Transcriptome sequencing reveals key potential long non-coding RNAs related to duration of fertility trait in the uterovaginal junction of egg-laying hens

Adeyinka Abiola Adetula1, Lantao Gu1, Chinedu Charles Nwafor2, Xiaoyong Du3, Shuhong Zhao1 & Shijun Li1

Duration of fertility, (DF) is an important functional trait in poultry production and IncRNAs have emerged as important regulators of various process including fertility. In this study we applied a genome-guided strategy to reconstruct the uterovaginal junction (UVJ) transcriptome of 14 egg-laying birds with long- and short-DF (n = 7); and sought to uncover key IncRNAs related to duration of fertility traits by RNA-sequencing technology. Examination of RNA-seq data revealed a total of 9977 IncRNAs including 2576 novel IncRNAs. Differential expression (DE) analysis of IncRNA identified 223 IncRNAs differentially expressed between the two groups. DE-IncRNA target genes prediction uncovered over 200 IncRNA target genes and functional enrichment tests predict a potential function of DE-IncRNAs. Gene ontology classification and pathway analysis revealed 8 DE-IncRNAs, with the majority of their target genes enriched in biological functions such as reproductive structure development, developmental process involved in reproduction, response to cytokine, carbohydrate binding, chromatin organization, and immune pathways. Differential expression of IncRNAs and target genes were confirmed by qPCR. Together, these results significantly expand the utility of the UVJ transcriptome and our analysis identification of key IncRNAs and their target genes regulating DF will form the baseline for understanding the molecular functions of IncRNAs regulating DF.

Duration of fertility (DF) is defined as the number of day’s post-fertilisation when viable eggs are produced. It is an important trait in egg-laying hens. For efficient and profitable poultry production prolonged duration of fertility for a breeder stock is vital. Studies on artificial insemination (AI) revealed the relationship between the number of spermatozoa inseminated and embryo survival influence duration of fertility1. Prolonged sperm storage and its longevity within the sperm storage tubules (SST) in the uterovaginal junction (UVJ) of laying hens were associated with DF trait2–4. The possibility of increasing the AI intervals by improving DF via selection of DF traits has been proven5–7.

To date, several proteins and enzyme-coding genes have been associated with fertility potential of laying birds, including, carbonic anhydrase8, avidin and avidin-related protein-2 (AVR2)9, aquaporins10,11, alkaline phosphatase12, progesterone receptor13, transforming growth factor-β (TGF-β) and its receptors14. Additionally, immune system activity has been implicated in maintaining sperm storage within the SST and at the time of fertilization15,16. Likewise, breeder and co-workers proposed that energy derived from mitochondrial fatty acid oxidation might be important for sperm longevity, motility, mobility and storage in the oviduct. Despite these extensive studies causes of long-DF, short-DF, and in some cases, infertility in egg-laying hens is largely unknown. Especially the transcriptional and post-transcriptional regulation of gene function during oviductal sperm selection, transport, and storage17.

Received: 5 March 2018
Accepted: 13 August 2018
Published online: 04 September 2018

1Key Laboratory of Agricultural Animal Genetics, Breeding, and Reproduction, Ministry of Education, Huazhong Agricultural University, Wuhan, Hubei Province, 430070, China. 2Faculty of Agriculture, Benson Idahosa University, Benin City, Edo State, Nigeria. 3College of Informatics, Huazhong Agricultural University, Wuhan, 430070, China. Correspondence and requests for materials should be addressed to S.L. (email: lishijun@mail.hzau.edu.cn)
Long non-coding RNAs (lncRNAs) have emerged as important molecules for the transcriptional and post-transcriptional regulation of gene expression evidenced by their tissue-specific expression patterns and subcellular localization. LncRNAs required for spermatogenesis and fertility have been identified in Drosophila. Also, regulatory elements such as enhancers and regulatory non-coding RNAs that are spermatozoa-specific in mammalian species have been identified, supporting the possibility that lncRNAs may impact DF in egg-laying hens. In chicken, only a small percentage of all annotated lncRNAs have been functionally characterized, but several lncRNAs have been identified including those associated with skeletal muscle development, liver and adipose tissues as well as several lncRNAs in the male testis with extreme sperm motility.

The sperm storage tubule (SST) located in the uterovaginal junction (UVJ) is considered the main site of the residence of spermatozoa and are highly related to fertility and reproductive traits. Furthermore, the ultrastructural analysis of UVJ tissue showed that more SSTs were embedded in the long-DF hens compared with short-DF hens. There was a significant difference in the number of SSTs between the two groups (21.5 ± 4.03 vs. 4.00 ± 0.91) in long and short-DF hens respectively. These results were able to distinguish the long and short-DF UVJ tissues.

Results
Duration of fertility trait in egg-laying hens. A total of 304 hens were recorded daily by the DF trait: DN: the days post-insemination until the last fertile and FN: the numbers of fertile eggs laid after artificial insemination (AI). Bell-shaped indicate the normal distribution and the density curve was determined by standard normal distribution parameters N (µ, σ) where µ is the mean and σ is the standard deviation.

Figure 1. Duration of fertility trait in hens. (a,b) Histogram illustrates the distribution of DF trait within the population of egg-laying hens (n = 304). DN: the days post-insemination until the last fertile and FN: the numbers of fertile eggs laid after artificial insemination (AI). Bell-shaped indicate the normal distribution and the density curve was determined by standard normal distribution parameters N (µ, σ) where µ is the mean and σ is the standard deviation.
General characteristics of the egg-laying hen’s transcriptome. RNA-sequencing was used to analyze the lncRNA expression from 14 cDNA libraries constructed from the UVJ of long- and short-DF hens (no pooled samples, n = 7). On the average, approximately 45.7 and 42.3 million raw reads, were generated from the long-DF and the short-DF groups respectively (Fig. 3). After data pre-processing and quality control, clean reads from the long-DF group ranges from (~80–90%) and short-DF group (~79–92%) were successfully aligned to chicken genome: Galgal4.0, map ratio of reads in most samples was about 73% in both groups; read coverage and mapped unique reads in most samples was more than 70%, see Supplementary Table S1 for the details of data pre-processing and read mapping. The reads alignment to chicken genome indicated that the sequencing reads were of good quality and the sequencing depth was sufficient for further analysis of lncRNA between the two groups of hens.

