



OPEN

## A Bayesian network structure learning approach to identify genes associated with stress in spleens of chickens

E. A. Videla Rodriguez<sup>1</sup>, John B. O. Mitchell<sup>2</sup> & V. Anne Smith<sup>1</sup>✉

Differences in the expression patterns of genes have been used to measure the effects of non-stress or stress conditions in poultry species. However, the list of genes identified can be extensive and they might be related to several biological systems. Therefore, the aim of this study was to identify a small set of genes closely associated with stress in a poultry animal model, the chicken (*Gallus gallus*), by reusing and combining data previously published together with bioinformatic analysis and Bayesian networks in a multi-step approach. Two datasets were collected from publicly available repositories and pre-processed. Bioinformatics analyses were performed to identify genes common to both datasets that showed differential expression patterns between non-stress and stress conditions. Bayesian networks were learnt using a Simulated Annealing algorithm implemented in the software Banjo. The structure of the Bayesian network consisted of 16 out of 19 genes together with the stress condition. Network structure showed CARD19 directly connected to the stress condition plus highlighted CYGB, BRAT1, and EPN3 as relevant, suggesting these genes could play a role in stress. The biological functionality of these genes is related to damage, apoptosis, and oxygen provision, and they could potentially be further explored as biomarkers of stress.

Stress has been defined as the physiological response triggered by an external or internal stimulus, with the aim of coping with or dealing with the stimulus<sup>1,2</sup>. The stress response is perceived, integrated, and displayed in the interplay of three major systems, the nervous system, the endocrine system, and the immune system, also known as the immune-neuroendocrine (INE) interplay<sup>3–5</sup>. Two axes of the INE interplay play the major roles in the stress response during the initial exposure to a stimulus: the Adrenergic Nervous System and the Hypothalamus-Pituitary-Adrenergic Axis are in charge of the fight or flight response and the activation of metabolic processes such as gluconeogenesis or proteolysis, respectively<sup>6–9</sup>. The activation of these two axes translates into a series of general and specific physiological as well as behavioral adaptations to deal with the influence of the potentially stressful stimulus<sup>3,10,11</sup>.

The stress response consists of a general as well as a specific response to a potentially stressful stimulus. While the general response is independent of the nature of the stimulus, the specific response is mostly dependent on the type of stimulus<sup>3,10,12,13</sup>. For example, a bird facing a potential predator might either display an escape behaviour, running as fast as possible from the predator, or display freezing behaviour, pretending to be dead by showing a tonic immobility response to avoid the threat. As another example, a bird under heat stress conditions can display behaviours such as panting or wing spreading to dissipate as much excess heat as possible<sup>12,14</sup>. In terms of the physiology, the exposure of birds to chronic heat stress can affect many systems, such as the digestive tract, the gut microbiota, the immune system, and the antioxidant system among others<sup>15–18</sup>. In the complexity of the stress response, and in particular heat stress, some of the effects of the stress response can have indirect consequences on other systems. To some extent, poor productive performance in poultry (e.g., poor body weight gain or egg production) can be associated with a lower feed intake as a consequence of the exposure to heat stress, or the alteration of the structure of immune organs can be associated with the imbalance of the oxidant/anti-oxidant status as a result of oxidative stress<sup>3,18–20</sup>.

Considering the complexity of the stress phenomenon, many indicators have been used to study the stress response. Corticosterone is considered as the main indicator of stress, as increased levels have been shown

<sup>1</sup>School of Biology, University of St Andrews, St Andrews KY16 9TH, Fife, UK. <sup>2</sup>EaStCHEM School of Chemistry and Biomedical Sciences Research Complex, University of St Andrews, St Andrews KY16 9ST, Fife, UK. ✉email: vas1@st-andrews.ac.uk

by birds exposed to stress protocols<sup>21–23</sup>. However, other indicators have been studied, such as changes in the microbial communities of the gut, immune suppression, oxidative status, and the morphology of the digestive tract among others<sup>17,20,21,24,25</sup>. Changes in the expression patterns of genes driven by the exposure to a potentially stressful stimulus have also been studied<sup>26,27</sup>. This indicator combines high-throughput technologies to measure the gene expression with bioinformatic tools to identify statistically significant differences between birds raised under non-stress conditions and birds exposed to the stressful stimulus<sup>28–31</sup>. Even though the set of significant genes can be used to identify molecular pathways associated with stress, one of the main drawbacks of this approach is that the number of statistically significant genes can be quite extensive and they can be related to several mechanisms and systems within the physiology of the birds<sup>29,32,33</sup>.

We can solve this problem with further exploration to discover hidden patterns behind the data: key genes among these lists that are most relevant to the stress condition of interest can be identified using Bayesian networks<sup>34–36</sup>. Bayesian networks represent a powerful mathematical tool that provide new insights by identifying probabilistic relationships between a given set of variables, distinguishing direct from indirect influences<sup>35,37–39</sup>. A Bayesian network is a directed acyclic graph (DAG), consisting of a set of nodes, that represent each of the variables, and edges, that display the relationships between the nodes<sup>40,41</sup>. Because the network is acyclic, i.e., has no loops, relationships among variables within a Bayesian network are often described using a ‘family analogy’ where a node of interest can have parents and children, considering those nodes with edges coming into or going out of the node of interest, respectively. Additionally, spouses are those nodes that share a common child with the node of interest<sup>40,42</sup>. The set of parents, children, and spouses represents one of the main properties of Bayesian networks known as a Markov Blanket. Considering that Bayesian networks are based on probability theory, this property allows identification of a small set of nodes in close association with a node of interest that makes this node probabilistically independent from the rest of the network<sup>40,42</sup>.

