Briefings in Bioinformatics, 00(00), 2021, 1–7

doi: 10.1093/bib/bbaa432
Database Review

Downloaded from https://academic.oup.com/bib/advance-article/doi/10.1093/bib/bbaa432/6120284 by University of California, San Francisco user on 26 May 2021

FifBase: a comprehensive fertility-associated indicators factor database for domestic animals

Hao Li†, Junyao Hou†, Ziyu Chen, Jingyu Zeng, Yu Ni, Yayu Li, Xia Xiao, Yaqi Zhou, Ning Zhang, Deyu Long, Hongfei Liu, Luyu Yang, Xinyue Bai, Qun Li, Tongtong Li, Dongxue Che, Leijie Li, Xiaodan Wang, Peng Zhang and Mingzhi Liao

Corresponding author: Mingzhi Liao, College of Life Sciences, Northwest A&F University, Yangling, Shaanxi 712100, China. Tel/Fax: +86 029 87092262; E-mail: liaomingzhi83@163.com
†These authors contributed equally to this work.

Hao Li is a master student at College of Life Sciences, Northwest A&F University, Yangling, Shaanxi, China. His research interests include bioinformatics, reproduction biology and developmental biology.
Junyao Hou is a master student at College of Life Sciences, Northwest A&F University, Yangling, Shaanxi, China. Her research interests include bioinformatics, systems biology and developmental biology.
Ziyu Chen is a master student at College of Life Sciences, Northwest A&F University, Yangling, Shaanxi, China. Her research interests include bioinformatics, reproduction biology and developmental biology.
Jingyu Zeng is an undergraduate student at College of Life Sciences, Northwest A&F University, Yangling, Shaanxi, China. His research interests include bioinformatics, systems biology and developmental biology.
Yu Ni is an undergraduate student at College of Information Engineering, Northwest A&F University, Yangling, Shaanxi, China. His research interests include bioinformatics, systems biology and developmental biology.
Yayu Li is an undergraduate student at College of Information Engineering, Northwest A&F University, Yangling, Shaanxi, China. Her research interests include bioinformatics, systems biology and developmental biology.
Xia Xiao is a master student at College of Life Sciences, Northwest A&F University, Yangling, Shaanxi, China. Her research interests include bioinformatics, systems biology and developmental biology.
Yaqi Zhou is a master student at College of Life Sciences, Northwest A&F University, Yangling, Shaanxi, China. Her research interests include bioinformatics, systems biology and developmental biology.
Ning Zhang is a PhD student at College of Life Sciences, Northwest A&F University, Yangling, Shaanxi, China. His research interests include bioinformatics, systems biology and developmental biology.
Deyu Long is a PhD student at College of Life Sciences, Northwest A&F University, Yangling, Shaanxi, China. His research interests include bioinformatics, systems biology and developmental biology.
Hongfei Liu is a master student at College of Animal Science and Technology, Northwest A&F University, Yangling, Shaanxi, China. His research interests include bioinformatics, systems biology and developmental biology.
Luyu Yang is a master student at College of Life Sciences, Northwest A&F University, Yangling, Shaanxi, China. Her research interests include bioinformatics, reproduction biology and developmental biology.
Xinyue Bai is a master student at School of Life Science and Technology, ShanghaiTech University, Shanghai, China. Her research interests include bioinformatics.
Qun Li is a PhD student at Institutes of Biomedical Sciences, Fudan University, Shanghai, China. His research interests include bioinformatics, reproduction biology and developmental biology.
Tongtong Li is a master student at College of Life Sciences, Northwest A&F University, Yangling, Shaanxi, China. Her research interests include bioinformatics, systems biology and developmental biology.
Dongxue Che is a master student at College of Life Sciences, Northwest A&F University, Yangling, Shaanxi, China. Her research interests include bioinformatics, systems biology and developmental biology.
Leijie Li is a PhD student at Department of Bioinformatics and Biostatistics, SJTU Yale Joint Center Biostatistics, Shanghai Jiao Tong University, Shanghai, China. His research interests include bioinformatics, systems biology and biostatistics.
Xiaodan Wang is a PhD student at School of Life Sciences, Tsinghua University, Beijing, China. Her research interests include bioinformatics, systems biology and developmental biology.
Peng Zhang is a master student at College of Life Sciences, Northwest A&F University, Yangling, Shaanxi, China. His research interests include bioinformatics, systems biology and developmental biology.
Mingzhi Liao is an associate professor at College of Life Sciences, Northwest A&F University, Yangling, Shaanxi, China. His research interests include bioinformatics, systems biology and developmental biology.
Submitted: 3 September 2020; **Received (in revised form):** 13 December 2020