Identification of lncRNAs in long-and short-DF egg-laying hens. In this study, we adopted a set of filtering criteria as described in the methods to identify putative lncRNAs from the alignments. Consequently, a total of 9977 lncRNAs were discovered in the assembled transcripts for which an average of 7038 and 6909 was expressed lncRNA transcripts in long- and short-DF groups respectively (Supplementary Table S2). Among the total lncRNA transcripts identified, 7401 and 2576 were known and unknown lncRNAs (Supplementary Table S2). Nevertheless, we observed that majority of lncRNAs were from the intergenic region (Supplementary Table S3). Next, the transcript lengths, exon number, and expression level between lncRNAs and mRNAs which

Table 1. Duration of fertility trait in hens. For all phenotypes, we examined multiple inseminations (n = 7).

| Phenotypes | DN' (days, Mean ± SD) | FN' (eggs, Mean ± SD) | Fertility rate' (%) |
|------------|------------------------|------------------------|---------------------|
| Long DF    | 18.71 ± 0.49           | 17.43 ± 0.53           | 94.57               |
| Short DF   | 7.14 ± 1.57            | 7.45 ± 1.40            | 23.74               |

*DN: the days post-insemination until the last fertile egg (up to about 40 wks of age); FN: the numbers of fertile eggs laid after artificial insemination (AI) (up to about 40 wks of age); % fertility = number of fertile eggs/number of total eggs produced or set.

Figure 2. The panel of a tissue section of the UVJ of egg-laying hen. (a,b) long-DF hens and (c,d) short-DF hens. A rounded-like shape known as sperm storage tubule (SST) was more embedded in the UVJs of long-DF hens compared with short-DF hens. Data are presented as mean ± SEM (n = 4). P < 0.01 (Student’s t-test). The red arrow indicates the SSTs, Scale bar = 100 μm.
generated from 14 individual chicken samples were compared and graphed, as shown in Fig. 4a–c. All the lncRNAs and all the mRNAs were regarded as two groups to compare their basic characteristics. The lncRNA transcript lengths were significantly shorter than those of mRNAs (Fig. 4a). A majority of lncRNAs had two or three exons, whereas mRNAs contained a broad range of exon numbers from two to greater than twenty (Fig. 4b). Additionally, the average expression level detected for lncRNAs (0.74 average FPKM) is significantly lower than that of mRNA (2.2 average FPKM) (Fig. 4c). Together, these results indicated that the lncRNAs in the UVJ of long- and short-DF exhibited relatively short length, low exon numbers, and low expression level.
Differentially expressed (DE) lncRNAs. After identification of the lncRNAs, the DE-lncRNAs were further analyzed based on the criteria of differential expression analysis. In total, the expression levels of 223 lncRNAs were significantly different between the long- and short-DF groups (q value < 0.05). Among these 223 lncRNAs, 81 were up-regulated and 142 were down-regulated in the long-DF group compared with the short-DF group (Fig. 5, Supplementary Table S4). The top 10 up-regulated and down-regulated lncRNAs in the long-DF hens were listed in Table 2. Among the downregulated ones, 85 lncRNAs were located in the intergenic region, 16 lncRNAs in bidirectional, 4 lncRNAs in introns (2 intronic-antisense and 2 intronic-sense) and 37 lncRNAs in exons (34 exonic-antisense and 3 exonic-sense). For the upregulated ones, 55 lncRNAs were mapped in intergenic position, 6 lncRNAs in bidirectional, 15 lncRNAs in exons (12 exonic-antisense and 3 exonic-sense), and 5 lncRNAs in introns (intronic-antisense) (Supplementary Table S4).

Figure 5. The differentially expressed (DE) lncRNAs between the long- and short-DF hens. The red and blue dots represent up- and down-regulated lncRNAs respectively. The grey dots represent the lncRNAs without significant differential expression.

| DE-lncRNA ID   | lncRNA type | length | log2FC | P value | Q value | Regulation type |
|---------------|-------------|--------|--------|---------|---------|----------------|
| MSTRG.2982.4  | intergenic  | 4713   | 10.55  | 0.000148| 0.032464| UP             |
| MSTRG.11712.1 | exonic_antisense | 460   | 9.38   | 1.7E-06 | 0.003118| UP             |
| MSTRG.838.2   | intergenic  | 5023   | 8.95   | 3.75E-31| 1.65E-27| UP             |
| MSTRG.10030.1 | intergenic  | 2695   | 7.83   | 1.13E-06| 0.000708| UP             |
| MSTRG.1096.2  | intergenic  | 1362   | 7.59   | 3.58E-05| 0.006568| UP             |
| MSTRG.29422.1 | exonic_antisense | 5615  | 7.56   | 1.79E-20| 2.72E-17| UP             |
| MSTRG.7526.1  | intronic_antisense | 911   | 7.45   | 7E-08   | 4.19E-05| UP             |
| MSTRG.22548.1 | intergenic  | 237    | 7.39   | 1.77E-06| 0.003118| UP             |
| MSTRG.5381.1  | exonic_antisense | 344   | 7.39   | 1.77E-06| 0.003118| UP             |
| MSTRG.6385.5  | intergenic  | 2632   | 7.01   | 7.13E-07| 0.000305| UP             |
| NONGGAT009046.2 | intergenic  | 2442   | -5.93  | 3.4E-11 | 4.99E-08| DOWN           |
| MSTRG.26352.3 | intergenic  | 901    | -6.04  | 3.61E-05| 0.012891| DOWN           |
| MSTRG.15771.3 | intergenic  | 7399   | -6.37  | 0.000107| 0.027697| DOWN           |
| MSTRG.23181.9 | exonic_antisense | 1229  | -6.43  | 0.000306| 0.027473| DOWN           |
| MSTRG.23244.2 | intergenic  | 7559   | -6.47  | 9.55E-05| 0.038805| DOWN           |
| MSTRG.15070.6 | intergenic  | 3272   | -7.03  | 2.08E-13| 6.09E-10| DOWN           |
| MSTRG.17455.2 | intergenic  | 1861   | -7.17  | 0.00018 | 0.019517| DOWN           |
| MSTRG.13256.1 | intergenic  | 210    | -7.60  | 8.93E-06| 0.011233| DOWN           |
| NONGGAT003930.2 | exonic_antisense | 1735  | -7.64  | 1.8E-08 | 1.35E-05| DOWN           |
| MSTRG.2059.3  | intergenic  | 6585   | -11.49 | 1.16E-21| 1.04E-17| DOWN           |

