
Genes and Bayesian networks: a strategic approach to poultry and stress.

Emiliano Ariel Videla Rodríguez¹

John B.O. Mitchell²

V. Anne Smith¹

¹School of Biology, University of St Andrews, St Andrews, Fife KY16 9TH United Kingdom.

²School of Chemistry, University of St Andrews, St Andrews, Fife KY16 9ST, United Kingdom

1 EXTENDED ABSTRACT

There are only a few studies implementing Bayesian network (BN) approaches in poultry science. These studies have been focused either on an individual animal level or on an industrial level. Individually, researchers applied BNs approaches to discover key performance variables of egg production in female quail, to identify hallmark single nucleotide polymorphisms of mortality in broiler breeders, and to understand the genetic interactions between genes involved in the metabolism of fatty acids in chickens [Felipe et al., 2014, Li et al., 2010, Long et al., 2009]. As a bridge between these two types of studies, BNs were implemented, among other machine learning techniques, to build an expert system to inspect the freshness of eggs at different storage periods, as a cheap, simple, rapid, non-destructive, and non-harmful method (Soltani Omid, 2015). In terms of the industrial applications, BNs have been implemented to understand i) antimicrobial resistance phenotypes and their interactions with genes, ii) the nutrition and food chain processes of chicken meat, iii) management and welfare-related variables in different housing systems, and finally, iv) poultry green wastes and the quality of the compost [Alvarenga et al., 2021, Comin et al., 2019, Faverial et al., 2016, Hartnack et al., 2019, Hidano et al., 2015, Lupo et al., 2013].

In our study, we implemented a BN approach to poultry genetics and stress, with the aim of identifying genes involved in the stress response or stress resilience in a poultry animal model, the chicken (*Gallus gallus*), using publicly available genetic databases. We initially identified a set of 25 differentially expressed genes (DEG) in the brains of 6 chickens raised under either non-stress ($n = 3$) or stress ($n = 3$) conditions. The number of observations was enough to perform bioinformatic analysis, but not ideal to implement BN algorithms. Even though the number of studies carried out in chicken measuring gene expression levels is relatively small, the previously identified DEGs were searched in 4 other studies sharing the same animal model, the same tissue

(brain), and the same high-throughput technology (microarray). This allowed us to augment the number of observations while increasing the robustness of our findings.

Each dataset was discretised into three-state variables (high-medium-low expression) to reduce the noise potentially associated with the different experimental designs [Balov., 2013]. The software Banjo was used to learn the structure of the BN, using a simulated annealing algorithm and the BDe score, exploring 250 million networks. Exploring the search space several times showed BNs with slight differences in the sets of links the algorithm found. Consequently, 50 BNs built with the top 100 highest scoring networks were combined into a weighted network by selecting those links present in at least 50 percent of the BNs (threshold = 25 out of 50 networks). The links between the DEGs found in at least 25 of the networks had the following strength values: DNAJA4 – BAG3: 49 out of 50 networks; DNAJA4 – HSPA4L: 48; RNPC3 – HSPA4L: 50; HSPH1 – HSPA4L: 49; HSPH1 – CRELD2: 50; XPO1 – HSP90B1: 37; C20orf96 – TNNT3: 34. Even though some links were found in both directions (e.g., C20orf96 – TNNT3 and TNNT3 – C20orf96), the direction of the arrows does not represent causation, but instead a statistical relationships. The biology behind the data was explored with the Database for Annotation, Visualisation, and Integrated Discovery (DAVID).

The consensus network included only 10 out of 25 DEGs, suggesting that even though these DEGs showed differential expression patterns, some genes might not display informative interactions. Among these 10 DEGs, 4 were Heat Shock Proteins (HSP; HSPH1, HSPA4L, DNAJA4, and HSP90B1) and they interacted with themselves as well as with the other 4 DEGs. HSPs play an important role as chaperones, protecting the correct functioning of other proteins, which can be crucial when the birds are dealing with exposure to stress [Goel et al., 2021, Perini et al., 2021]. Among the DEGs, there was an open reading frame whose function is yet unknown, and it interacted with another DEG, TNNT3. This DEG is involved in regulating some nervous process in the

brain of *Drosophila* [Li and Gao., 2003], highlighting the power of BN approaches in knowledge discovery.

Our approach, identifying genes with differential expression patterns, reusing and combining publicly available genetic dataset, and building a BN, brought further informative insights into the stress phenomenon, complementary to those provided by the individual studies. Our results highlight specific genes which could be investigated in future studies to better understand the genetics of stress and its implications for the physiology of poultry species.

