

OPEN

# Gene co-expression network analysis reveals key potential gene modules in utero-vaginal junction associated with duration of fertility trait of breeder hens

Lantao Gu<sup>1,2</sup>, Ruoxi Jing<sup>1</sup>, Yanzhang Gong<sup>1</sup>, Mei Yu<sup>1</sup>, Abdelmotalieb Elokil<sup>1,3</sup>  & Shijun Li<sup>1</sup>

The number of days (DN) when hens lay fertile eggs as well as the number of fertile eggs (FN) were produced after a single artificial insemination (AI), including the two duration of fertility (DF) traits. Indeed, they are the key production performance that associates with the production cost of hatching egg when its determination the interval between successive artificial inseminations. However, the relevant genes response for regulating the DF has not been uncovered yet. Therefore, we performed a weighted gene co-expression network analysis (WGCNA) to investigate the insight into co-expression gene modules on DF process in hens. The total mRNA was extracted from the utero-vaginal junction (UVJ, with the sperm storage function in hen's oviduct which is the biological basis for DF) of 20 hens with several levels of DF traits, and performed transcriptome sequences of mRNA. As a result, three co-expression gene modules were identified to be highly correlated with DF traits. Moreover, the expression changes of top 5 hub genes in each module with DF traits were further confirmed in other 20 hens by RT-PCR. These findings highlighted the co-expression modules and their affiliated genes as playing important roles in the regulation of DF traits.

The avian females have the ability to store sperm in their reproductive tracts following the natural copulation or AI for days or weeks (depending on the species), locating utero-vaginal junction (UVJ). Thus, its produce a series of fertile eggs following a single copulation event or artificial insemination (AI)<sup>1</sup>. This period of sustained fertility is generally identified as the duration of fertility (DF, days number between insemination and the last fertile egg)<sup>2</sup>. The long or short DF are identified the frequency of AI because any reduction in this frequency without loss of fertility directly reduces the cost of producing hatching eggs, thus to cut down the number of breeder males and labor costs associated with AI<sup>3</sup>. Meanwhile, it could alleviate the stress and suffering of hens from frequent AI. Therefore, study on the molecular mechanism that regulating DF traits can not only develop a clear understanding of this process, but also can improve the economic efficiency for production the hatching eggs. In mature hens' oviduct, many specialized simple tubulars invaginate the utero-vaginal junction (UVJ) mucosal folds<sup>4</sup>. These tubulars protects a long-time sperm storage function, thus were referred to as sperm storage tubules (SSTs). The utero-vaginal sperm storage tubules functions as a sperm reservoir in hen's oviduct<sup>5</sup>. Here, sperms are ultimately released and upward transport to the infundibulum for ova fertilization. This function gives hens a senseful capacity that hens could lay fertile eggs for days or weeks without frequently insemination<sup>6,7</sup>. Therefore, investigation the relationship of gene expression in UVJ to the duration of fertility traits could add our understanding about the molecular regulatory mechanism participating in the duration of fertility process.

In this study, RNA sequenced and WGCNA presented the key potential gene co-expression modules associated with breeder hens' duration of fertility traits and their possible implications in duration of fertility process.

<sup>1</sup>Key Laboratory of Agricultural Animal Genetics, Breeding and Reproduction of Ministry of Education, College of Animal Science and Technology, Huazhong Agricultural University, Wuhan, Hubei, China. <sup>2</sup>Guilin Medical University, Guilin, Guangxi, China. <sup>3</sup>Department of Animal Production, Faculty of Agriculture, Benha University, Moshtohor, Egypt. Correspondence and requests for materials should be addressed to S.L. (email: [lishijun@mail.hzau.edu.cn](mailto:lishijun@mail.hzau.edu.cn))

## Material and Methods

**Ethics statement.** The experimental procedures used in this study met the guidelines of the Care and Use of Laboratory Animals of the Standing Committee of Hubei People's Congress (No. 5) and approved by the Biological Studies Animal Care Committee of Hubei Province, P.R. China, and the Ethics Committee of Huazhong Agricultural University, PR China. All efforts were made to minimize the animal suffering.