Table 2. List of top differentially expressed lncRNAs in the long-DF hens identified in the UVJ transcriptome (FDR-corrected P-value < 0.05).
DE-lncRNAs target genes and Gene Ontology (GO). To further investigate the potential biological functions and regulatory relationship label between DE-lncRNAs and target genes, target genes of the DE-lncRNAs were predicted with the Pearson correlation coefficient (PCC) method. Considering PCC > 0.9 and p-value < 0.05, eight DE-lncRNAs between long- and short-DF hens (5 downregulated and 3 upregulated ones) were identified and targeted more than 200 genes. The regulatory network of 8 DE-lncRNAs consisted 200 genes and 530 regulatory relationships (Supplementary Table S5). These results demonstrate a strong regulatory relationship between the 8 DE-lncRNAs and targets in DF trait regulation (Fig. 6). Among the 8 DE-lncRNAs, downregulated MSTRG.20950.12, MSTRG.17455.2, MSTRG.23167.4, and MSTRG.28487.2 targeted the majority of genes, such as SOX2, FAM20C, EVPL, PHGDH, and PRPS2. Upregulated MSTRG.14719.6 targeted set of genes encoding small nucleolar RNAs, C/D box (SNORD) and small nucleolar RNAs, H/ACA box (SNORA), such as SNORD38, SNORD14, SNORA32, SNORA77, SNORA66, SNORA55, and SNORA14. MSTRG.8138.1 regulated genes like CYP51A1 and 17.5. MSTRG.2982.4 also positively targeted genes such as IFIT5, and MSTRG.24725.1 positively regulate genes like HMGI-C (Supplementary Table S5). The most represented target gene (SOX2) of MSTRG.20950.12, MSTRG.17455.2, MSTRG.23167.4, MSTRG.28487.2, MSTRG.14719.6, MSTRG.2982.4, MSTRG.24725.1 and MSTRG.8138.1 were enriched in biological function such as reproductive structure development, reproductive system development, positive regulation of cell differentiation, and response to growth factor (Fig. 7a, Supplementary Table S6). In addition, the target genes of downregulated lncRNA (MSTRG.20950.12) such as CTR9 and DDIAS were enriched in the GO functions like response to cytokine, positive regulation of MAPK cascade, chromatin, and developmental process involved in reproduction (Supplementary Table S6). Furthermore, according to pathway enrichment analyses, the target genes of MSTRG.8138.1 (e.g., BRAF, and PRKAA2) were significantly enriched in the KEGG functions like mTOR signaling pathway (Fig. 7b, Supplementary Table S7). Together these results showed that these 8 key DE-lncRNAs may regulate DF trait, UVJ formation, and reproductive processes through their potential targets (Table 3).

Validation of DE-lncRNAs and target genes by qPCR. A total of 4 DE-lncRNAs and 5 target genes were selected and validated by qPCR including two downregulated lncRNAs (MSTRG.28487.2, and MSTRG.17455.2), two upregulated lncRNAs (MSTRG.8138.1 and MSTRG.2982.4) and five target genes (SOX2, FAM20C, PLCB1, BRAF and PRKAA2) between the long and short-DF hens. Among the selected genes, downregulated MSTRG.28487.2 and MSTRG.17455.2 expression fold change by qPCR were similar to the expression fold change obtained from the RNA-Seq (Fig. 8a,b). Similarly, upregulated MSTRG.8138.1 and MSTRG.2982.4 also appeared to be expressed at the same level when compared with the RNA-Seq data (Fig. 8c,d). Together, these data indicated that there was a strong agreement between the qPCR and RNA sequenced data. We found that target genes (SOX2, PLCB1, and FAM20C) were highly expressed in long-DF hens compared with the short-DF hens (Fig. 8e–g). High expression of BRAF and PRKAA2 were also detected in the short-DF hens compared with the long-DF hens (Fig. 8h,i). According to these findings, long- and short-DF hens significantly differed at P < 0.01 and P < 0.05 levels based on assuming unequal variances Student’s t-test.
Duration of fertility (DF) trait is a vital factor in poultry production. In the current study, the estimated DF-trait for the two composite characteristics (DN and FN) was approximately $(18.71 \pm 0.49; 7.14 \pm 1.57)$ and $(17.43 \pm 0.53; 7.43 \pm 1.40)$ in the long- and short-DF hens respectively (Table 1). During the reproductive phase, there was a significant difference in the DF traits between the two groups (Supplementary Table S8). In addition, a different morphological change was observed in their respective UVJ samples. Numerous SSTs were embedded in the long-DF groups compared with the short-DF, indicating that the mechanism regulating prolonged DF could be mostly due to the observed SST as revealed in the previous work conducted in broiler and turkey hens 30. Several studies associated with period of fertility of egg-laying hens have been reported31,32. However, the molecular mechanisms underlying prolonged-DF remain obscure.

During the reproductive season, we compared the lncRNA expression profiles in the long-DF hens with those of short-DF hens and identified a set of lncRNA genes associated with DF trait (Supplementary Table S2). In agreement with previous reports28, lncRNAs shared similar features, such as shorter in length, lower in exon number, and lower expression level than protein-coding transcripts. A total of 223 lncRNAs were differentially expressed (DE), of these, 81 lncRNAs were upregulated and 142 lncRNAs were downregulated in the long-DF groups, indicating that the majority of the DE-lncRNAs were downregulated throughout the reproductive phase. We anticipate that some of these DE-lncRNAs might play key roles in prolonging the DF in hens as shown in other studies19,33.