In this complex scenario, the aim of this study was to identify a small set of genes in close association with the stress condition in a poultry animal model, the chicken (*Gallus gallus*). We further increased the statistical power of our method by reusing and combining data from separate experiments combined with the accuracy of bioinformatic analysis and the power of Bayesian networks in a multi-step interdisciplinary approach. Publicly available repositories were explored to collect data coming from two experiments where the expression values of genes were measured in the spleen of chickens exposed to heat stress by RNA-sequencing. Bioinformatic analyses were implemented to identify a set of genes driven by the exposure to the stressor, followed by learning the structure of a Bayesian network to display the relationships and interactions between the genes and the stressful condition. The structure of the network was divided into communities of densely connected nodes, with a special focus on the community of nodes related to the stress condition. Finally, the biological meaning of the discovered interactions and relationships was explored.

## Results

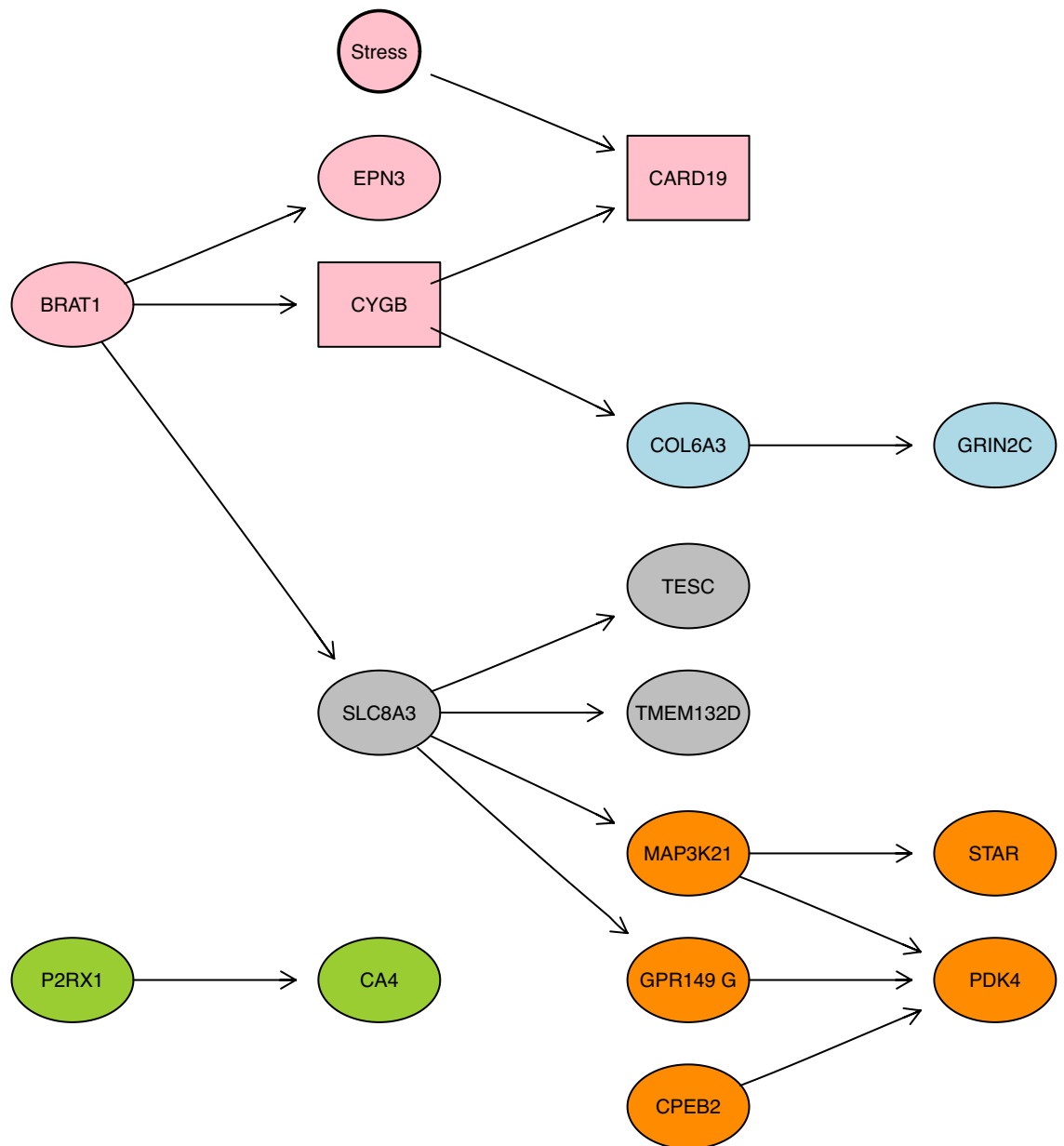
A total of 19 genes having differences in the expression patterns between non-stress and stress chickens were common to two datasets evaluating the effects of stress in the spleen of chickens. The overall structure of the network revealed that 16 out of 19 genes were part of the network in addition to the stress condition (Fig. 1). The stress condition was directly connected in the network with only one gene, CARD19. The Markov Blanket property of the condition revealed that in addition to CARD19, CYGB was also related to the stress condition (Fig. 1, rectangle-shaped nodes).