**Animal management and trait measurement.** A total of 450 healthy Jinghong breeder hens (parents) at 30 weeks old were obtained from the poultry farm of Huadu Yukou Poultry Industry Co. Ltd (Beijing, China). All birds were raised in individual cages, kept in identical light/dark cycles and had ad libitum access to healthy water and a commercial diet until the end of the experiment. All hens were artificially inseminated once with  $2.00 \times 10^8$  sperms issued from pooled ejaculates. To prevent any undesirable effects of the interval between inseminations and oviposition on subsequent fertility, insemination was done identically in afternoon. Eggs were collected and marked daily for 20 days after AI and checked the fertility by candling eggs after 10 days from incubation (dead embryos were considered as fertile). Two duration of fertility traits DN and FN were measured in terms (DN = the number of days post-insemination until the last fertile egg was produced; and FN = the number of fertile eggs were produced after a single AI). DN and FN were expressed as average of three measurements<sup>8</sup>.

**RNA sequencing and data preprocessing.** A total of 21 hens with their phenotypes distributed throughout duration of fertility traits were sorted out from the experimental population (Table S2). On the 13<sup>th</sup> day after AI, these hens were euthanized by decapitation under anesthesia. UVJ tissues were dissected immediately and removed adhering connective tissues<sup>9</sup>. Total RNA was isolated using Trizol reagent (Invitrogen, Foster City, CA, USA), following the recommended manufacturers protocol. The quality and quantity of RNA samples were detected by 1.0% agarose gel electrophoresis and absorbance optical density (OD) at a 260/280 nm ratio, respectively<sup>8</sup>. Pair-end 150-bp reads were generated on a Hiseq 2500 sequencing machine at the Shanghai biotechnology corporation. Additionally, aligned to the chicken reference genome by HISAT2(version: 2.0.9). A raw counts data was obtained from the HTSEQ-count (version: 3.0.8) Soft-ware and the genes with row counts less than 15 in all samples were removed<sup>10,11</sup>. These steps finally resulted in more than 17000 genes to infer next study. Then the raw counts data were processed and normalized by edgeR Soft-ware to generate an FPKM expression matrix<sup>12</sup>. Sample clustering was performed using the hierarchical average linkage clustering function implemented in the WGCNA R package to identify possible outliers. One sample (Sample ID: 7) was identified as an outlier and thus excluded from the final dataset (Supplementary File 1, Fig. S1).

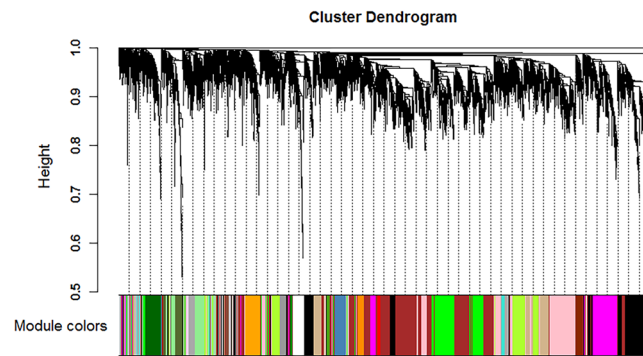
**Co-expression network construction and meta-modules detection.** Networks were formed following the protocols of WGCNA<sup>13</sup>. Briefly, a similarity matrix was constructed using the Pearson correlation coefficients created between the normalized expression levels of the input transcripts. By raising the absolute value of the Pearson correlation coefficients to a power of 9, we could get a scale-free topology index above 0.9, achieving a network with few, large correlations at the expense of lowly correlated transcripts. This allows for the fewer, highly connected and biologically relevant hub genes. An adjacency network was then created using topological overlap measure (TOM, a measure of neighborhood connectivity between two genes)<sup>14</sup>. To create modules, the adjacency network was converted into a dissimilarity measure (1-TOM) and clustered using flashClust (a hierarchical clustering function). Cluster branches were cut to identify modules. Module size was set to a minimum of 30 transcripts and modules with a 75% similarity were merged using dynamic tree cutting, resulting in 54 modules<sup>10</sup>.

Module eigengene (Me) was defined as the first principal component of a given module, which represented the gene expression profiles in a module. Generally, a module with high correlation between Me and DF traits could be meta-module that associated with DF traits. Hub genes tends to be in the center of a network, highly connected with other genes and hence of high functional significance. The intramodular connectivity (K.in) was calculated as the summation of adjacency performed over all genes in a network. A gene with high K.in means that this gene is the most important elements of modules associated with relevant trait. Therefore, a gene with high K.in in a module was hub gene.

