoumi = 1 versus Leghorn = 0, p = 0.001) and positively correlated with the Harderian gland (0.45 for Harderian gland = 1 versus trachea and lung = 0, p = 4e−71). Although not significant, this module weakly correlated with challenge status (0.12 for challenged = 1 versus non-challenged = 0, p = 0.1). The steelblue module may also include transcripts potentially associated with response to NDV.

Genes that are co-expressed are likely to be regulated by similar mechanisms. Within each module, the transcripts with high gene significance (GS) and module membership (MM), i.e., driver genes, were likely to have the most influence on the expression values of the entire module. The top driver genes were identified for the module that was most highly correlated with each factor (**Table 2**). Several long intervening/intergenic non-coding RNAs (lincRNAs) and unnamed protein coding genes were identified as driver genes (**Table 2**). The *NKX2-1* was among the top 10 bottom-loading genes for PC1, and *CLDN18* among the top 10 top-loading genes for PC2 calculated by pcaExplorer (Marini, 2016).

TABLE 2 | Top 3 driver genes from the module most highly correlated with each factor for birds infected with lentogenic NDV.

Factor / Module	Transcript ID / Gene Name	Gene significance	Module membership
Line/lightyellow	ENSGALT00000082473/ lincRNA	0.97	0.96
	ENSGALT00000065772/ lincRNA	−0.97	−0.97
	ENSGALT00000087010/ protein coding	0.96	0.97
DPI/darkolivegreen	ENSGALT00000017183/ EIF2AK2	−0.37	0.83
	ENSGALT00000016499/ CMTR1	−0.37	0.74
	ENSGALT00000066953/ protein coding	−0.35	0.90
Challenge status/ darkred	ENSGALT00000027228/ TNFSF13B	0.53	0.82
	ENSGALT00000044004/ MPEG1	0.52	0.82
	ENSGALT00000003281/ protein coding	0.51	0.74
Sex/orange	ENSGALT00000077789/ Nipped-B homolog-like	−0.98	−0.95
	ENSGALT00000047883/ protein coding	−0.98	−0.96
	ENSGALT00000080994/ protein coding	−0.97	−0.95
Trachea/ midnightblue	ENSGALT00000021346/akr	0.97	0.96
	ENSGALT00000049178/ DNAL4	0.97	0.95
	ENSGALT00000009017/F3	0.97	0.94
Lung/lightcyan1	ENSGALT00000089244/ protein coding	0.99	0.98
	ENSGALT00000010644/ CLDN18	0.99	0.97
	ENSGALT00000057049/ lincRNA	0.98	0.97
Harderian gland/ dark magenta	ENSGALT00000059647/ LRRC30	0.98	0.98
	ENSGALT00000006028/ PABPC4	0.97	0.94
	ENSGALT00000059800/ NKX2-1	−0.97	−0.94

Functional Analysis of Modules of Interest for Birds Infected With Lentogenic NDV

The darkolivegreen, darkred, and steelblue modules were of particular interest because of their correlations with either challenge status or line. The darkolivegreen module's correlations with dpi and challenge status suggest that these genes were important in the early response to NDV. Of the 80 transcripts, 59 had an associated gene name. No modules were significantly correlated with both the challenge status and line; however, the darkred and steelblue modules may include genes associated with response to NDV because of their correlations with challenge status or line. Gene ontology (GO) term and network analysis was performed on these three modules.