The structure of the Bayesian networks was further explored by dividing the overall structure into smaller communities of densely connected nodes within the community but scarcely connected with nodes in other communities<sup>43</sup>. The application of a divisive cluster algorithm that uses the structure of the Bayesian networks as the input revealed five communities of densely connected nodes. The community of the stress condition consisted of four genes: in addition to the genes belonging to the Markov Blanket, BRAT1 and EPN3 displayed a possible interaction with the stress condition, representing a group of genes densely connected amongst themselves, but scarcely connected with the rest of the genes (Fig. 1, nodes highlighted in pink).

The Database for Annotation, Visualization, and Integrated Discovery (DAVID) was applied to explore the biological functionality of these four genes. DAVID overrepresentation analysis identified three terms, although all were not significant after adjustments for multiple tests: Calcium signaling pathway (KEGG Pathway, P-value = 0.026, Benjamini adjusted P-value = 0.28), sarcolemma (GO term, P-value = 0.045, Benjamini adjusted P-value = 1), and membrane (keyword, P-value = 0.06, Benjamini adjusted P-value = 1). The DAVID Functional Annotation Table (Table 1) shows the KEGG pathways<sup>44</sup> and GO TERMS<sup>45</sup> for each one of the genes being part of the community analysis of the stress condition. Terms particularly relevant to the stress condition are related to regulation of apoptotic process and caspase recruitment (CARD19), oxygen transporter activity and oxygen binding (CYGB), apoptosis process and cellular response to DNA damage stimulus (BRAT1), and endocytosis (EPN3).

## Discussion

This study was aimed at identifying a reduced number of genes closely associated with a stressful condition in the chicken as a poultry animal model. To have a more accurate approximation to the stress phenomenon, two publicly available datasets involving the measurement of gene expression in the spleen of chicken exposed to heat stress were combined into a larger dataset. After bioinformatic pre-processing and analysis, we identified a set of 19 genes common to both datasets with a differential expression pattern; these genes were used for learning the structure of the Bayesian networks. With the Bayesian network in place, its structure was divided into smaller communities of densely connected nodes. By the implementation of this approach, two genes were identified as part of the Markov Blanket property of the stress condition. In addition to these two genes, two other genes were



**Figure 1.** Bayesian network and community analysis of a set of genes. 19 genes showing differences in expression pattern were initially included at the time of learning the structure of the network in addition to the stressful condition; however, only 16 out of those 19 were linked in a network structure. Nodes represent each one of the genes and the stressful condition (circle-shaped node, thick outline), the edges represent probabilistic dependencies between the nodes. Note the direction of the arrows do not represent causation, but instead a statistical relationship. The Markov Blanket of the stress condition (rectangle-shaped nodes) consisted of two genes, CARD19 (child) and CYGB (spouse). Five communities of densely connected nodes were identified (different colours represent different communities). The community of the condition consisted of 4 genes (CARD19, EPN3, CYGB, and BRAT1, highlighted in pink).

part of the community of the condition, giving a total of 4 out of the 19 initial genes displaying a close relationship with the stress condition. Our results identified a small set of relevant genes related to stress that can be used to extract meaningful information regarding the genetics of this complex phenomenon.

Stress involves the perception of the stimulus in the immune-neuroendocrine interplay, triggering the stress response, and displaying physiological and behavioural adaptations with the aim of dealing with the stressful stimulus<sup>2–4,6</sup>. Heat stress has been widely studied and its effects on immune organs and immune responses have been reported<sup>13,15,18,46</sup>. In particular, Hirakawa et al.<sup>47</sup> found that the mass of the spleen was severely affected by the exposure to heat stress<sup>47</sup>. Additionally, heat stress altered the structure of the spleen, having an impact on the humoral immune responses that modulate the lymphocyte populations<sup>47</sup>. Chickens under high environmental temperatures have also shown imbalances in the oxidant/anti-oxidant status as a consequence of the alteration of some by-products or end products of lipid peroxidation such as malondialdehyde (MDA) and thiobarbituric

<b>Caspase recruitment domain family member 19 (CARD19)</b>	
GO TERMS	<b>Regulation of apoptotic process</b> Integral component of membrane
<b>Cytoglobin (CYGB)</b>	
GO TERMS	Neuron projection, neuronal cell body <b>Oxygen transporter activity</b> , iron ion binding, <b>oxygen binding</b> , heme binding
<b>BRCA1 associated ATM activator 1 (BRAT1)</b>	
GO TERMS	Positive regulation of protein phosphorylation, glucose metabolic process, apoptotic process, <b>cellular response to DNA damage stimulus</b> , cell proliferation, response to ionizing radiation, cell growth, cell migration, mitochondrion localization Nucleus, cytoplasm, membrane
<b>Epsin 3 (EPN3)</b>	
KEGG PATHWAY	<b>Endocytosis</b>

**Table 1.** Functional Annotation Table provided by the Database for Annotation, Visualization, and Integrated Discovery (DAVID) corresponding to the four genes found to be in close relationship with the stressful condition. In bold are highlighted terms particularly relevant to the stress condition.

acid reacting substances (TBARS)<sup>20,48</sup>. The imbalance is created by the excess of oxidant molecules, such as reactive species containing oxygen, nitrogen, and/or chlorine, potentially affecting the structure of proteins, lipids, and DNA and RNA. Consequently, the functioning of the cell might be affected in terms of energy availability, calcium homeostasis, and mitochondrial functionality, leading to cell damage, and therefore to the survival of the cell being threatened by apoptosis or necrosis<sup>19,49</sup>.