**RNA expression analyses by quantitative RT-PCR.** At last, total RNA from UVJ tissues of the 20 hens (one of 21 hens was found to be outliers) were isolated. cDNA was synthesized using TransScript One-Step gDNA Removal and cDNA Synthesis SuperMix (Trans, Beijing) as recommended by the manufacturer. The resulting cDNA was subjected to real-time quantitative PCR with SYBR Green Mix (QIAGEN, Germany) and specific primer sets (Table S4) in accordance with the manufacturer instructions. PCR products were verified by melting curve analysis. The relative quantification expression was calculated using the delta-delta Ct method with each gene normalized to GAPDH.

## Results

**RNA sequencing and weighted gene co-expression network construction.** Respectively, the 450 hens' duration of fertility traits were  $14.31 \pm 2.38$  days (Coefficient of variation: 16.66%; Percentiles 2.5–97.5%: 8.35–18.00) for DN, and  $10.66 \pm 2.53$  eggs (Coefficient of variation: 23.71%; Percentiles 2.5–97.5%: 4.67–15.00) for FN. The Correlation coefficients of FN and DN phenotypic values was 0.53. Thereafter, the duration of fertility of hens were divided into 10 levels according to the DN and FN respectively. To make sure that there are at least one sample in each level, a total of 21 hens with their duration of fertility traits distributed throughout ten levels were sorted out for RNA sequencing (Tables S1 and S2). More than 40 million sequencing reads were generated for each sample, and more than 70% of these reads were exactly mapped onto the chicken reference genome (Table S3). After filter genes with row counts less than 15 in all samples, we finally obtained the expression profiling of more than 17 000 genes in UVJ. But one sample (Sample ID: 7) was found to be an outlier and thus excluded from the final dataset.



**Figure 1.** Identification of gene co-expression modules in utero-vaginal junction of hens. Module size was set to a minimum of 30 transcripts and modules with a 75% similarity were merged using dynamic tree cutting, resulting in 54 color-coded modules.

After preprocessing the RNA sequencing data, we applied the WGCNA package to compile the co-expression network. Keeping to the scale-free topology criterion,  $\beta = 9$  was considered in this study. Following dynamic tree cut, the topological overlap clustering dendrogram identified 54 distinct gene modules, as can be seen from the color-band underneath the cluster tree (Fig. 1). The gray module consisted of genes that do not group into any specific module.

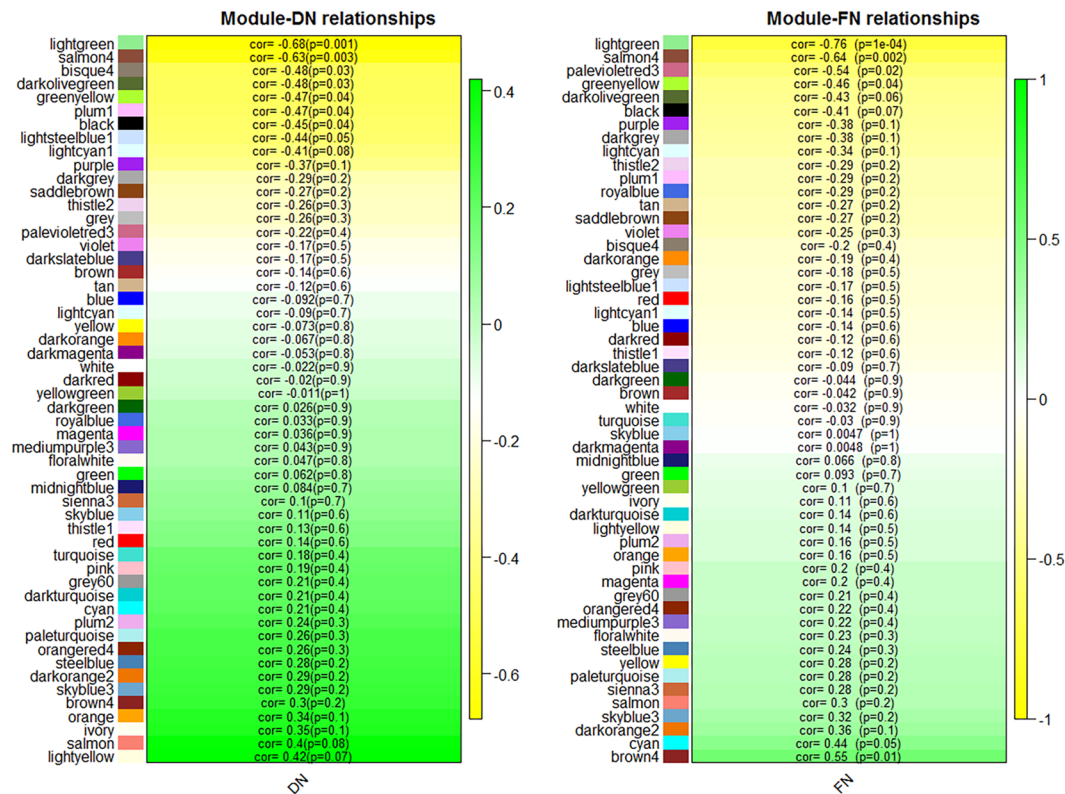
**Identification of meta-modules associated with duration of fertility.** To identify co-expression modules associated with duration of fertility, we assessed the relationship of duration of fertility traits DN and FN with the module eigengene (Me). Two modules (lightgreen and salmon4) showed strong association with DN ( $P < 0.01$ ) while five modules (bisque4, darkolivegreen, greenyellow, plum1 and black) showed association with DN ( $P < 0.05$ ). The two modules (lightgreen and salmon4) also showed strong association with FN ( $P < 0.01$ ), and three modules (palevioletred3, greenyellow and brown4) were showed associated with FN ( $P < 0.05$ ). Overall, significant correlations of module eigengene with DN and FN both were observed in three modules (lightgreen, salmon4 and greenyellow), thereafter named duration of fertility modules (Fig. 2). The visualized the connectivity patterns and hub genes of the meta-modules were presented in Fig. 3.