Panther identified significant top-level GO terms for the darkolivegreen, darkred, and steelblue modules (**Figure 3**). The significant GO terms for the darkolivegreen and darkred modules, both significantly correlated with challenge status,
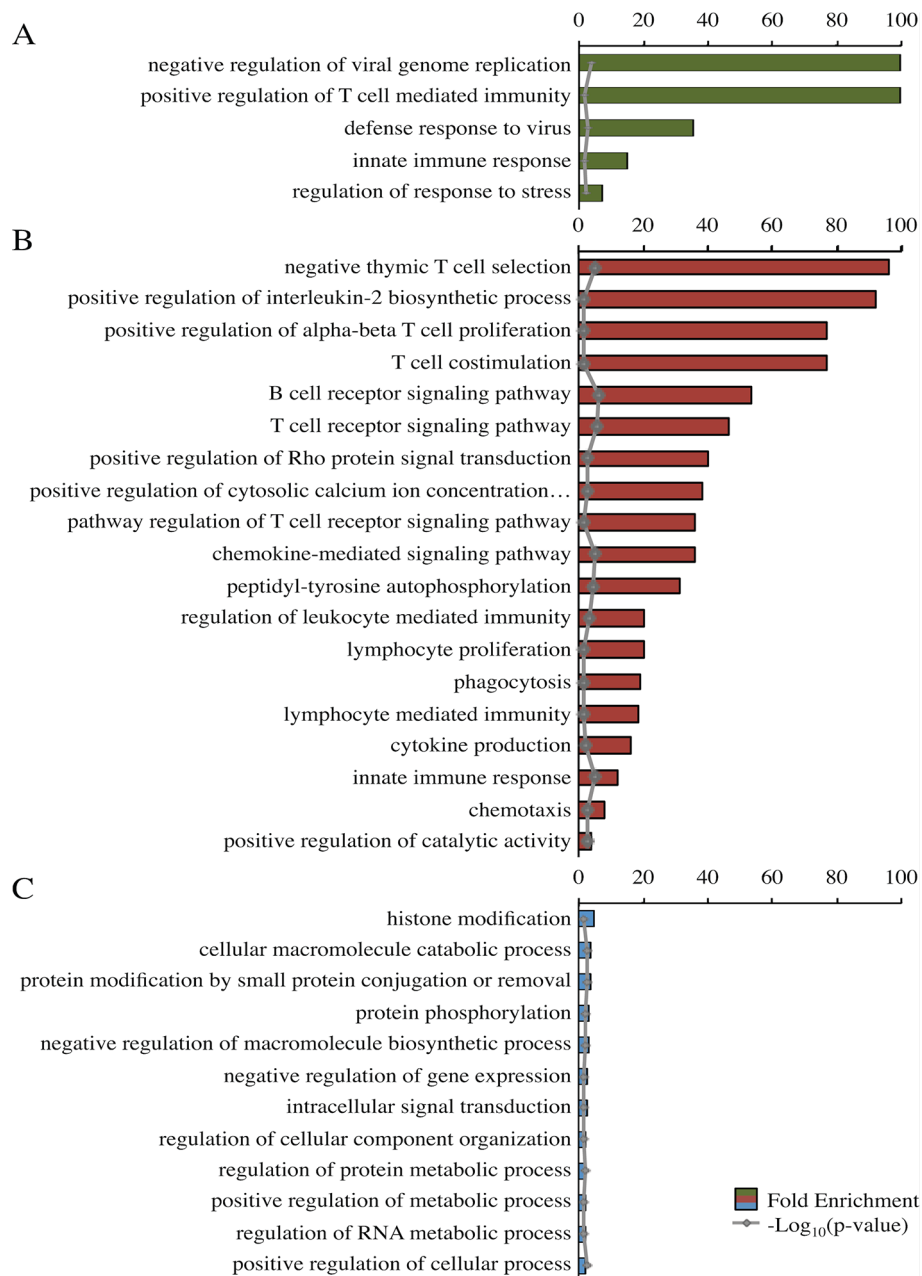
FIGURE 3 | Fold enrichment of top-level overrepresented GO terms within modules of interest. Panther overrepresentation test using associated gene names for each of three modules: darkolivegreen module **(A)**, darkred module **(B)**, and steel blue module **(C)**. Fold enrichment was capped at 100. −Log10(p value) shown in gray. All p values <0.05. Ellipses indicate those involved in phospholipase C-activating G-protein-coupled signaling.

were clearly immune related. The significant GO terms for the steelblue module were more general in nature (**Figure 3**).

To generate protein interaction networks for the transcripts in each module of interest, STRING was utilized (**Figure 4**). The darkolivegreen, darkred, and steelblue networks had significantly more connections than would be expected by chance. The darkolivegreen module included 55 nodes and 31 edges and was significantly associated with several KEGG pathways that are