The DE-lncRNA-target prediction and functional interaction network revealed eight key DE-lncRNAs related to long-DF hens, including MSTRG.20950.12, MSTRG.17455.2, MSTRG.23167.4, MSTRG.28487.2, MSTRG.14719.6, MSTRG.2982.4, MSTRG.24725.1, and MSTRG.8138.1 (Supplementary Table S5). Many of the long-DF-enriched target genes were categorized into identical protein binding, response to growth factor, reproductive system development, positive regulation of cell differentiation, response to cytokine, carbohydrate binding, response to cytokine, positive regulation of MAPK cascade, peptidyl-lysine modification, chromatin, histone modification, regulation of MAPK cascade, MAPK cascade, positive regulation of protein modification process, regulation of cell cycle, and ossification.

A total of nine overlapped targets were potentially associated with long-DF hens (Table 3). These included CLDN1, SOX2, FAM20C, PLCB1, CTR9, DDIA1, 17.5, PRKAA2, and BRAF. For instance, SOX2 was expressed at high levels in the long-DF hens compared with the short-DF hens. SOX2 is a member of the sex-determining region (SRY-related) having analogous HMG domains for DNA binding and is highly involved in regulation of cell differentiation. SOX2 expression patterns in early chicken embryos, consistently mark neural primordial cells at various stages of development36. Mutations in SOX2 have previously been associated with male genital tract abnormalities37. Another interesting gene, named family with sequence similarity 20, member C (FAM20C), was also highly expressed in the long-DF hens. FAM20C is a physiological Golgi casein kinase that phosphorylates multiple secreted proteins38. The molecular activity of FAM20C has been implicated in mammalian reproduction39. In this study, the observation of significant high expression of FAM20C implied that it may...
Table 3. Summary of the eight key differentially expressed long non-coding RNA (lncRNA) associated with duration of fertility trait in the egg-laying hens.

| DE-lncRNA   | Category | Function                                      | Targets                                                                 |
|-------------|----------|-----------------------------------------------|-------------------------------------------------------------------------|
| MSTRG.23167.4 | MF       | Identical protein binding                     | AADAT, TOX3, VWF, CEBPB, HNF1A, SHMT1, CDL1N1, ACFP                      |
| MSTRG.28487.2, MSTRG.17455.2, MSTRG.8138.1, MSTRG.2982.4 | BP       | Response to growth factor                     | SOX2, NOS1, FAM20C, BMPER, NPNT                                          |
| MSTRG.28487.2, MSTRG.8138.1 | BP       | Reproductive system development               | SOX2, BIRC6, HNF1A, VWF, CEBPB                                          |
| MSTRG.14719.6, MSTRG.28487.2, MSTRG.17455.2, MSTRG.8138.1, MSTRG.28487.2, MSTRG.2982.4 | BP       | Positive regulation of cell differentiation   | SOX2, BNP2, ZNF335, FAM20C, CEBPB, NPNT, PLCB1                       |
| MSTRG.20950.12, MSTRG.14719.6, MSTRG.24725.1 | BP       | Cellular response to cytokine stimulus        | CTR9, PLCB1, DDIA5, NPNT, PID1, FASN                                      |
| MSTRG.24725.1, MSTRG.28487.2, MSTRG.8138.1 | BP       | Developmental process involved in reproduction | VWF, CEBPB, BIRC6, HNF1A, DDIA5, SOX2                                   |
| MSTRG.28487.2, MSTRG.17455.2, MSTRG.8138.1, MSTRG.28487.2 | BP       | Regulation of osteoblast differentiation      | NPNT, CEBPB, SOX2, FAM20C, BMPER                                        |
| MSTRG.28487.2, MSTRG.8138.1 | BP       | Response to acid chemical                     | LGMN, SOX2, PID1, CEBPB                                                 |
| MSTRG.28487.2, MSTRG.17455.2, MSTRG.8138.1, MSTRG.28487.2 | BP       | Osteoblast differentiation                    | FAM20C, SOX2, NPNT, CEBPB                                               |
| MSTRG.8138.1 | MF       | Carbohydrate-binding                          | 17.5, CD93                                                              |
| MSTRG.20950.12, MSTRG.14719.6, MSTRG.24725.1 | BP       | Response to cytokine                         | CTR9, FASN, NPNT, PID1, DDIA5, PLCB1                                     |
| MSTRG.14719.6, MSTRG.28478.2, MSTRG.8138.1 | BP       | Positive regulation of MAPK cascade          | PLCB1, NPNT, BMPER, BNP2, SOX2                                         |
| MSTRG.20950.12 | BP       | Peptidyl-lysine modification                 | NOS1, YEATS2, ZNF335, CTR9                                              |
| MSTRG.20950.12, MSTRG.14719.6 | CC       | Chromatin                                    | CTR9, CEBPB, PLCB1, H3F3B                                               |
| MSTRG.20950.12 | BP       | Chromatin organization                       | ZNF335, CTR9, H3F3B, YEATS2, NOS1, HNF1A                                 |
| MSTRG.20950.12 | BP       | Chromatin modification                       | HNF1A, NOS1, CTR9, ZNF335, YEATS2                                       |
| MSTRG.20950.12 | BP       | Histone modification                         | YEATS2, NOS1, CTR9, ZNF335                                              |
| MSTRG.14719.6, MSTRG.24725.1, MSTRG.28487.2, MSTRG.8138.1 | BP       | Regulation of MAPK cascade                   | BNP2, SOX2, BMPER, NPNT, PLCB1                                          |
| MSTRG.14719.6, MSTRG.24725.1, MSTRG.28487.2, MSTRG.8138.1 | BP       | MAPK cascade                                 | BMPER, NPNT, PLCB1, SOX2, BNP2                                          |
| MSTRG.14719.6, MSTRG.24725.1, MSTRG.28487.2, MSTRG.8138.1 | BP       | Positive regulation of protein modification process | NOS1, SOX2, BNP2, NPNT, PLCB1, CTR9, BMPER                              |
| MSTRG.14719.6, MSTRG.24725.1, MSTRG.28487.2, MSTRG.8138.1 | BP       | Regulation of cell cycle                     | PID1, E2FS, PLCB1, DDIA5, BIRC6, SOX2                                   |
| MSTRG.28487.2, MSTRG.17455.2, MSTRG.8138.1, MSTRG.2982.4 | BP       | Osmification                                 | NPNT, CEBPB, FAM20C, SOX2                                               |
| MSTRG.8138.1 | KEGG     | mTOR signaling pathway                       | PRKAA2, BRAF                                                            |

Contribute to the DF regulation of long-DF hens. The immunological response in the UVJ of mated hens requires up-regulated phospholipase C eta 1 (PLCH1), an immune-modulatory gene in the previous studies. Likewise in our data, the high expression level of phospholipase C beta 1 (PLCB1) was observed in the long-DF hens which suggest that PLCB1 has a potential role in DF regulation.