The Markov Blanket property together with the community analysis revealed a total of four genes in close association with the stress condition. One of the genes, CARD19, showed a direct interaction with the stress condition, while the other three genes were part of the Markov Blanket and/or the community of the stress condition. CARD proteins belong to the family of caspase recruitment domains and they are proteins that mediate apoptosis as well as the activation of the NF- $\kappa$ B signaling pathways<sup>50–53</sup>. Cytoglobin (CYGB) belongs to the globin family, whose major role is related to the provision of oxygen in different tissues and organs, in addition to a potential protective activity against reactive oxygen species<sup>54–56</sup>. Considering that heat stress leads to oxidative stress, cell damage, apoptosis, and immune dysfunction, CARD19 and CYGB could be identified as key genes associated with these mechanisms that chickens trigger as a consequence of the influence of the stressor. Under exposure to other stressors, such as an immune challenge with mycotoxin or hypoxic conditions, apoptotic signaling pathways were also activated in splenic cell<sup>49,57</sup>. Specially under hypoxic conditions, Chen et al.<sup>49</sup> identified that splenic cells initiated apoptotic signaling pathways as a result of oxidative stress involving inflammatory mechanisms and the NF- $\kappa$ B pathway<sup>49</sup>.

Considering the learnt structure of the Bayesian network, the further analysis of smaller communities of densely connected nodes showed that the stress condition potentially interacted with two other genes: BRAT1 and EPN3. BRCA1-associated ATM activator 1 (BRAT1) was previously identified by Qui et al.<sup>58</sup> in the spleen of layer chickens undergoing an infection with avian leukosis virus (subgroup J)<sup>58</sup>. In humans, this gene interacts with two other genes, BRCA1 and ATM, mediating cell pathways associated with DNA damage as well as apoptosis<sup>59–62</sup>. Epsin-3 (EPN3) is a member of the endocytosis protein adapter gene family, and its main function is related to endocytosis<sup>63</sup>. Additionally, EPN3 has been identified in pathological or damaged tissues requiring wound healing<sup>64</sup>. In our study, it seems that both the Markov Blanket and the community of densely connected nodes of the stress condition are pointing towards key genes related to apoptosis and tissue damage. It is then plausible to highlight that when chickens are exposed to a complex phenomenon such as heat stress, one of the main immune organs, the spleen, reflects some morphological and physiological alterations as a consequence of undergoing apoptotic-related mechanisms, potentially translating into the reported suppression and dysfunction of the immune responses<sup>32,47</sup>.

Regarding the network approach implemented in our study, the combination of two strategies was applied with the aim of identifying genes in close relationship with the stress condition: the Markov Blanket property of Bayesian networks and the community of highly connected nodes<sup>42,43</sup>. Initially, the overall structure of the network, the relationships and the interaction between the given set of variables (the genes and the stressful condition) were learnt from the data. As a following step, the already learnt structure of the Bayesian network was divided into smaller groups of densely connected nodes<sup>43</sup>. By combining these two strategies, an initial set of 19 genes were further reduced to a small set of genes that showed a close association with the stress and that can be further studied. Among this small set of genes, the structure of the Bayesian network revealed that CARD19 showed a close interaction with the stressful condition, suggesting this gene could be explored as a potential biomarker of stress. Therefore, further research can be developed with short-term goals, such as using these genes to identify chickens raised under non-stress or stress conditions, and consequently, using them as indicators of stress, raising the alarm to monitor and manage the breeding conditions to mitigate the detrimental effects of stress on poultry production<sup>65,66</sup>. On the other hand, further research can have long-term goals, such as artificial selection and breeding programs in order to enhance the resilience or resistance of chicken breeds to stress, such as the Fayoumi chickens that have been used in studies as a heat stress and disease resistant breed<sup>66–68</sup>.

In conclusion, this study implemented a series of steps aimed at reducing an initial number of genes obtained from high-throughput technologies to a small number of genes, and unravelling their relationships and interactions. We combined two previous studies that evaluated the effects of stress on the spleen of chickens to get a more accurate approximation to the stress phenomenon. The series of steps involved the combination of: (i) bioinformatic tools to identify differentially expressed genes, (ii) Bayesian networks to learn the overall structure of the network, (iii) the Markov Blanket together with the community analysis to identify a small set of genes in close association with the stress condition, and (iv) the database for biological knowledge discovery DAVID. Such a sequence of computational approaches could be applicable to many studies of gene expression, across many measurement platforms, enabling combination of power from multiple experiments to identify small sets of genes for further study. This work is complementary to the original studies, in that the genes we identify are not an overall picture (this would be found in those studies), but instead provide information about a small set of genes with a strong signal, across multiple studies, suggesting relation to the condition of interest. Here, the outcome of this series of steps identified two genes as being part of the Markov Blanket and two additional genes as being part of the community analysis for the stress condition in poultry. The biological processes of these four genes were related to damage and apoptosis, and they could potentially be further used as biomarkers of heat stress. The exploratory nature of our study demands future research to determine the discovered protein–protein interactions *in-vivo*, comparing the differences between chickens raised under non-stress conditions and chickens raised under stress conditions.