related to response to virus, including influenza A, Herpes simplex infection, and the RIG-I-like receptor signaling pathway (**Figure 4A**). Within the darkolivegreen module, *STAT1* had the second highest gene significance for challenge status and a high module membership (GS = 0.50, MM = 0.88), making the highly connected protein a potential driver gene. The darkred module network included 110 nodes and 60 edges and was significantly associated with the following KEGG pathways:
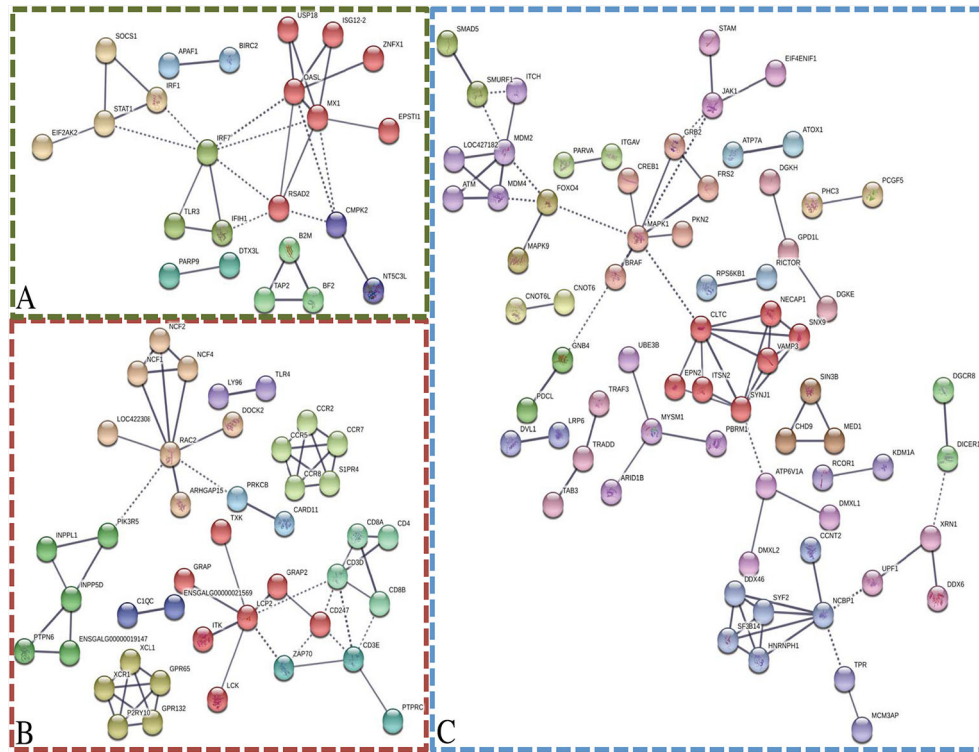
FIGURE 4 | Network analysis of modules of interest. STRING was utilized to generate networks based on the associated gene name of the transcripts within the modules of interest: darkolivegreen module **(A)**, darkred module **(B)**, and steelblue module **(C)**. Disconnected nodes were removed, high confidence (0.700), and MCL clustering (inflation = 3). Solid lines indicate connections within cluster and dotted lines indicate connections between clusters.

cytokine–cytokine receptor interaction, cell adhesion molecules (CAMs), phagosome, Jak-STAT signaling pathway, intestinal immune network for IgA production, and phosphatidylinositol signaling system. The fourth highest driver gene, *INPP5D* (GS = 0.49, MM = 0.84), was moderately well connected (**Figure 4B**). The network generated based on the steelblue module included 251 nodes and 78 edges and was significantly associated with only one KEGG pathway: endocytosis. *DGKE* and *RICTOR* were among the top 4 driver genes in the steelblue module based on gene significance for line (**Figure 4C**).

DISCUSSION

The combined tissue analysis of the current study provided a comprehensive picture of the Fayoumi and Leghorn transcriptomic response to lentogenic NDV challenge in tissues where the NDV vaccine replicates. The Harderian gland, lung, and tracheal epithelial cells come into direct contact with the lentogenic virus when infected *via* the eyes and/or airways and provide the best opportunity to examine the host–pathogen interaction. Previously reported analysis of each individual tissue's transcriptome resulted in unique findings for each tissue, as briefly summarized here. The trachea epithelial cells had the highest amount of detectable virus (Deist et al., 2017b).

As time progressed, the number of DEG decreased between the challenged and non-challenged birds in the trachea, suggesting the trachea was consistently active in host response and in recovery from the virus in both lines (Deist et al., 2017b). The transcriptomes of the lung and Harderian gland were more impacted by line than by the challenge (Deist et al., 2017a; Deist et al., 2018). The Leghorn lung appeared to be non-responsive to the virus, whereas the Fayoumi lung produced a large number of DEG at 10 dpi when comparing the challenged to non-challenged birds (Deist et al., 2017a). The Harderian gland of the Leghorns showed a large number of DEG between the challenged and non-challenged birds at 6 dpi, and the Fayoumis had very few DEG at all time points (Deist et al., 2018). The challenge by tissue interaction resulted in more DEG if there were more DEG between challenged and non-challenged birds within a tissue. It is difficult to interpret the results of the challenge by tissue interaction because the tissues had time- and line-dependent responses to the virus as measured by DEG and had different amounts of detectable virus (Deist et al., 2017a; Deist et al., 2017b; Deist et al., 2018).

The complex design of the current combined, three-tissue experiment did not readily lend itself to a differential expression analysis because with three tissues, three time points, two genetic lines, and two challenge statuses, there was a very large number of possible contrasts. Using WGCNA for co-expression analysis does

not rely on specific contrasts (differential expression) and can identify correlations between co-expressed genes and factors of importance in the study design. Using WGCNA to analyze multiple tissues is a novel approach that was successful. Previously, WGCNA was used to analyze the lung tissue individually (Deist et al., 2017a), which had some results in concordance with the combined tissue analysis. ENSGALT00000087010 was among the top driver genes for genetic line in the lung analysis as well as the joint analysis. In addition, *Nipped-B homolog-like* and ENSGALT00000080994 were among the top driver genes for sex in both analyses.