One important signaling pathway associated with DF trait is mTOR signaling pathway and the target genes involved were over-represented in KEGG database, including mTOR signaling pathway, Vascular smooth muscle contraction, insulin signaling pathway, FoxO signaling pathway, Regulation of actin cytoskeleton, and Focal adhesion pathway (Supplementary Table S7). The elevated expression of BRAF and PRKAA2 genes via mTOR signaling pathway was seen in the short-DF hens. The mTOR signaling pathways are critical regulators of ovarian function including quiescence, activation, and survival of primordial follicles (PFs), granulosa cell (GC), proliferation and differentiation, and meiotic maturation of oocytes. mTOR signaling pathway has been implicated in fertile egg production and female fertility in mice, suggesting that BRAF and PRKAA2 genes via mTOR signaling pathways may be involved in regulating the DF trait.

In the lncRNA-target regulatory network, several set of genes encoding small nucleolar RNAs, such as SNORD38, SNORD14, SNORA32, SNORA77, SNORA66, SNORA55, and SNORA14 also exhibited a strong correlation with the upregulated MSTRG.14719.6, suggesting the important role of MSTRG.14719.6 in the post-transcriptional modification of ribosomal RNAs (rRNAs) which are essential for sperm viability and male fertility. Additional studies are needed to determine the precise function of these small nucleolar RNAs in the DF regulation.

Furthermore, in the regulatory network, MSTRG.2982.4 and MSTRG.24725.1 targeted cytokines related genes such as Mx (Interferon-induced GTP-binding protein Mx) and IFIT5 respectively. Previous studies have that ovarian functions are regulated by cytokines. We speculate MSTRG.2982.4 and MSTRG.24725.1 may participate in prolonging the DF, likely via their potential target genes. MSTRG.8138.1 regulated lectin-like proteins (17.5). It has been reported that the sperm-oocyte binding in the mouse occurred as a result of the existence of lectin-like proteins on the sperm plasma membrane which binds to zona pellucida glycoproteins. MSTRG.8138.1 also played an important role in gene-gene networks that are involved in reproductive system development and positive regulation of cell differentiation, suggesting the crucial roles of MSTRG.8138.1 in the DF regulation.

The enrichment of target in cellular response to cytokine stimulus indicates that it may contribute to the immune function within the oviduct of hens during reproductive seasons. In addition to these genes, several genes interacted with another in the gene pathways. Of those, 20 additional genes were found to have functions for DF trait regulation, such as AADAT, TOX3, VWF, CEBPB, SHMT1, ACPP, NOS1, BMPER, NPNT, BNP2, ZNF335, BIRC6, HNF1A, PID1, FASN, LGMN, CD93, YEATS2, H3F3B, and E2FS.
Conclusions

In summary, based on the RNA-Seq and bioinformatics analysis results, a list of DE-lncRNAs related to the cellular response to cytokine, reproductive structure development, regulation of protein modification, carbohydrate binding, co-factor binding, chromatin organization and modification, response to growth factors, cytokine stimulus and immune pathways were identified in the UVJ of long and short-DF hens. Eight key DE-lncRNAs were identified to have the most probable role in prolonging the DF in egg-laying hens, thus, they are novel lncRNAs that are related to DF in hens. These results provided the starting point for studies aimed at understanding the molecular mechanism of the DF trait in egg-type chickens.

Material and Methods

Egg-laying hens. A total of 304 egg-type chickens obtained from the poultry farm of Huadu Yukou Poultry Industry Co. Ltd, Beijing, China were raised in individual cages, kept in identical light/dark cycles and fed standard diets ad libitum from 25 weeks until the end of the experiment aiming at study their duration of fertility. All hens were artificially inseminated once with $2 \times 10^8$ pooled sperms ejaculates collected from viable rooster flocks. Eggs were collected and marked daily from day 2–20 after artificial insemination, all experiment was executed in three replicates and lasted 60 days. The number of egg per hen over the period was recorded and the fertilized eggs were examined by candling on day 10 of incubation (dead embryos were considered as fertile). DF trait was expressed in terms of DN (the number of days post-insemination until last fertile egg) and FN (the number of fertile eggs after a single AI) and hens with DN or FN traits $\geq 18$ were selected, coded as long and DN or FN $< 10$ as short-DF groups. Seven long-DF and seven short-DF hens were selected for sampling UVJ. All selected hens were anesthetized with sodium pentobarbital (20 mg/kg) administered intraperitoneally, and the narrow band known as UVJ located at the cranial anterior end of the vagina were used as a sample for RNA collection. In other to characterize the UVJ structure of long- and short-DF hens. The sample size for tissues analysis was four UVJs selected from 4 different individuals of long- and short-DF groups respectively. The UVJ samples were formalin-fixed for 48h and paraffin-embedded. The 100 μm of the tissues sectioned were stained with hematoxylin and counterstained with eosin (hematoxylin for 1 min and 1% eosin for 10 sec) for observation under a light microscope (OLYMPUS; TH4–200; Tokyo, Japan).