## Methods

**Dataset.** Two datasets were collected from a publicly available data repository (Gene Expression Omnibus—GEO), under the following accession numbers: GSE119387 and GSE85434. Briefly, chickens studied in the GSE119387 dataset came from two different regions of Ethiopia: low altitude regions are hot and humid, with chickens adapted to heat conditions, and high altitude regions, on the other hand, are cooler, with chickens susceptible to heat conditions. The effects of heat stress conditions were evaluated in chickens coming from both regions, but they were raised in low altitude regions (hot and humid). Chickens studied in GSE85434 were exposed either to thermoneutral condition (25 °C—control) or to thermal treatment (35 °C—heat stress) for 3.5 h. Therefore, both datasets evaluated the effects of heat stress on gene expression in the spleen of chickens, measured by RNA-sequencing technologies. Each dataset was individually analyzed with the aim of identifying genes relevant for stress, determined by differential expression patterns between non-stress and stress conditions. The .txt files were downloaded, imported into R<sup>69</sup>, and pre-processed using the R package “edgeR”<sup>70</sup>, normalizing and removing any possible background noise associated with the data. Thereafter, the *lmfit* function, from the R package “limma”<sup>71</sup>, was implemented to fit a linear model according to the experimental design of each dataset. The *eBayes* function was applied to calculate the statistics that would identify the set of genes. Finally, the top highly significant genes were selected using the *topTable* function. A list of 677 and 483 relevant genes were independently identified for each dataset. To combine the power of the two studies, we examined those genes in common between the two datasets, giving a total of 19 genes shared by the two studies, and representing a manageable number of variables to learn the structure of Bayesian networks. The expression values for each gene were extracted from the datasets and were used to create the final dataset, which consisted of 19 genes and 50 individuals. This dataset was discretized into three categories, low, medium or high, applying the function *discretizeDF*, within the “arules” package<sup>72</sup>, following a quantile discretization distribution. The discretization of the data was implemented to reduce noise possibly related to differences between experimental designs as well as to increase the statistical power<sup>35,40</sup>. Once the data had been discretized, the stress condition was included as a binary variable, taking a value of 0 for chickens raised under non-stress conditions and a value of 1 for chickens exposed to the stress condition.

**Bayesian networks and community analysis.** To learn the structure of Bayesian networks, the software Banjo was implemented (available for free for academic purposes from <http://www.cs.duke.edu/~amink/software/banjo/>)<sup>35,37</sup>. Banjo implements heuristic searches with two possible algorithms, Greedy or Simulated Annealing, and scores each network with a BDe score. Banjo allows the possibility of selecting the top highest scoring network or combining the *n* top highest-scoring networks into one consensus network as the output. In this study, the search space was explored with a total of 250 million networks, using a Simulated Annealing algorithm with local random moves as the proposer. A consensus network was built combining the top 100 high-scoring networks. Considering that Bayesian networks implement heuristic searches, adding, removing, or reversing edges with the aim of finding the highest scoring network, the final set of edges was slightly different after running the algorithm several times. Therefore, with the aim of finding the Bayesian network that best fitted the data, the following strategy was implemented to solve this challenge. The search space was explored a total of 1000 times, resulting in 1000 consensus networks. These 1000 networks were divided into 10 groups of a hundred networks (10 × 100 = 1000). Within each set of a hundred networks, all the arcs identified among these 100 networks were used to create a matrix of presence/absence. Thereafter, for each individual network, if an arc was present in the set of arcs, a value of one was assigned; while, if the arcs was absent, a value of zero was assigned. Ten absence/presence matrices were further used for calculating the weight of the arcs: the presence/absence values of each of the arcs (either one or zero) were added across networks; the weight of an arc could take values between 1 and 100. Those arcs whose weight values were equal to or higher than 50 were selected to build the network that best fitted the data. Once the consensus Bayesian networks were built, a further step was taken using the structure of the networks to implement a community analysis. The aim of this community analysis is to identify clusters of nodes densely connected amongst themselves but scarcely connected with nodes between clusters<sup>43</sup>. The approach implemented is similar to the one used when performing cluster analysis, with

the difference that the input is a network. It uses a hierarchical divisive clustering, beginning with the structure of the network, identifying the least connected nodes between two given variables and then, removing these arcs for the next step<sup>43</sup>. The process is repeated several times, until the whole network is divided into smaller communities of densely connected nodes<sup>43</sup>. The R package “igraph”<sup>73</sup> was implemented to identify the communities within the consensus Bayesian networks. Initially, the function *cluster\_edge\_betweenness* was applied to group nodes densely connected, and then the function *dendPlot* was implemented to visualize the results.

**DAVID bioinformatics database.** In order to provide further insights into the biological meaning of the genes in close association with the stress condition, the database for annotation, visualization and integrated discovery (DAVID) was explored<sup>74</sup>. DAVID is a bioinformatic resource, publicly available (<https://david.ncifcrf.gov>), that combines different sources of information, such as protein–protein interactions, bio-pathways, GO terms, homology, literature, among many others, with the aim of closing the gap between a list of statistically significant genes and their functional and biological meaning<sup>74</sup>. DAVID implements some search algorithms to classify the genes into groups of genes that have similar annotation terms, determine which of these genes have overrepresented biological terms, and identify annotations and terms related to a particular gene<sup>74</sup>. DAVID provides a useful bioinformatic resource to further explore and visualize a list of genes, focusing not only on individual genes, but also on groups of genes that might be related to each other based on their annotations. In this study, we used the corresponding ENSEMBL GENE IDs as the input; the Functional Annotation Chart and the Functional Annotation Table provided by DAVID were used as the outputs. The former provides a list of overrepresented annotation terms considering the list of genes as a whole, while the latter is focused on each individual gene and provides their corresponding annotation terms.