The darkolivegreen module was composed of genes that had a clear functional role in host defense. The unannotated transcripts within the darkolivegreen module should be explored further to identify their functions because they were co-expressed with genes important to the host response to NDV. GO term analysis of the genes in the darkolivegreen module showed they were involved in the host response to virus. *EIF2AK2* (also known as *PKR*) was the top driver gene for dpi for the darkolivegreen module. *EIF2AK2* is interferon-inducible and has important antiviral effects (Sadler and Williams, 2008). The EIF2 signaling pathway is known to inhibit NDV replication (Zhang et al., 2014). This pathway was shown to be significantly different between the two lines in both the trachea (Deist et al., 2017b) and spleen transcriptomes (Zhang et al., 2018). *EIF2AK2* phosphorylates *EIF2α*, which regulates protein synthesis (Deb et al., 2001). Also, treatment with poly(I:C) led to expression changes of *EIF2AK2* in chickens (Kim and Zhou, 2015). The negative correlation of the darkolivegreen module with dpi suggests *EIF2AK2* is likely more highly expressed at earlier times post-infection. Differential expression confirmed that the challenged birds had significantly higher *EIF2AK2* expression levels at 2 dpi when compared to 6 or 10 dpi across all tissues (data not shown). At 2 dpi, the highest levels of viral transcripts were detected in the trachea (Deist et al., 2017b) and Harderian gland (Deist et al., 2018), as measured by RNA-seq; high expression of *EIF2AK2* at 2 dpi would aid in viral clearance.

The functions of the genes in the darkred module involve both the innate and adaptive immune system, as determined by the GO term analysis and the top 2 driver genes. It is well known that TNF genes play an important role in response to pathogens. *TNFSF13B* (also known as *BAFF*) specifically aids in survival and proliferation of chicken B cells (Schneider, 2004), and was the top driver gene for challenge status in the darkred module. *TNFSF13B* receptors are expressed predominantly on chicken B cells (Schneider, 2004). One receptor for *TNFSF13B*, *TNFRSF13B*, was among the genes significant in the trachea challenge by line interaction at 6 dpi (Deist et al., 2017b) and was differentially expressed between the Fayoumi and Leghorn chickens in the lung in both non-challenged birds and birds challenged with avian influenza (Wang et al., 2014). Antibody production is required for NDV clearance (Reynolds and Maraqa, 2000; Kapczynski et al., 2013), and expression of *TNFSF13B* may critically influence B cells and their antibody production in response to NDV.

Another driver gene for challenge status in the darkred module was *MPEG1*. *MPEG1* (also known as *Perforin-2*) is expressed highly in macrophages and is an important component of the innate

immune system (McCormack and Podack, 2015). The protein encoded by *MPEG1* works with complement to form membrane attack complexes and create holes in a target cell's membrane (McCormack and Podack, 2015). *MPEG1* was 1 of 19 DEG in the Harderian gland (Deist et al., 2018) and 1 of 16 DEG in the lung (Deist et al., 2017a) between challenged and non-challenged Leghorns at 2 dpi. The infiltration of macrophages into tissues infected with NDV may result in an increase in RNA expression of MPEG1.

The two chicken lines have previously shown differences in the expression levels of *TNFSF13B* and *MPEG1* in response to viral pathogens, which may suggest a potential role in disease resistance (Deist et al., 2017b; Deist et al., 2018). *MPEG1* and *TNFSF13B*, however, were not included in the darkred network. This suggests that these proteins may have unknown or low confidence connections in STRING to the other proteins in the darkred module, or that co-expression of these genes was due to their functional similarity and that their proteins may not interact.

The steelblue module significantly correlated with chicken line and weakly correlated with challenge status. Although no immune-related GO terms were significant, network analysis revealed immune-related genes in this module, including *JAK*, *TRAF3*, *DICER1*, and *MAPK1*. This network was also associated with the endocytosis pathway, which is a mechanism by which NDV enters the host cell. It may be possible for a host to modify gene expression of this pathway to minimize viral entry. These inbred chicken lines have been used as models in multiple studies; further exploration of the lightyellow module will help to identify differences between the two lines.