**Validation of hub genes.** We validated the relative expression levels of hub genes from the three duration of fertility modules using RT-PCR (Fig. 4). The results were consistent with those obtained by RNA-seq, which the expression levels of 12 genes (*SPG11*, *CTDSPL2*, *MFAP3L*, *LRP3*, *RALGAPA1*, *MBNL1*, *SGSM2*, *SCML2*, *KLHL13*, *DCAF6*, *KBTBD2* and *CLK2*) were all significantly down-regulated in the UVJ correlated with duration of fertility traits DN or FN, and the expression levels of 3 genes (*ZCRB1*, *PSMA7* and *ANAPC13*) were all significantly up-regulated in the UVJ correlated with duration of fertility traits.

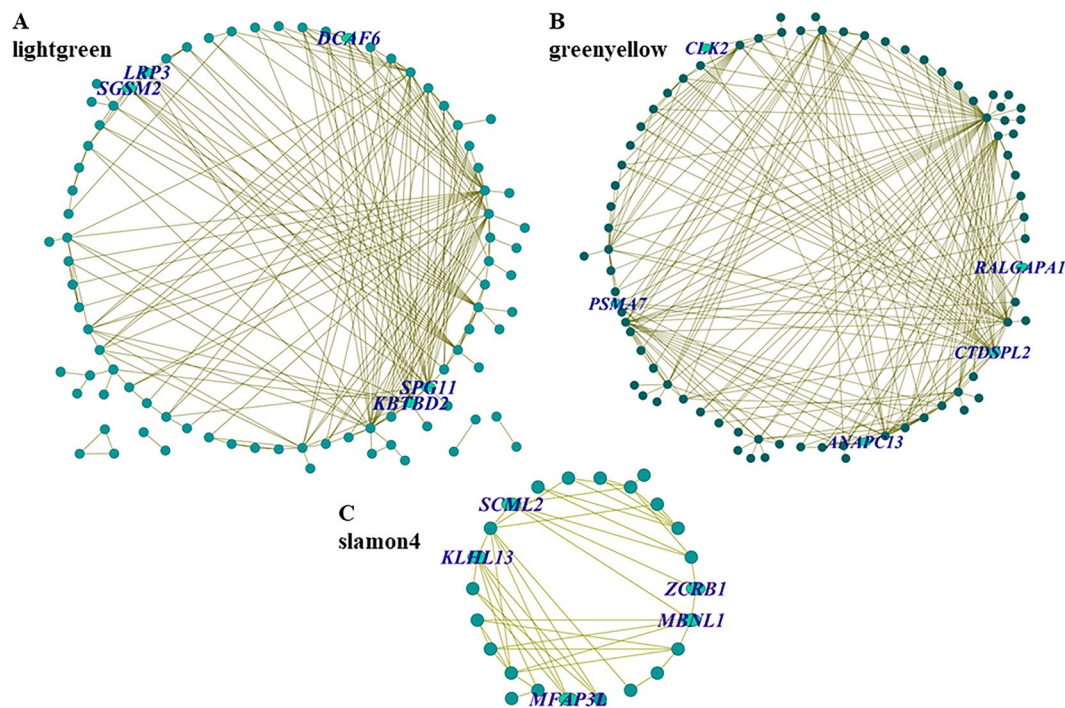
## Discussion

In the present study, we performed RNA sequencing and a systematic WGCNA to gain a better understanding of the transcriptomic changes in UVJ tissues of hens with duration of fertility trait on a genome-wide scale. Our findings revealed that fifty-four co-expression gene modules were discovered in UVJ of hens and three modules (named light-green, salmon4 and green-yellow) were found significantly associated with duration of fertility traits in this study. It is believed that modules are co-expressed gene clusters with integrated function in biological process. Therefore, functional modules can facilitate the finding of hub genes that are key drivers of consistent biological process<sup>15</sup>. To sum up, the duration of fertility was a complex and systematic process: surviving sperms from the UVJ environment and immunological response, retaining sperms in the SST, activating and taxiing sperm towards oocytes, and removing non-fertilizing sperm<sup>16,17</sup>. The investigation of these modules revealed novel hub genes and pathways that are likely biologically relevant to these events. The green-yellow module consisted of 351 genes, and the top five hub genes were *CTDSPL2*, *RALGAPA1*, *CLK2*, *PSMA7* and *ANAPC13*. These genes (*ANAPC13*, *PSMA7* and *CLK2*) have been reported to mainly involving adaptive immune and play an important regulatory mechanism required for promoting cell survival<sup>18–21</sup>. In previous study, it is reported that the female sperm allergy is an important cause leading to the problem of sperm survival, this allergy was present in the hens with short duration of fertility<sup>22,23</sup>. It indicated that the necessary of immune privilege for sperms residing and surviving in SSTs. Along these lines, these genes driving the adaptive immune and cell survival might response for this event. The rest, *ANAPC13* participated in progesterone-mediated oocyte maturation<sup>24</sup>, similarly the *CTDSPL2* has been found to participate in gene regulation during erythroid maturation<sup>25</sup>, and the mRNA level of *RALGAPA1* was reported to specifically associated with total egg number at 500 d of age or egg rate after the first egg<sup>26</sup>. They also seem to be associated with duration of fertility.