Figure 8. Validation of differentially expressed lncRNAs by qPCR. (a,b) Fold change expression of down-regulated lncRNAs in long DF hens compared with the short DF hens; (c,d) Fold change expression of up-regulated lncRNAs in long DF hens compared with short DF hens. Data are reported for three technical replicates in all biological samples. (e–i) The expression levels of lncRNA target genes between long and short DF hens. Each histogram represents the level of the target gene relative expression level. Error bars indicate standard error of the mean (SEM). Asterisks symbolized P-value significance *p ≤ 0.05, and **p ≤ 0.01, based on assuming unequal variances Student’s t-test.
RNA isolation and sequencing procedures. Total RNA was isolated from the UVJ of seven long-DF and seven short-DF hens using the RNeasy mini kit (Qiagen, Germany). The quality and concentration of RNA were analyzed by NanoDrop ND2000 (Thermo Fisher Scientific, Waltham, MA, USA) spectrophotometer and gel electrophoresis. Fourteen cDNA libraries were constructed using the TruSeq™ Stranded Total RNA Sample Preparation kit (Illumina, USA). The library construction was performed by a commercial company (Shanghai Biotechnology Corporation (SBC), China. The datasets (14 files) supporting the conclusions of this article is available in the Gene Expression Omnibus (GEO) database NCBI under the accession number GSE101163 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE101163).

Transcriptome analysis for IncRNA data. Mapping of all clean reads was carried on the chicken reference genome (assembly Gallus gallus_4.0) using TopHat (version:2.0.9) with parameters set to default values58. Unmapped reads were trimmed and remapped, additionally, Gene Transfer Format (GTF) of the Ensembl gene annotation was included in the read mapping as previously described by57. Next, RNA-Seq alignments for all sample was assembled using String Tie (version: 1.3.0) and Gifcompare (version: 0.9.8) was used to evaluate the accuracy of the assembled transcripts (i.e., putative transcripts containing both coding and noncoding transcripts) and compare the assembly with known transcripts (putative transcripts for each sample assembly against a set of combined gene annotations), to extract assemblies that fully match known and unknown annotations57,28. Prior to that, low-quality assemblies were removed based on the Fragments Per Kilobase of exon per Million fragments mapped (FPKM) threshold, and multi-exon transcripts were retained for downstream processing. The FPKM threshold for classifying complete and partial transcripts in our experiment was established according to57. Furthermore, transcripts with short read length <200 bp, <2 exons including those with maximum putative ORFs <300 bp were filtered out. Next, protein-coding transcripts were identified and removed with the software programme Coding-Non-Coding Index (CNCI > 0)57 and Coding Potential Calculator (CPC > 0)59. Similarly, Pfam (http://pfam.xfam.org) was used to remove transcripts with significant protein domain. From the remaining transcripts, a putative IncRNA is defined as novel IncRNAs if its Pfam hits returned insignificantly. The analysis protocols are described step by step in Supplementary Fig. S1. Afterward, the novel IncRNAs were classified according to60 and characterized based on transcription length, exon number and expression level between IncRNA and mRNA transcripts (RefSeq transcripts, NCBI). Furthermore, differentially expressed IncRNAs (DE-IncRNAs) in the comparison groups of long-DF versus short-DF were identified using the EdgeR package (freely available from the Bioconductor web site [http://bioconductor.org]61. Only the IncRNAs with the criteria of log2FC (fold change) >2 and p-value ≤ 0.01 and FDR ≤ 0.05 were identified as DE-IncRNAs.

Prediction and functional annotation of DE-IncRNAs target genes. To further predict the DE-IncRNA target genes, the Pearson correlation coefficient (PCC) was calculated to evaluate the co-expression relationships between DE-IncRNAs and target genes. The co-expression pairs with PCC > 0.9 and p-value < 0.05 were selected for the construction of the regulatory network, which was visualized by Cytoscape 3.3.0 (free download: available at http://www.cytoscape.org/). Functional enrichment analysis and pathway analysis of DE-IncRNAs target genes were analyzed using Go (Gene Ontology) function in the Database for Annotation, Visualization and Integrated Discovery (DAVID, from the National Institute of Allergy and Infectious Diseases (NIAID), NIH (https://david.ncifcrf.gov/). and Kyoto Encyclopedia of Genes and Genomes (KEGG) (http://www.genome.ad.jp/kegg/) databases62,63. Quantitative real-time PCR (qPCR) validation of DE-IncRNAs and target genes. The RNA sequencing data were validated by qPCR, using a standard EasyScript™ one-step gDNA Removal, cDNA Synthesis SuperMix, for cDNA production and SYBR Green real-time PCR (Toyobo co., Ltd, Osaka, Japan) following manufacturer’s instructions. Two up-regulated and two down-regulated DE-IncRNAs and five target genes were randomly selected to quantify the expression levels. Quantitative PCR (qPCR) was then performed in Bio-Rad thermal cycler, CFX-384, real-time system. Gene expression in each sample was normalized to beta-actin expression. The qPCR amplifications were conducted using an independent set of seven biological replicates and three technical replicates per sample. Relative quantification of IncRNAs and target genes expression was evaluated by the 2(−ΔΔCt) methods64,65. Primers sequences of IncRNAs and target genes for qPCR are listed in Supplementary Table S9.

Ethics approval and consent to participate. The dissection of the UVJ sample and collection was conducted according to the guidelines established for the Care and Use of Laboratory Animals by the Ethics Committee of Huazhong Agricultural University, P. R. China, and Standing Committee of Hubei People's Congress (No. 5) approved by the Animal Care Committee of Hubei Province, P.R. China.

Data Availability
All replicates sequencing reads used in this study have been submitted to Gene Expression Omnibus (GEO) under the accession number GSE101163.

References
1. Brillard, J. Practical aspects of fertility in poultry. World's Poultry Science Journal 59, 441–446 (2003).
2. 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 (2008).
3. Bakst, M. R. & Vinyard, B. T. Oviducal sperm storage in turkeys: Spatial distribution of sperm within the uterovaginal junction sperm-storage tubules. J Exp Zool 292, 206–209 (2002).
4. Liu, G., Zhu, J., Wang, Z., Jiang, X. & Dalalla, M. Analysis of sperm storage ability using duration of fertility in hens. British poultry science 49, 770–775 (2008).
5. Cheng, Y. S. et al. Selection responses for the number of fertile eggs of the Brown Tsaiya duck (Anas platyrhynchos) after a single artificial insemination with pooled Muscovy (Cairina moschata) semen. *Genetics Selection Evolution* **34**, 597 (2002).