The combined analysis of three target tissues generated a more extensive understanding of response to infection with lentogenic NDV at the whole organism level. Other studies have compared differences between the expression of key immune genes in these genetic lines after an NDV challenge in embryo (Schilling et al., 2018; Schilling et al., 2019) and at 3 weeks of age under heat stress conditions (Saelao et al., 2018). Also, genetic markers of interest have been identified in a commercial line (Rowland et al., 2018; Saelao et al., 2019) and African ecotypes (Walugembe et al., 2019) after infection with this same lentogenic strain. Data are being generated at a fast rate; new tools and types of analyses are needed to analyze the data in a more comprehensive way. The current study attempted to identify critical genes associated with response to NDV based on RNA-seq data from three different tissues. This new bioinformatics approach was limited by number of tissues and number of individuals per tissue. Further studies to confirm these findings are necessary. Differences between the lines demonstrated how host genetics impacts vaccine response. The darkred and steelblue modules were the modules most likely to contain genes that relate to pathogen response, and *EIF2AK2*, *MPEG1*, and *TNFSF13B* are promising candidate genes for lentogenic NDV resistance. These were not identified as candidate genes in the individual tissue analyses. Future studies are required to determine how the expression of these genes is impacted by velogenic challenge in the Fayoumi and Leghorn. This study demonstrated how WGCNA could be used to efficiently combine transcriptome data from different tissues, genetic lines, challenge statuses, and dpi

to identify important gene clusters associated with factors of interest. Novel clusters and driver genes were identified in this analysis that can be used as a reference for comparison with future studies that use different NDV strains, as well as serving as a basis for investigations of host modulation of vaccine efficacy.

MATERIALS AND METHODS

This study used data from experiments approved by the Iowa State University Institutional Animal Care and Use Committee (IACUC log number 1-13-7490-G). All experiments were performed in accordance with the committee's relevant guidelines and regulations. The experimental design for the lentogenic NDV challenge has been previously described in detail (Deist et al., 2017b). Briefly, 3-week-old chicks were either inoculated with 10^7 EID$_{50}$ La Sota NDV (challenged) or given PBS (non-challenged) via an ocular–nasal route. The relatively resistant Fayoumi and relatively susceptible Leghorn were intermixed within the challenged and non-challenged groups. Blocking for challenge status and line, at each time of 2, 6, and 10 dpi, one-third of the birds were euthanized for tissue collection. This design resulted in a total of 12 treatment groups: two challenge groups (challenged and non-challenged), two lines (Fayoumi and Leghorn), and three time points (2, 6, and 10 dpi).

Using the RNAqueous kit Total RNA Isolation Kit (Thermo Fisher Scientific, Waltham, MA, USA), RNA was isolated from the trachea epithelial cells, lungs, and Harderian glands of about four birds per treatment group. The RNA was treated with DNAse and the quality of each sample was confirmed (RQN > 8). The Illumina TruSeq RNA Library Prep Kit v2 (Illumina, San Diego, CA, USA) was used to generate cDNA libraries. Libraries were sequenced on the HiSeq2500 for 100-bp, single-end reads at the Iowa State University DNA Facility (Ames, IA, USA). The sequencing data was analyzed with a standard pipeline and the *Gallus-gallus* 5 reference genome (Gal5; GCA_000002315.3): FASTX, TopHat2 (Kim et al., 2013), and HTSeq (Anders et al., 2015). Transcripts with less than four counts across all samples were removed prior to normalization. The RNA-seq data are publicly available from ArrayExpress at EMBL-EBI under accession numbers: E-MTAB-5431, E-MTAB-5859, and E-MTAB-6038.

For data visualization of the 143 sequenced samples, pcaExplorer (Marini, 2016) was used to generate PCA plots based on count data normalized using DESeq2 (Love et al., 2014), accounting for tissue and treatment group. The variance associated with each principal component was calculated with the 1,000 most variant transcripts (Marini, 2016).

Differential expression analysis was performed to determine how the tissues responded differently to the NDV challenge using the GLM in edgeR (Robinson et al., 2010), accounting for every combination of tissue, line, dpi, and challenge status resulting in 36 levels. Within each line and dpi, the following contrasts were written to identify transcripts that were significantly impacted (FDR < 0.05) by the interaction between challenge and tissue: (Tissue i, Line j, dpi k, Challenged–Tissue i, Line j, dpi k, Non-challenged)–(Tissue i', Line j, dpi k, Challenged–Tissue i', Line j, dpi k, Non-challenged). With three tissues, two lines, and three time points, a total of 18 contrasts were analyzed.