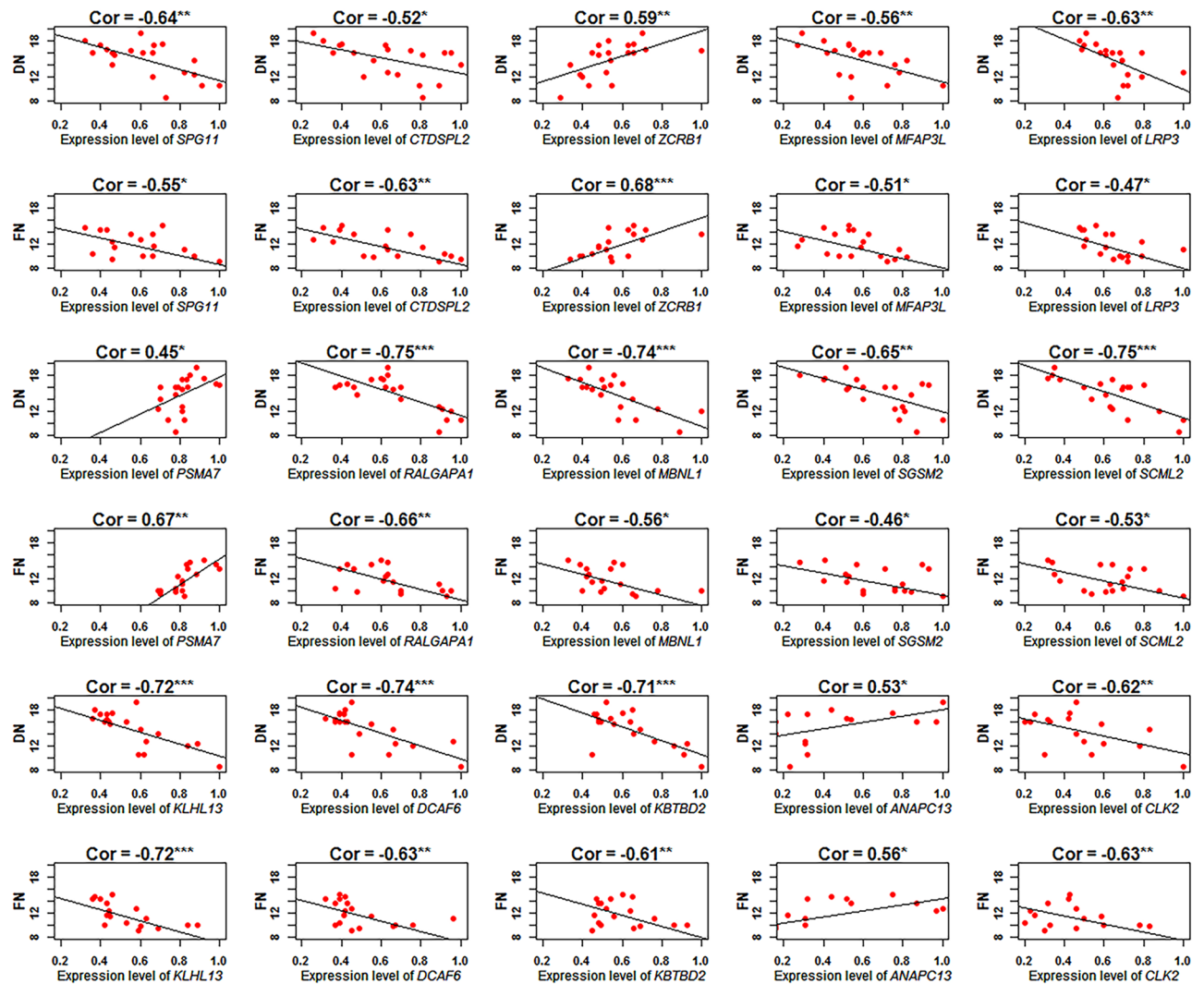
The salmon4 module consisted of 40 genes, and the top five hub gene including *MBNL1*, *MFAP3L*, *ZCRB1*, *KLHL13* and *SCML2*. Among these genes, the *MBNL1* and *ZCRB1* both mediate the wide range of tissue-specific RNA processing steps<sup>27</sup>. Particularly, *MBNL1* was reported to negatively regulates *TGF $\beta$ 3* production auto-crine<sup>28</sup>, which coincide to the association of the expression level of *TGF $\beta$ 3* in UVJ with duration of fertility<sup>8</sup>. Otherwise, the *KLHL13* was reported to participated in adaptive immune system, while the *SCML2* was involved



**Figure 2.** Heatmap reporting correlations (and corresponding p-values) between modules and duration of fertility traits. DN = the number of days post-insemination until the last fertile egg was produced; and FN = the number of fertile eggs were produced after a single AI.



**Figure 3.** Interaction of gene co-expression network of the duration of fertility related modules. The modules (A: lightgreen, B: greenyellow, C: slamon4) was visualized using Cytoscape 3.0 software. Every node represents a gene, and the nodes labeled by gene symbols were the top five hub genes as ranked by intra-module connectivity.



**Figure 4.** The hub gene expression changes correlated with duration of fertility traits DN and FN. DN = the number of days post-insemination until the last fertile egg was produced; and FN = the number of fertile eggs were produced after a single AI. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ .

in transcriptional regulation<sup>29</sup> and the *MFAP3L* was reported to participate in the nuclear signaling of EGFR which involving angiogenesis. We have known that the SST was specialized simple tubular invaginations of the utero-vaginal junction (UVJ) mucous epithelium, and its tissue morphology is like the blood vessels. The genesis of SST might involve this process.

The light-green module consisted of 228 genes, and the top five hub gene including *KBTBD2*, *SPG11*, *DCAF6*, *SGSM2* and *LRP3*. These genes show a mainly biological function of material transport and metabolism. Such as *SGSM2* and *SPG11* was reported to participated in intercellular vesicle transport<sup>30</sup>. The *LRP3* plays a role in endocytosis<sup>31</sup>. The *KBTBD2* participated in lip metabolism<sup>32</sup>, and the *DCAF6* show effects on protein ubiquitination and degradation<sup>33</sup>. These processes might be mediated the supplication of nutrients to the sperm, removing of waste products of sperm metabolism, and the phagocytosis of non-fertilizing sperm.

In conclusion, we found three gene co-expression network and their affiliated hub genes that are associated with duration of fertility traits. Although interpretation of the results in the present study is limited and more surveys will be needed to determine the physiological functions of these genes, it contributes to the understanding of gene regulatory mechanism on duration of fertility.