© The Author(s) 2021. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oup.com

Abstract

Fertility refers to the ability of animals to maintain reproductive function and give birth to offspring, which is an important indicator to measure the productivity of animals. Fertility is affected by many factors, among which environmental factors may also play key roles. During the past years, substantial research studies have been conducted to detect the factors related to fecundity, including genetic factors and environmental factors. However, the identified genes associated with fertility from countless previous studies are randomly dispersed in the literature, whereas some other novel fertility-related genes are needed to detect from omics-based datasets. Here, we constructed a fertility index factor database FifBase based on manually curated published literature and RNA-Seq datasets. During the construction of the literature group, we obtained 3301 articles related to fecundity for 13 species from PubMed, involving 2823 genes, which are related to 75 fecundity indicators or 47 environmental factors. Eventually, 1558 genes associated with fertility were filtered in 10 species, of which 1088 and 470 were from RNA-Seq datasets and text mining data, respectively, involving 2910 fertility-gene pairs and 58 fertility-environmental factors. All these data were cataloged into FifBase (http://www.nwsuaf1mz.com/FifBase/), where the fertility-related factor information, including gene annotation and environmental factors, can be browsed, retrieved and downloaded with the user-friendly interface.

Key words: biological database; domestic animal; fertility; genetic factors; environmental factors

Downloaded from https://academic.oup.com/bib/advance-article/doi/10.1093/bib/bbaa432/6120284 by University of California, San Francisco user on 26 May 2021

Introduction

Fertility characterizes the ability of animals to maintain normal reproductive function and produce offspring [1]. The level of fecundity directly affects the economic benefits of animal husbandry [2–5], so the genetic basis of variation about fertility between individuals is of great interest in mammals, particularly in humans and livestock [6]. Furthermore, infertility and subfertility also impact the success of embryo transfer in animals and humans [7, 8]. In agricultural production, the transition from late pregnancy to early lactation in domestic animals has important implications for animal health, milk production and reproductive performance, and consequently the profitability of poultry industry [9, 10]. Therefore, detecting the genes and environmental factors that play roles in reproduction will facilitate to accelerate genomic selection for improving fertility of livestock and poultry and reducing the high rates of early embryonic mortality [6, 11].

Though tremendous efforts have been devoted to elucidate factors that influence fertility, there are evidences about that fertility in domestic animals has declined significantly in recent decades [12, 13]. The causes of this decline are multifactorial [12, 13], including nutrition [14], food intake [15], temperature [16], drug [17], illumination [18] reproductive management level [19, 20], etc. These evidences indicate environmental factors increasingly become a major obstacle to animal reproduction. Beyond them, there are growing evidences that genes also play key roles in animal fertility [21–24]. With the development of RNA-Seq technologies, it becomes possible to detect fertility-related genes on a genome-wide scale. Numerous studies have been performed by RNA-Seq to identify gene associated with fertility [15, 25–28].

Based on the above basis, researchers have developed several databases for the storage of fertility data. For example, DevOmics (http://www.devomics.org) integrated epigenetic and transcriptional data on human and mouse germ cells. SpermatogenesisOnline 1.0 [29] collected and predicted genes that have been reported to be involved in spermatogenesis in 37 species. Gametogenesis molecule online (GMO) [30] detected the dynamic process of gametogenesis from the perspective of systems biology based on protein–protein interaction networks and functional analysis and provided a computational perspective and frame to the analysis of the gametogenesis dynamics and modularity in both human and mouse. However, all of them focused on gametogenesis, and there is a lack of research on domestic animals. Most importantly, these research studies did not address factors related to fertility.