Received: 17 December 2021; Accepted: 27 April 2022

Published online: 06 May 2022

## References

- De Kloet, E. R. Hormones, brain and stress. *Endocr. Regul.* **37**, 51–68 (2003).
- Selye, H. & Professor, F. R. S. C. Stress and the general adaptation syndrome. *Br. Med. J.* **1**, 1383–1392 (1950).
- Calefi, A. S., Quinteiro-Filho, W. M., Ferreira, A. J. P. & Palermo-Neto, J. Neuroimmunomodulation and heat stress in poultry. *Worlds. Poult. Sci. J.* **73**, 493–504 (2017).
- Ashley, N. T. & Demas, G. E. Neuroendocrine-immune circuits, phenotypes, and interactions. *Horm. Behav.* **87**, 25–34 (2017).
- Nazar, F. N., Barrios, B. E., Kaiser, P., Marin, R. H. & Correa, S. G. Immune neuroendocrine phenotypes in *Coturnix coturnix*: Do avian species show LEWIS/FISCHER-like profiles? *PLoS ONE* **10**, e0120712 (2015).
- Sapolsky, R., Romero, L. & Munck, A. How do glucocorticoids influence stress responses? Integrating permissive, suppressive, stimulatory, and preparative actions. *Endocr. Rev.* **21**, 55–89 (2000).
- Webster Marketon, J. I. & Glaser, R. Stress hormones and immune function. *Cell Immunol.* **252**, 16–26 (2008).
- Puvadolpirod, S. & Thaxton, J. P. Model of physiological stress in chickens 1. Response parameters. *Poult. Sci.* **79**, 363–369 (2000).
- Davison, F. *Avian Immunology* (Elsevier, 2014).
- Siegel, H. S. Adrenals, stress and the environment. *Worlds. Poult. Sci. J.* **27**, 327–349 (1971).
- Romero, L. M., Dickens, M. J. & Cyr, N. E. The reactive scope model—A new model integrating homeostasis, allostasis, and stress. *Horm. Behav.* **55**, 375–389 (2009).
- Gallup, G. G. Tonic immobility: The role of fear and predation. *Psychol. Rec.* **27**, 41–61 (1977).
- Adelman, J. S. & Martin, L. B. Vertebrate sickness behaviors: Adaptive and integrated neuroendocrine immune responses. *Integr. Comp. Biol.* **49**, 202–214 (2009).
- Ma, X. *et al.* Heat stress impairs the nutritional metabolism and reduces the productivity of egg-laying ducks. *Anim. Reprod. Sci.* **145**, 182–190 (2014).
- Quinteiro-Filho, W. M. *et al.* Heat stress impairs performance and induces intestinal inflammation in broiler chickens infected with *Salmonella* Enteritidis. *Avian Pathol.* **41**, 421–427 (2012).
- BelhadjSlimen, I., Najar, T., Ghram, A. & Abdrrabba, M. Heat stress effects on livestock: Molecular, cellular and metabolic aspects, a review. *J. Anim. Physiol. Anim. Nutr.* **100**, 401–412 (2016).
- Shi, D. *et al.* Impact of gut microbiota structure in heat-stressed broilers. *Poult. Sci.* **98**, 2405–2413 (2019).
- Mashaly, M. M. *et al.* Effect of heat stress on production parameters and immune responses of commercial laying hens. *Poult. Sci.* **83**, 889–894 (2004).
- Akbarian, A. *et al.* Association between heat stress and oxidative stress in poultry; mitochondrial dysfunction and dietary interventions with phytochemicals. *J. Anim. Sci. Biotechnol.* **7**, 37. <https://doi.org/10.1186/s40104-016-0097-5> (2016).
- Altan, O., Pabuçcuoğlu, A., Altan, A., Konyalıoğlu, S. & Bayraktar, H. Effect of heat stress on oxidative stress, lipid peroxidation and some stress parameters in broilers. *Br. Poult. Sci.* **44**, 545–550 (2003).
- Nazar, F. N., Estevez, I., Correa, S. G. & Marin, R. H. Stress induced polarization of immune-neuroendocrine phenotypes in *Gallus gallus*. *Sci. Rep.* **7**, 8102. <https://doi.org/10.1038/s41598-017-08733-0> (2017).
- Ericsson, M. & Jensen, P. Domestication and ontogeny effects on the stress response in young chickens (*Gallus gallus*). *Sci. Rep.* **6**, 35818. <https://doi.org/10.1038/srep35818> (2016).
- McFarlane, J. M. & Curtis, S. E. Multiple concurrent stressors in chicks. 3. Effects on plasma corticosterone and the heterophil:lymphocyte ratio. *Poult. Sci.* **68**, 522–527 (1989).
- Lin, H., Decuyper, E. & Buyse, J. Oxidative stress induced by corticosterone administration in broiler chickens (*Gallus gallus domesticus*): 1. Chronic exposure. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **139**, 737–744 (2004).
- Quinteiro-Filho, W. M. *et al.* Heat stress impairs performance parameters, induces intestinal injury, and decreases macrophage activity in broiler chickens. *Poult. Sci.* **89**, 1905–1914 (2010).
- Elfving, M. *et al.* Early stress causes sex-specific, life-long changes in behaviour, levels of gonadal hormones, and gene expression in chickens. *PLoS ONE* **10**, e0125808. <https://doi.org/10.1371/journal.pone.0125808> (2015).
- Løtvedt, P., Fallahshahroudi, A., Bektic, L., Altimiras, J. & Jensen, P. Chicken domestication changes expression of stress-related genes in brain, pituitary and adrenals. *Neurobiol. Stress* **22**, 121–133 (2017).
- Greene, C. S., Tan, J., Ung, M., Moore, J. H. & Cheng, C. Big data bioinformatics. *J. Cell Physiol.* **229**, 2896–2900 (2014).
- Bélteky, J., Agnvall, B., Johnsson, M., Wright, D. & Jensen, P. Domestication and tameness: Brain gene expression in red junglefowl selected for less fear of humans suggests effects on reproduction and immunology. *R. Soc. Open Sci.* **3**, 160033. <https://doi.org/10.1098/rsos.160033> (2016).
- Zilliox, M. J. & Irizarry, R. A. A gene expression bar code for microarray data. *Nat. Methods* **4**, 911–913 (2007).