## References

- Bakst, M. R. Oviducal sperm selection, transport, and storage in poultry. *Poultry science Rev* 5, 7 (1994).
- Bakst, M. R. & Vinyard, B. T. Oviducal sperm storage in turkeys: spatial distribution of sperm within the uterovaginal junction sperm-storage tubules. *The Journal of experimental zoology* 292, 206–209 (2002).
- Beaumont, C., Brillard, J. P., Millet, N. & De Reviere, M. Comparison of various characteristics of duration of fertility in hens. *British poultry science* 33, 649–661, <https://doi.org/10.1080/00071669208417503> (1992).
- Bakst, M. R. Anatomical basis of sperm-storage in the avian oviduct. *Scanning microscopy* 1, 1257–1266 (1987).
- Blesbois, E. & Brillard, J. P. Specific features of *in vivo* and *in vitro* sperm storage in birds. *Animal: an international journal of animal bioscience* 1, 1472–1481, <https://doi.org/10.1017/S175173110700081X> (2007).

6. Bakst, M. R. Physiology and endocrinology symposium: role of the oviduct in maintaining sustained fertility in hens. *Journal of animal science* **89**, 1323–1329, <https://doi.org/10.2527/jas.2010-3663> (2011).
7. Birkhead, T. R. & Brillard, J. P. Reproductive isolation in birds: postcopulatory prezygotic barriers. *Trends in ecology & evolution* **22**, 266–272, <https://doi.org/10.1016/j.tree.2007.02.004> (2007).
8. Gu, L., Sun, C., Gong, Y., Yu, M. & Li, S. Novel copy number variation of the TGFbeta3 gene is associated with TGFbeta3 gene expression and duration of fertility traits in hens. *PLoS one* **12**, e0173696, <https://doi.org/10.1371/journal.pone.0173696> (2017).
9. Bakst, M.R. Anatomical basis of sperm-storage in the avian oviduct. *Scanning microscopy* **1**, 9 (1987).
10. Salleh, S. M., Mazzoni, G., Lovendahl, P. & Kadarmideen, H. N. Gene co-expression networks from RNA sequencing of dairy cattle identifies genes and pathways affecting feed efficiency. *BMC bioinformatics* **19**, 513, <https://doi.org/10.1186/s12859-018-2553-z> (2018).
11. Zhang, H. *et al.* Genome-wide maps of ribosomal occupancy provide insights into adaptive evolution and regulatory roles of uORFs during Drosophila development. *PLoS biology* **16**, e2003903, <https://doi.org/10.1371/journal.pbio.2003903> (2018).
12. Chen, Y., Pal, B., Visvader, J. E. & Smyth, G. K. Differential methylation analysis of reduced representation bisulfite sequencing experiments using edgeR. *F1000Research* **6**, 2055, <https://doi.org/10.12688/f1000research.13196.2> (2017).
13. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC bioinformatics* **9**, 559, <https://doi.org/10.1186/1471-2105-9-559> (2008).
14. Pflieger, L. T. *et al.* Gene co-expression network analysis for identifying modules and functionally enriched pathways in SCA2. *Human molecular genetics* **26**, 3069–3080, <https://doi.org/10.1093/hmg/ddx191> (2017).
15. Lou, Y. *et al.* Characterization of transcriptional modules related to fibrosing-NAFLD progression. *Scientific reports* **7**, 4748, <https://doi.org/10.1038/s41598-017-05044-2> (2017).
16. Publicover, S., Harper, C. V. & Barratt, C. [Ca<sup>2+</sup>]<sub>i</sub> signalling in sperm-making the most of what you've got. *Nature cell biology* **9**, 235–242, <https://doi.org/10.1038/ncb0307-235> (2007).
17. Das, S. C., Isobe, N. & Yoshimura, Y. Mechanism of prolonged sperm storage and sperm survivability in hen oviduct: a review. *American journal of reproductive immunology* **60**, 477–481, <https://doi.org/10.1111/j.1600-0897.2008.00651.x> (2008).
18. Hearn, A., York, I. A., Bishop, C. & Rock, K. L. Characterizing the specificity and cooperation of aminopeptidases in the cytosol and endoplasmic reticulum during MHC class I antigen presentation. *Journal of immunology* **184**, 4725–4732, <https://doi.org/10.4049/jimmunol.0903125> (2010).
19. Minnicozzi, M., Sawyer, R. T. & Fenton, M. J. Innate immunity in allergic disease. *Immunological reviews* **242**, 106–127, <https://doi.org/10.1111/j.1600-065X.2011.01025.x> (2011).