Co-expression analysis was performed using WGCNA (Langfelder and Horvath, 2008). WGCNA clusters genes into modules based on expression levels and correlates a module's eigengene to traits of interest. All transcripts with more than four counts across all samples were included. Transcript counts were normalized using the variance stabilizing transformation by DESeq2 (Love et al., 2014). The soft power threshold was set to 20 and a minimum module size of 30 was used (Langfelder and Horvath, 2008). The module's eigengenes were correlated to factors of interest. Coding for the factors was as follows: Line [Fayoumi (1), Leghorn (0)], days post-infection [dpi; 2 dpi (0), 6 dpi (1), 10 dpi (2)], NDV [challenged (1), non-challenged (0)], sex [male (1), female (0)], Trachea [Trachea (1), Lung (0), Harderian gland (0)], Lung [Trachea (0), Lung (1), Harderian gland (0)], and Harderian gland [Trachea (0), Lung (0), Harderian gland (1)]. For each factor, the driver genes for a module were based on the transcripts with the highest absolute gene significance and module membership calculated by WGCNA (Langfelder and Horvath, 2008). Gene significance was estimated by correlating a transcript's expression profile with a sample trait. A correlation of a module eigengene and the expression profile of each transcript was used to determine a transcript's module membership score. All reported *p* values were taken directly from the WGCNA output.

GO term and network analyses were performed to ascertain the predicted function of specific modules. All transcript IDs were converted to their associated gene name using Ensembl BioMart, then NCBI and Uniprot were used to find gene names for those transcripts unidentified by Ensembl. Panther (Mi et al., 2017) was utilized for GO term overrepresentation analysis using a Bonferroni correction. Only the top-level GO terms, based on the hierarchical labeling by Panther, were presented (**Figure 3**). Associated gene names were used as input into STRING (Szklarczyk et al., 2017) for network analysis based on protein–protein interactions described in the literature. Disconnected nodes were removed and a high confidence (0.700) was required for connection of the nodes. The nodes were colored based on MCL clustering (inflation = 3) (Szklarczyk et al., 2017).

DATA AVAILABILITY STATEMENT

The RNA-seq data are publicly available from ArrayExpress at EMBL-EBI under accession numbers: E-MTAB-5431, E-MTAB-5859, and E-MTAB-6038.

ETHICS STATEMENT

This study used data from experiments approved by the Iowa State University Institutional Animal Care and Use Committee (IACUC log number 1-13-7490-G). All experiments were performed in accordance with the committee's relevant guidelines and regulations.

AUTHOR CONTRIBUTIONS

MD: Conceptualization of RNA-seq experiment, investigation, collected samples, isolated RNA, constructed cDNA libraries, formal analysis of RNA-seq, methodology, data visualization, writing—original draft preparation. RG: Conceptualization of experiment, experimental design, investigation—prepared viral isolate for inoculation, writing—review and editing. JD: Conceptualization of experiment, experimental design, writing—review and editing. HZ: Conceptualization of experiment, experimental design, funding acquisition, project administration, writing—review and editing. SL: Conceptualization of experiment, experimental design, provided resources (chicken genetic lines), investigation, funding acquisition, project administration, supervised analysis of data, writing—review and editing. All authors have read and approved the manuscript.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.00011/full#supplementary-material

SUPPLEMENTARY TABLE S1 | The S1 Table (table 1.xlsx) includes the Ensembl transcript ID and corresponding log fold change and FDR value for all transcripts for all 18 challenge*line contrasts.

REFERENCES

Afonso, C. L., Miller, P. J., Grund, C., Koch, G., Peeters, B. P., Selleck, P. W., et al. (2012). *Manual of diagnostic tests and vaccines for terrestrial animals* (Paris, France: World Organisation for Animal Health).

Anders, S., Pyl, P. T., and Huber, W. (2015). HTSeq–a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31 (2), 166–169. doi: 10.1093/bioinformatics/btu638

Cheeseman, J. H., Kaiser, M. G., Ciraci, C., Kaiser, P., and Lamont, S. J. (2007). Breed effect on early cytokine mRNA expression in spleen and cecum of chickens with and without Salmonella enteritidis infection. *Dev. Comp. Immunol.* 31 (1), 52–60. doi: 10.1016/j.dci.2006.04.001

Deb, A., Haque, S. J., Mogensen, T., Silverman, R. H., and Williams, B. R. G. (2001). RNA-dependent protein kinase PKR is required for activation of NFKB by IFN-γ in a STAT1-independent pathway. *J. Immunol.* 166 (10), 6170–6180. doi: 10.4049/jimmunol.166.10.6170

Deist, M. S., Gallardo, R. A., Bunn, D. A., Dekkers, J. C. M., Zhou, H., and Lamont, S. J. (2017a). Resistant and susceptible chicken lines show distinctive responses to Newcastle disease virus infection in the lung transcriptome. *BMC Genomics* 18 (1), 989. doi: 10.1186/s12864-017-4380-4