31. Matthäus, F., Smith, V. A. & Gebicke-Haerter, P. J. Some useful mathematical tools to transform microarray data into interactive molecular networks. in *Systems Biology in Psychiatric Research: From High-Throughput Data to Mathematical Modeling* (John Wiley & Sons, 2010).
32. Guo, Y. *et al.* Identification of genes related to effects of stress on immune function in the spleen in a chicken stress model using transcriptome analysis. *Mol. Immunol.* **124**, 180–189 (2020).
33. Goerlich, V. C., Nätt, D., Elfving, M., Macdonald, B. & Jensen, P. Transgenerational effects of early experience on behavioral, hormonal and gene expression responses to acute stress in the precocial chicken. *Horm. Behav.* **61**, 711–718 (2012).
34. Guzmán, D. A. *et al.* The fractal organization of ultradian rhythms in avian behavior. *Sci. Rep.* **7**, 684. <https://doi.org/10.1038/s41598-017-00743-2> (2017).
35. Milns, I., Beale, C. M. & Anne Smith, V. Revealing ecological networks using Bayesian network inference algorithms. *Ecology* **91**, 1892–1899 (2010).
36. Felipe, V. P. S., Silva, M. A., Valente, B. D. & Rosa, G. J. M. Using multiple regression, Bayesian networks and artificial neural networks for prediction of total egg production in European quails based on earlier expressed phenotypes. *Poult. Sci.* **94**, 772–780 (2014).
37. Mitchell, E. G., Wallace, M. I., Smith, V. A., Wiesenthal, A. A. & Brierley, A. S. Bayesian network analysis reveals resilience of the jellyfish *Aurelia aurita* to an Irish Sea regime shift. *Sci. Rep.* **11**, 3707. <https://doi.org/10.1038/s41598-021-82825-w> (2021).
38. Comin, A., Jeremiasson, A., Kratzer, G. & Keeling, L. Revealing the structure of the associations between housing system, facilities, management and welfare of commercial laying hens using Additive Bayesian Networks. *Prev. Vet. Med.* **164**, 23–32 (2019).
39. Balov, N. A categorical network approach for discovering differentially expressed regulations in cancer. *BMC Med. Genom.* **6**(Suppl 3), S1. <https://doi.org/10.1186/1755-8794-6-S3-S1> (2013).
40. Heckerman, D., Geiger, D. & Chickering, D. M. Learning Bayesian networks: The combination of knowledge and statistical data. *Mach. Learn.* **20**, 197–243 (1995).
41. Yu, J., Smith, V. A., Wang, P. P., Hartemink, A. J. & Jarvis, E. D. Advances to Bayesian network inference for generating causal networks from observational biological data. *Bioinformatics* **20**, 3594–3903 (2004).
42. Aliferis, C. F., Statnikov, A., Tsamardinos, I., Mani, S. & Koutsoukos, X. D. Local causal and markov blanket induction for causal discovery and feature selection for classification part I: Algorithms and empirical evaluation. *J. Mach. Learn. Res.* **11**, 171–234 (2010).
43. Newman, M. E. J. & Girvan, M. Finding and evaluating community structure in networks. *Phys. Rev. E Stat. Nonlinear Soft Matter Phys.* **69**, 026113. <https://doi.org/10.1103/PhysRevE.69.026113> (2004).
44. Kanehisa, M., Goto, S., Sato, Y., Furumichi, M. & Tanabe, M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* **40**, D109–D114 (2012).
45. Ashburner, M. *et al.* Gene ontology: Tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).
46. Honda, B. T. B. *et al.* Effects of heat stress on peripheral T and B lymphocyte profiles and IgG and IgM serum levels in broiler chickens vaccinated for Newcastle disease virus. *Poult. Sci.* **94**, 2375–2380 (2015).
47. Hirakawa, R. *et al.* Heat stress causes immune abnormalities via massive damage to effect proliferation and differentiation of lymphocytes in broiler chickens. *Front. Vet. Sci.* **7**, 46. <https://doi.org/10.3389/fvets.2020.00046> (2020).
48. Lin, H., Decuyper, E. & Buyse, J. Acute heat stress induces oxidative stress in broiler chickens. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* **144**, 11–17 (2006).
49. Chen, D., Ning, F., Zhang, J., Tang, Y. & Teng, X. NF- $\kappa$ B pathway took part in the development of apoptosis mediated by miR-15a and oxidative stress via mitochondrial pathway in ammonia-treated chicken splenic lymphocytes. *Sci. Total Environ.* **729**, 139017. <https://doi.org/10.1016/j.scitotenv.2020.139017> (2020).
50. Hofmann, K. The modular nature of apoptotic signaling proteins. *Cell Mol. Life Sci.* **55**, 1113–1128 (1999).
51. Bertin, J. *et al.* CARD11 and CARD14 are novel caspase recruitment domain (CARD)/membrane-associated Guanylate Kinase (MAGUK) family members that interact with BCL10 and activate NF- $\kappa$ B. *J. Biol. Chem.* **276**, 11877–11882 (2001).
52. Bertin, J. *et al.* CARD9 is a novel caspase recruitment domain-containing protein that interacts with BCL10/CLAP and activates NF- $\kappa$ B. *J. Biol. Chem.* **275**, 41082–41086 (2000).