20. Zheng, X. *et al.* Salivary exosomal PSMA7: a promising biomarker of inflammatory bowel disease. *Protein & cell* **8**, 686–695, <https://doi.org/10.1007/s13238-017-0413-7> (2017).
21. Nam, S. Y. *et al.* Phosphorylation of CLK2 at serine 34 and threonine 127 by AKT controls cell survival after ionizing radiation. *The Journal of biological chemistry* **285**, 31157–31163, <https://doi.org/10.1074/jbc.M110.122044> (2010).
22. Das, S. C., Nagasaka, N. & Yoshimura, Y. Changes in the localization of antigen presenting cells and T cells in the utero-vaginal junction after repeated artificial insemination in laying hens. *The Journal of reproduction and development* **51**, 683–687 (2005).
23. Lu, J. C., Huang, Y. F. & Lu, N. Q. Antisperm immunity and infertility. *Expert review of clinical immunology* **4**, 113–126, <https://doi.org/10.1586/1744666X.4.1.113> (2008).
24. Sokolowski, B., Orchard, S., Harvey, M., Sridhar, S. & Sakai, Y. Conserved BK channel-protein interactions reveal signals relevant to cell death and survival. *PLoS one* **6**, e28532, <https://doi.org/10.1371/journal.pone.0028532> (2011).
25. Wani, S., Sugita, A., Ohkuma, Y. & Hirose, Y. Human SCP4 is a chromatin-associated CTD phosphatase and exhibits the dynamic translocation during erythroid differentiation. *Journal of biochemistry* **160**, 111–120, <https://doi.org/10.1093/jb/mvw018> (2016).
26. Shen, X. *et al.* The GTPase activating Rap/RanGAP domain-like 1 gene is associated with chicken reproductive traits. *PLoS one* **7**, e33851, <https://doi.org/10.1371/journal.pone.0033851> (2012).
27. Fish, L. *et al.* Muscleblind-like 1 suppresses breast cancer metastatic colonization and stabilizes metastasis suppressor transcripts. *Genes & development* **30**, 386–398, <https://doi.org/10.1101/gad.270645.115> (2016).
28. LeMasters, K. E., Blech-Hermoni, Y., Stillwagon, S. J., Vajda, N. A. & Ladd, A. N. Loss of muscleblind-like 1 promotes invasive mesenchyme formation in endocardial cushions by stimulating autocrine TGFbeta3. *BMC developmental biology* **12**, 22, <https://doi.org/10.1186/1471-213X-12-22> (2012).
29. Bonasio, R. *et al.* Interactions with RNA direct the Polycomb group protein SCML2 to chromatin where it represses target genes. *eLife* **3**, e02637, <https://doi.org/10.7554/eLife.02637> (2014).
30. Perez-Branguli, F. *et al.* Dysfunction of spatacsin leads to axonal pathology in SPG11-linked hereditary spastic paraplegia. *Human molecular genetics* **23**, 4859–4874, <https://doi.org/10.1093/hmg/ddu200> (2014).
31. Ishii, H. *et al.* cDNA cloning of a new low-density lipoprotein receptor-related protein and mapping of its gene (LRP3) to chromosome bands 19q12-q13. 2. *Genomics* **51**, 132–135, <https://doi.org/10.1006/geno.1998.5339> (1998).
32. Zhang, Z. *et al.* Insulin resistance and diabetes caused by genetic or diet-induced KBTBD2 deficiency in mice. *Proceedings of the National Academy of Sciences of the United States of America* **113**, E6418–E6426, <https://doi.org/10.1073/pnas.1614467113> (2016).
33. Chen, H. H. *et al.* NRIP/DCAF6 stabilizes the androgen receptor protein by displacing DDB2 from the CUL4A-DDB1 E3 ligase complex in prostate cancer. *Oncotarget* **8**, 21501–21515, <https://doi.org/10.18632/oncotarget.15308> (2017).

## Acknowledgements

The current research was financially supported by the National Nature Scientific Foundation of China (31772585), and the Fundamental Research Funds for the Central Universities of China (number 2662015PY145 and 2662018PY088). The funders did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Author Contributions

Lantao Gu and Shijun Li. conceived and designed the study. Ruoxi Jing, Yanzhang Gong, Mei Yu and Lantao Gu analyzed data, and Lantao Gu wrote the paper while Abdelmotaleb Elokil revised the article in English. All authors contributed to revision and review of the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-019-50148-6>.

**Competing Interests:** The authors declare no competing interests.

**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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019