Deist, M. S., Gallardo, R. A., Bunn, D. A., Kelly, T. R., Dekkers, J. C. M., Zhou, H., et al. (2017b). Novel mechanisms revealed in the trachea transcriptome of resistant and susceptible chicken lines following infection with Newcastle disease virus. *Clin. Vaccine Immunol.* 24 (5), 1–17. doi: 10.1128/CVI.00027-17

Deist, M. S., Gallardo, R. A., Bunn, D. A., Kelly, T. R., Dekkers, J. C. M., Zhou, H., et al. (2018). Novel analysis of the Harderian gland transcriptome response to Newcastle disease virus in two inbred chicken lines. *Sci. Rep.* 8 (1), 6558. doi: 10.1038/s41598-018-24830-0

Dimitrov, K. M., Afonso, C. L., Yu, Q., and Miller, P. J. (2017). Newcastle disease vaccines-A solved problem or a continuous challenge? *Vet. Microbiol.* 206, 126–136. doi: 10.1016/j.vetmic.2016.12.019

Fleming, D. S., Koltes, J. E., Fritz-Waters, E. R., Rothschild, M. F., Schmidt, C. J., Ashwell, C. M., et al. (2016). Single nucleotide variant discovery of highly inbred Leghorn and Fayoumi chicken breeds using pooled whole genome resequencing data reveals insights into phenotype differences. *BMC Genomics* 17 (1), 812. doi: 10.1186/s12864-016-3147-7

Kapczynski, D. R., Afonso, C. L., and Miller, P. J. (2013). Immune responses of poultry to Newcastle disease virus. *Dev. Comp. Immunol.* 41 (3), 447–453. doi: 10.1016/j.dci.2013.04.012

Kim, D., Pertea, G., Trapnel, C., Pimental, H., Kelley, R., and S. L, S. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* 14, 13. Kim et al., 2013 doi: 10.1186/gb-2013-14-4-r36

Kim, T. H., and Zhou, H. (2015). Functional analysis of chicken IRF7 in response to dsRNA analog poly(I:C) by integrating overexpression and knockdown. *PloS One* 10 (7), e0133450. doi: 10.1371/journal.pone.0133450

Lakshmanan, N., Kaiser, M. G., and Lamont, S. J. (1996). "Marek's disease resistance in MHC-congenic lines from Leghorn and Fayoumi breeds," in *Current Research on Marek's Disease.* Eds. R. F. Silva, H. H. Cheng, P. M. Coussens, L. F. Lee and L. F. Velicer (Kennett Square, PA: American Association of Avian Pathologists), 57–62.

Lamont, S. J., and Chen, Y. (1992). Endogenous viral genes in thirteen highly inbred chicken lines and lines selected for immune response traits. *Poult. Sci.* 71, 530–538. doi: 10.3382/ps.0710530

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinf.* 9, 559. doi: 10.1186/1471-2105-9-559

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15 (12), 550. doi: 10.1186/s13059-014-0550-8

Marini, F. (2016). "pcaExplorer: interactive visualization of RNA-seq data using a principal components approach.")

McCormack, R., and Podack, E. R. (2015). Perforin-2/Mpeg1 and other pore-forming proteins throughout evolution. *J. Leukocyte Biol.* 98 (5), 761–768. doi: 10.1189/jlb.4MR1114-523RR

Mi, H., Huang, X., Muruganujan, A., Tang, H., Mills, C., Kang, D., et al. (2017). PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Res.* 45 (D1), D183–D189. doi: 10.1093/nar/gkw1138

Miller, P. J., and Koch, G. (2013). "Newcastle disease, other avian paramyxoviruses, and avian metapneumovirus infections," in *Diseases of Poultry, 13th ed.* Eds. D. E. Swayne, J. R. Glisson, L. R. McDougald, L. K. Nolan, D. L. Suarez and V. Nair. (Ames, IA. :John Wiley & Sons, Ltd), 89–138.

Nagai, Y., Klenk, H., and Rott, R. (1976). Proteolytic cleavage of the viral glycoproteins and its significance for the virulence of Newcastle disease virus. *Virology* 72, 494–508. doi: 10.1016/0042-6822(76)90178-1

Nagai, Y. (1995). Virus activation by host proteinases. a pivotal role in the spread of infection, tissue tropism and pathogenicity. *Microbiol. Immunol.* 39 (1), 1–9. doi: 10.1111/j.1348-0421.1995.tb02161.x

Pinard-van der Laan, M. H., Bed'hom, B., Coville, J. L., Pitel, F., Feve, K., Leroux, S., et al. (2009). Microsatellite mapping of QTLs affecting resistance to coccidiosis (Eimeria tenella) in a Fayoumi x White Leghorn cross. *BMC Genomics* 10, 13. doi: 10.1186/1471-2164-10-31