53. Wang, L. *et al.* CARD10 is a novel caspase recruitment domain/membrane-associated guanylate kinase family member that interacts with BCL10 and activates NF- $\kappa$ B. *J. Biol. Chem.* **276**, 21405–21409 (2001).
54. Burmester, T., Ebner, B., Weich, B. & Hankeln, T. Cytoglobin: A novel globin type ubiquitously expressed in vertebrate tissues. *Mol. Biol. Evol.* **19**, 416–421 (2002).
55. Schmidt, M. *et al.* Cytoglobin is a respiratory protein in connective tissue and neurons, which is up-regulated by hypoxia. *J. Biol. Chem.* **279**, 8063–8069 (2004).
56. Kugelstadt, D., Haberkamp, M., Hankeln, T. & Burmester, T. Neuroglobin, cytoglobin, and a novel, eye-specific globin from chicken. *Biochem. Biophys. Res. Commun.* **325**, 719–725 (2004).
57. Ren, Z. *et al.* Deoxynivalenol induces apoptosis in chicken splenic lymphocytes via the reactive oxygen species-mediated mitochondrial pathway. *Environ. Toxicol. Pharmacol.* **39**, 339–346 (2015).
58. Qiu, L. *et al.* Comprehensive transcriptome analysis reveals competing endogenous RNA networks during avian leukosis virus, subgroup J-induced tumorigenesis in chickens. *Front. Physiol.* **9**, 996. <https://doi.org/10.3389/fphys.2018.00996> (2018).
59. Aglipay, J. A., Martin, S. A., Tawara, H., Lee, S. W. & Ouchi, T. ATM activation by ionizing radiation requires BRCA1-associated BAAT1. *J. Biol. Chem.* **281**, 9710–9718 (2006).
60. Okada, S. & Ouchi, T. Cell cycle differences in DNA damage-induced BRCA1 phosphorylation affect its subcellular localization. *J. Biol. Chem.* **278**, 2015–2020 (2003).
61. Scully, R. *et al.* Dynamic changes of BRCA1 subnuclear location and phosphorylation state are initiated by DNA damage. *Cell* **90**, 425–435 (1997).
62. Aglipay, J. A. *et al.* A member of the Pryn family, IFI16, is a novel BRCA1-associated protein involved in the p53-mediated apoptosis pathway. *Oncogene* **22**, 8931–8938 (2003).
63. Xie, X., Cho, B. & Fischer, J. A. *Drosophila* Epsin's role in Notch ligand cells requires three Epsin protein functions: The lipid binding function of the ENTH domain, a single Ubiquitin interaction motif, and a subset of the C-terminal protein binding modules. *Dev. Biol.* **363**, 399–412 (2012).
64. Spradling, K. D., McDaniel, A. E., Lohi, J. & Pilcher, B. K. Epsin 3 is a novel extracellular matrix-induced transcript specific to wounded epithelia. *J. Biol. Chem.* **276**, 29257–29267 (2001).
65. Lara, L. & Rostagno, M. Impact of heat stress on poultry production. *Animals* **3**, 356–369 (2013).
66. Renaudeau, D. *et al.* Adaptation to hot climate and strategies to alleviate heat stress in livestock production. *Animal* **6**, 707–728 (2012).
67. Van Goor, A. *et al.* Unique genetic responses revealed in RNAseq of the spleen of chickens stimulated with lipopolysaccharide and short-term heat. *PLoS ONE* **12**, e0171414. <https://doi.org/10.1371/journal.pone.0171414> (2017).
68. Redmond, S. B., Chuammitri, P., Andreasen, C. B., Palić, D. & Lamont, S. J. Chicken heterophils from commercially selected and non-selected genetic lines express cytokines differently after in vitro exposure to *Salmonella enteritidis*. *Vet. Immunol. Immunopathol.* **132**, 129–134 (2009).
69. R Core Team. R: A language and environment for statistical computing. *R Found. Stat. Comput. Vienna, Austria*. URL <http://www.R-project.org> (2021).

70. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2009).
71. Ritchie, M. E. *et al.* Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
72. Hahsler, M., Grün, B. & Hornik, K. Arules—A computational environment for mining association rules and frequent item sets. *J. Stat. Softw.* **14**, 1–25 (2005).
73. Csardi, G. & Nepusz, T. The igraph software package for complex network research. *InterJournal of Complex Systems*, 1695; [http://interjournal.org/manuscript\\_abstract.php?361100992](http://interjournal.org/manuscript_abstract.php?361100992) (2006).
74. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).

## Acknowledgements

This work was supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 812777.

## Author contributions

E.A.V.R. identified the specific study, performed the analyses, prepared the figure, and wrote the initial draft of the manuscript; V.A.S. conceptualized the general study idea; J.B.O.M. and V.A.S. advised on analyses. All authors revised and agreed on the final manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to V.A.S.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

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



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

© The Author(s) 2022