6. Brun, J.-M., Mialon, M. M., Sellier, N., Brillard, J.-P. & Rouvier, R. Inheritance of duration of fertility in female common ducks (Anas platyrhynchos) inseminated in pure breeding or in inter-generic crossbreeding with Muscovy drakes (Cairina moschata). *Animal* **6**, 1731–1737 (2012).

7. Cheng, Y.-S. et al. Eleven generations of selection for the duration of fertility in the intergeneric crossbreeding of ducks. *Genetics Selection Evolution* **41**, 32 (2009).

8. Holm, L., Riddrstråle, Y. & Knutsson, P.-G. Localisation of carbonic anhydrase in the sperm storing regions of the domestic hen oviduct. *Cells Tissues Organs* **156**, 253–260 (1996).

9. Daryabari, H. et al. Reproductive performance and oviductal expression of avdin and avdin-related protein-2 in young and old broiler breeder hens orally exposed to supplementary biotin. *Poultry science* **93**, 2289–2293 (2014).

10. Bakst, M. Physiology and Endocrinology Symposium: Role of the oviduct in maintaining sustained fertility in hens. *Journal of animal science* **89**, 1323–1329 (2011).

11. Zaniboni, L. & Bakst, M. Localization of aquaporins in the sperm storage tubules in the turkey oviduct. *Poultry science* **83**, 1209–1212 (2004).

12. Bakst, M. & Akuffo, V. Alkaline phosphatase reactivity in the vagina and uterotubal junction sperm-storage tubules of turkeys in egg production: implications for sperm storage. *British poultry science* **48**, 515–518 (2007).

13. Yoshimura, Y., Koike, K. & Okamoto, T. Immunolocalization of progesterone and estrogen receptors in the sperm storage tubules of laying and diethylstilbestrol-injected immature hens. *Poultry science* **79**, 94–98 (2000).

14. Das, S. C., Isobe, N., Nishibori, M. & Yoshimura, Y. Expression of transforming growth factor-β isoforms and their receptors in utero-vaginal junction of hen oviduct in presence or absence of resident sperm with reference to sperm storage. *Reproduction* **132**, 781–790 (2006).

15. Gu, L., Sun, C., Gong, Y., Yu, M. & Li, S. Novel copy number variation map of the TGFβ3 gene is associated with TGFβ3 gene expression and duration of fertility traits in hens. *PloS one* **12**, e0173697 (2017).

16. Bakst, M. R., Wishart, G. & Brillard, J.-P. Oviducal sperm selection, transport, and storage in poultry. *Poul. Sci. Rev* **5**, 117–143 (1994).

17. Bakst, M. R. & Brillard, J. P. Role of the oviduct in maintaining sustained fertility in hens. *J Dairy Sci* **93**, 504–504 (2010).

18. Gomes, A. Q., Nolasco, S. & Soares, H. Non-coding RNAs: multi-tasking molecules in the cell. *International journal of molecular sciences* **14**, 16010–16039 (2013).

19. Ken, K. et al. Critical roles of long noncoding RNAs in Drosophila spermatogenesis. *Genome Res* **26**, 1233–1244 (2016).

20. Yoneda, R. et al. A genomic region transcribed into a long non-coding RNA interacts with the Prss42/Tessp-2 promoter in spermatocytes during mouse spermatogenesis, and its flanking sequences can function as enhancers. *Molecular reproduction and development* **83**, 541–557 (2016).

21. Ran, M. et al. Systematic identification of long noncoding RNAs in immature and mature porcine testes. *Biol Reprod* **94**(77), 71–79 (2016).

22. Lin, X. et al. Expression dynamics, relationships, and transcriptional regulations of diverse transcripts in mouse spermatogenic cells. *Rna Biol* **13**, 1011–1024 (2016).

23. Kurishara, M., Shiraiishi, A., Satake, H. & Kimura, A. P. A conserved noncoding sequence can function as a spermatocyte-specific enhancer and a bidirectional promoter for a ubiquitously expressed gene and a testis-specific long noncoding RNA. *Journal of molecular biology* **426**, 3069–3093 (2013).

24. Wang, Y.-N. et al. Annotating long integenic non-coding RNAs under artificial selection during chicken domestication. *BMC evolutionary biology* **17**, 192 (2017).

25. Li, T. et al. Identification of long non-protein coding RNAs in chicken skeletal muscle using next generation sequencing. *Genomics* **99**, 292–298 (2012).

26. Li, Z. et al. Integrated Analysis of Long Non-coding RNAs (LncRNAs) and mRNA Expression profiles reveals the potential role of LncRNAs in skeletal muscle development of the chicken. *Front Physiol* **7** (2016).

27. Muret, K. et al. Long noncoding RNA repertoire in chicken liver and adipose tissue. *Genetics Selection Evolution* **49**, 6 (2017).

28. Liu, Y. et al. Analyses of Long Non-Coding RNA and mRNA profiling using RNA sequencing in chicken testis with extreme sperm motility. *Sci Rep-Uk* 7, 9055 (2017).

29. Degen, A. & Hawes, R. Fertility in the domestic hen following the surgical removal of the utero-vaginal junction. *Poultry science* **51**, 464–470 (1972).

30. Bakst, M. et al. Comparisons of sperm storage tubule distribution and number in 4 strains of mature broiler breeders and in turkey hens before and after the onset of photostimulation. *Poultry science* **89**, 986–992 (2010).

31. Sasanami, T., Matsuizaki, M., Mizushima, S. & Hiyam, G. Sperm Storage in the Female Reproductive Tract in Birds. *Journal of Reproduction and Development* **59**, 334–338 (2013).

32. Birkhead, T. & Møller, A. Numbers and size of sperm storage tubules and the duration of sperm storage in birds: a comparative study. *Trends in evolutionary biology* **17**, 192 (2017).

33. Li, L. et al. A long non-coding RNA interacts with Gfra1 and maintains survival of mouse spermatogonial stem cells. *Cell death & disease* **7**, e2140 (2016).