Reynolds, D. L., and Maraqa, A. D. (2000). Protective immunity against Newcastle disease: the role of cell-mediated immunity. *Avian Dis.* 44 (1), 145–154. doi: 10.2307/1592518

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26 (1), 139–140. doi: 10.1093/bioinformatics/btp616

Rowland, K., Saelao, P., Wang, Y., Fulton, J. E., Liebe, G. N., McCarron, A. M., et al. (2018). Association of candidate genes with response to heat and newcastle disease virus. *Genes (Basel)* 9 (11), 1–14. doi: 10.3390/genes9110560

Sadler, A. J., and Williams, B. R. (2008). Interferon-inducible antiviral effectors. *Nat. Rev. Immunol.* 8 (7), 559–568. doi: 10.1038/nri2314

Saelao, P., Wang, Y., Chanthavixay, G., Yu, V., Gallardo, R. A., Dekkers, J. C. M., et al. (2018). Integrated proteomic and transcriptomic analysis of differential expression of chicken lung tissue in response to NDV infection during heat stress. *Genes (Basel)* 9 (12), 1–16. doi: 10.3390/genes9120579

Saelao, P., Wang, Y., Chanthavixay, G., Gallardo, R. A., Wolc, A., Dekkers, J. C. M., et al. (2019). Genetics and genomic regions affecting response to newcastle disease virus infection under heat stress in layer chickens. *Genes (Basel)* 10 (1), 1–17. doi: 10.3390/genes10010061

Sandford, E. E., Orr, M., Li, X., Zhou, H., Johnson, T. J., Kariyawasam, S., et al. (2012). Strong concordance between transcriptomic patterns of spleen and peripheral blood leukocytes in response to avian pathogenic Escherichia coli infection. *Avian Dis.* 56 (4), 732–736. doi: 10.1637/10261-060512-Reg.1

Schilling, M. A., Katani, R., Memari, S., Cavanaugh, M., Buza, J., Radzio-Basu, J., et al. (2018). Transcriptional innate immune response of the developing chicken embryo to newcastle disease virus infection. *Front. Genet.* 9, 61. doi: 10.3389/fgene.2018.00061

Schilling, M. A., Memari, S., Cavanaugh, M., Katani, R., Deist, M. S., Radzio-Basu, J., et al. (2019). Conserved, breed-dependent, and subline-dependent innate immune responses of Fayoumi and Leghorn chicken

embryos to Newcastle disease virus infection. *Sci. Rep.* 9 (1), 7209. doi: 10.1038/s41598-019-43483-1

Schneider, K. (2004). Chicken BAFF–a highly conserved cytokine that mediates B cell survival. *Int. Immunol.* 16 (1), 139–148. doi: 10.1093/intimm/dxh015

Sun, H., Bi, R., Liu, P., Nolan, L. K., and Lamont, S. J. (2016). Combined analysis of primary lymphoid tissues' transcriptomic response to extra-intestinal Escherichia coli (ExPEC) infection. *Dev. Comp. Immunol.* 57, 99–106. doi: 10.1016/j.dci.2015.12.013

Szklarczyk, D., Morris, J. H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., et al. (2017). The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res.* 45 (D1), D362–D368. doi: 10.1093/nar/gkw937

Walugembe, M., Mushi, J. R., Amuzu-Aweh, E. N., Chiwanga, G. H., Msoffe, P. L., Wang, Y., et al. (2019). Genetic analyses of tanzanian local chicken ecotypes challenged with newcastle disease virus. *Genes* 10 (7), 1–16. doi: 10.3390/genes10070546

Wang, Y., Lupiani, B., Reddy, S. M., Lamont, S. J., and Zhou, H. (2014). RNA-seq analysis revealed novel genes and signaling pathway associated with disease resistance to avian influenza virus infection in chickens. *Poult. Sci.* 93 (2), 485–493. doi: 10.3382/ps.2013-03557

Zhang, S., Sun, Y., Chen, H., Dai, Y., Zhan, Y., Yu, S., et al. (2014). Activation of the PKR/eIF2α signaling cascade inhibits replication of Newcastle disease virus. *Virol. J.* 11 (62), 1–11. doi: 10.1186/1743-422X-11-62

Zhang, J., Kaiser, M. G., Deist, M. S., Gallardo, R. A., Bunn, D. A., Kelly, T. R., et al. (2018). Transcriptome analysis in spleen reveals differential regulation of response to newcastle disease virus in two chicken lines. *Sci. Rep.* 8 (1), 1278. doi: 10.1038/s41598-018-19754-8