References

- T.C. Alvarenga, Lima R.R., J.S.S. Bueno Filho, S.D. Simão, F.C.Q. Mariano, R.R. Alvarenga, and P.B. Rodrigues. Application of Bayesian networks to the prediction of the AMEn: A new methodology in broiler nutrition. *Translational Animal Science*, 5:txaa215, 2021. doi: <https://doi.org/10.1093/tas/txaa215>.
- N. Balov. A categorical network approach for discovering differentially expressed regulations in cancer. *BMC Medical Genomics*, 6(Suppl3):S1, 2013. doi: <https://doi.org/10.1186/1755-8794-6-S3-S1>.
- A. Comin, A. Jeremiasson, G. Kratzer, and L. Keeling. Revealing the structure of the associations between housing system, facilities, management and welfare of commercial laying hens using Additive Bayesian Networks. *Preventive Veterinary Medicine*, 164:23–32, 2019. doi: <https://doi.org/10.1016/j.prevetmed.2019.01.004>.
- J. Faverial, D. Cornet, J. Paul, and J. Sierra. Multivariate analysis of the determinants of the end-product quality of manure-based composts and vermicomposts using Bayesian network modelling. *PLoS ONE*, 11(6): e0157884, 2016. doi: <https://doi.org/10.1371/journal.pone.0157884>.
- V.P.S. Felipe, M.A. Silva, B.D. Valente, and G.J.M. Rosa. Using multiple regression, Bayesian networks and artificial neural networks for prediction of total egg production in European quails based on earlier expressed phenotypes. *Poultry Science*, 94(4):772–780, 2014. doi: <https://doi.org/10.3382/ps/pev031>.
- A. Goel, C.M. Ncho, and Y.H. Choi. Regulation of gene expression in chickens by heat stress. *Journal of Animal Science and Biotechnology*, 12:11, 2021. doi: <https://doi.org/10.1186/s40104-020-00523-5>.
- S. Hartnack, T. Odoch, G. Kratzer, R. Furrer, Y. Wasteson, T.M. L'Abée-Lund, and E. Skjerve. Additive Bayesian networks for antimicrobial resistance and potential risk factors in non-typhoidal *Salmonella* isolates from layer hens in Uganda. *BMC Veterinary Research*, 15:212, 2019. doi: <https://doi.org/10.1186/s12917-019-1965-y>.
- A. Hidano, T. Yamamoto, Y. Hayama, N. Muroga, S. Kobayashi, T. Nishida, and T. Tsutsui. Unraveling antimicrobial resistance genes and phenotype patterns among *Enterococcus faecalis* isolated from retail chicken products in Japan. *PLoS ONE*, 10(3):e0121189, 2015. doi: <https://doi.org/10.1371/journal.pone.0121189>.
- H. Li, G. Wu, J. Zhang, and N. Yang. Identification of the heart-type fatty acid-binding protein as a major gene for chicken fatty acid metabolism by bayesian network analysis. *Poultry Science*, 89(9):1825–1833, 2010. doi: <https://doi.org/10.3382/ps.2010-00699>.
- W. Li and F.B. Gao. Actin filament-stabilizing protein tropomyosin regulates the size of dendritic fields. *Journal of Neuroscience*, 23(15):6171–6175, 2003. doi: <https://doi.org/10.1523/JNEUROSCI.23-15-06171.2003>.
- N. Long, D. Gianola, G.J. Rosa, K.A. Weigel, and S. Avendaño. Comparison of classification methods for detecting associations between SNPs and chick mortality. *Genetics Selection Evolution*, 41:18, 2009. doi: <https://doi.org/10.1186/1297-9686-41-18>.
- C. Lupo, S. Le Bouquin, L. Balaine, V. Michel, J. Péraste, I. Petetin, P. Colin, L. Jouffe, and C. Chauvin. Bayesian network as an aid for Food Chain Information use for meat inspection. *Preventive Veterinary Medicine*, 109(1-2):25–36, 2013. doi: <https://doi.org/10.1016/j.prevetmed.2012.09.004>.
- F. Perini, F. Cendron, G. Rovelli, C. Castellini, M. Cas-sandro, and E. Lasagna. Emerging genetic tools to investigate molecular pathways related to heat stress in chickens: A review. *Animals*, 11(1):46, 2021. doi: <https://doi.org/10.3390/ani11010046>.