34. Kamachi, Y., Uchikawa, M. & Kondoh, H. Pairing SOX off: with partners in the regulation of embryonic development. *Trends in Genetics* **16**, 182–187 (2000).

35. Peavy, L. H. & Lovell-Badge, R. Sox genes find their feet. *Current opinion in genetics & development* **7**, 338–344 (1997).

36. Uchikawa, M., Ishida, Y., Takemoto, T., Kamachi, Y. & Kondoh, H. Functional analysis of chicken Sox2 enhancers highlights an array of diverse regulatory elements that are conserved in mammals. *Developmental cell* **4**, 509–519 (2003).

37. Kelberman, D. et al. Mutations within Sox2/SOX2 are associated with abnormalities in the hypothalamo-pituitary-gonadal axis in mice and humans. *Proceedings of the National Academy of Sciences* **116**, 2442–2455 (2016).

38. Tibaldi, E. et al. Golgi apparatus casein kinase phosphorylates bioactive Ser-6 of bone morphogenetic protein 15 and growth and differentiation factor 9. *FEBS letters* **584**, 801–805 (2010).

39. Tagliabracci, V. S. et al. Dynamic regulation of FGFR3 by Fam20C phosphorylation, GalNAc-T3 glycosylation, and furin proteolysis. *Proceedings of the National Academy of Sciences* **111**, 5520–5525 (2014).

40. Atikuzzaman, M., Bhui, R. M., Fogelholm, J., Wright, D. & Rodriguez-Martinez, H. Matting induces the expression of immune- and pH regulatory genes in the utero-vaginal junction containing mucosal sperm-storage tubuli of hens. *Reproduction* **150**, 473–483 (2015).

41. Yoshimura, Y. Significance of local immunity in hen reproductive organs. *Animal Science Journal* **75**, 183–191 (2004).

42. Makkere, A., Goel, M. M. & Mahdi, A. A. P13K/PTEN/Akt and TSC/mTOR signaling pathways, ovarian dysfunction, and infertility: an update. *Journal of molecular endocrinology* **53**, R103–R118 (2014).

43. Jiang, Z. et al. Survivin is essential for fertile egg production and female fertility in mice. *Cell death & disease* **5**, e1154 (2014).

44. Brohi, R. D. & Hsu, L.-J. Posttranslational Modifications in Spermatozoids and Effects on Male Fertility and Sperm Viability. *OMICS: A Journal of Integrative Biology* **21**, 245–256 (2017).

45. Dupuis-Sandoval, F., Poirier, M. & Scott, M. S. The emerging landscape of small nucleolar RNAs in cell biology. *Wiley Interdisciplinary Reviews: RNA* **6**, 381–397 (2015).
46. Sirotnik, A. V. Cytokines: signalling molecules controlling ovarian functions. The international journal of biochemistry & cell biology 43, 857–861 (2011).
47. Peska, S., Krause, C. D. & Walter, M. R. Interferons, interferon-like cytokines, and their receptors. Immunological reviews 202, 8–32 (2004).
48. Hedon, B., Bringer, J. & Mares, P. Fertility and sterility: a current overview. Proceedings of the 15th World Congress on Fertility and Sterility Montpellier France 17–22 September 1995 (1995).
49. Clark, G. F. Molecular models for mouse sperm-oocyte binding. Glycobiology 21, 3–5 (2011).
50. Radhakrishnan, P. & Fedorka, K. M. Immune activation decreases sperm viability in both sexes and influences female sperm storage. Proceedings of the Royal Society of London B: Biological Sciences 279, 3577–3583 (2012).
51. McNamara, K., Lieshout, E. & Simmons, L. Females suffer a reduction in the viability of stored sperm following an immune challenge. Journal of evolutionary biology 27, 133–140 (2014).
52. Brillard, J. & Antoine, H. Storage of sperm in the uterovaginal junction and its incidence on the numbers of spermatozoa present in the perivitelline layer of hens’ eggs. British poultry science 31, 635–644 (1990).
53. Goerzen, P., Julsrud, W. & Robinson, F. Duration of fertility in ad libitum and feed-restricted caged broiler breeders. Poultry science 75, 962–965 (1996).
54. Silber, S., Kagawa, N., Kowayama, M. & Gosden, R. Duration of fertility after fresh and frozen ovary transplantation. Fertility and sterility 94, 2191–2196 (2010).
55. Wishart, G. Regulation of the length of the fertile period in the domestic fowl by numbers of oviducal spermatozoa, as reflected by those trapped in laid eggs. Journal of reproduction and fertility 80, 493–498 (1987).
56. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25, 1105–1111 (2009).
57. Sun, L. et al. Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. Nucleic acids research 41, e166–e166 (2013).
58. Knauss, J. L. & Sun, T. Regulatory mechanisms of long noncoding RNAs in vertebrate central nervous system development and function. Neuroscience 235, 200–214 (2013).
59. 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 (2010).
60. Dennis, G. et al. DAVID: database for annotation, visualization, and integrated discovery. Genome biology 4, R60 (2003).
61. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature protocols 4, 44–57 (2009).
62. 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 (2010).
63. Dennis, G. et al. DAVID: database for annotation, visualization, and integrated discovery. Genome biology 4, R60 (2003).
64. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2−ΔΔCT method. methods 25, 402–408 (2001).
65. Schmittgen, T. D. & Livak, K. J. Analyzing real-time PCR data by the comparative CT method. Nature protocols 3, 1101 (2008).

Acknowledgements
Authors are thankful to Professor Chunyan Mou for her invaluable suggestions. This project was supported by the National Natural Science Foundation of China (No. 31372301), and the Fundamental Research Funds for the Central Universities (No. 0900202930). 

Author Contributions
Conceptualization: S.L. and S.Z. Data Curation: A.A.A. and L.G. Formal analysis: A.A.A., L.G. and X.D. Funding acquisition: S.L. Supervision: S.L. and S.Z. Project administration: A.A.A. Writing-original draft: A.A.A. Writing-review & editing: A.A.A., C.C.N., S.Z. and S.L.

Additional Information
Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-31301-z.
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.