To fill this gap, we built FifBase as benchmark for future animal fertility-related research. In order to achieve this, more than 100 000 from PubMed were fetched, among which 3301 papers were manually surveyed. Finally, 1279 literatures were filtered and kept in our database. What's more, 473 potential RNA-Seq datasets were extracted from gene expression omnibus (GEO). All the above datasets were related to 75 fertility indicators and 47 environmental factors or their lexicological variants of keywords. In our database, 1547 genes associated with fertility were obtained in 10 species, of which 1077 and 481 were from RNA-Seq dataset and text mining data, respectively, involving 2910 fertility-gene pairs and 58 fertility-environmental factors. Finally, FifBase provides a user-friendly interface to conveniently browse, retrieve and download the list of genes related to animal fertility. This elaborate database can serve as an important and valuable resource for facilitating the exploration of the poultry and livestock.

Materials and methods

Data collection, curation and processing

The data we collected were divided into two parts. The first part is mined from literatures in PubMed. A list of keywords, involving 17 species, 75 fertility indicators and 47 environmental factors or their lexicographical variants was used to search PubMed. The keywords are connected by the logical word 'AND' or 'OR'. Then we called E-search and E-fetch programs in the E-utilities interface to retrieve literature information in PubMed by Perl script. Altogether, more than 100 000 related papers from PubMed were obtained and imported into MySQL(version 5.6.40). Then, we connected literatures and genes by mapping PubMed IDs to gene list with gene2pubmed file from NCBI, and a total of

3301 publications were obtained. We manually surveyed these publications with their abstracts and checked their relevant supplement information. Finally, we further checked the full text of the filtered public literature manually and extracted the pertinent information including gene symbols, PubMed ID, fertility indicator, tissues, and species. At last, 1279 reproduction-related articles, 470 genes, were obtained after manual retrieval.

In the second part of the data collection, we systematically extracted RNA-Seq datasets related to poultry and livestock fertility from NCBI short read archive (SRA, http: //www.ncbi.nlm.nih.gov/sra/) and NCBI GEO (http://www.ncbi.nlm.nih.gov/geo/) based on scientific name of species and keywords such as 'fertility' or 'fecundity' or 'reproduction'. In total, 14 RNA-Seq datasets related to fertility were obtained for further selection, involving 4 species and 473 potential SRA runs.

Analysis of RNA-Seq data

All obtained raw RNA-Seq datasets were processed through the same pipeline. Raw RNA-Seq data were downloaded from SRA and mapped to corresponding reference genomes with HISAT2 (version2.0.5, [31]. We sorted HISAT2 output using SAMtools (version 1.8, [32], converted sam format to bam format, and then assembled and integrated transcripts of all samples using StringTie (version 1.3.6, [33]. Eventually, Ballgown package (version 2.16.0, [34] was used to standardize reads counts and identified the differentially expressed genes between the high and low fertility groups. We also normalized read counts into the FPKM values (fragments per kilobase of transcript per million mapped reads). We defined genes with adjusted *P*-value <0.05 and$|\log2$ (H/L ratio)$| \geq 2$ in each group as genes associated with fertility. At last, 1077 genes-associated fertility was obtained.

For quality control, we focused on the repeatability among biological replicates. We defined a suppositional standard sample for each species/fertility/tissue as the median value of normalized read counts of their combined samples. Based on the standard samples, we calculated the Spearman correlation coefficients of normalized read counts between each sample and its corresponding standard sample. Samples with a correlation over 0.8 were defined as qualified and remained for the following analysis.

Database implementation

FifBase database runs entirely on open source software (MySQL 5.6.40 database server, Apache web server with the interface written in PHP, Ubuntu Linux operating system). Perl scripts were written for data collection and processing in text mining. What's more, FifBase has been tested in Mozilla Firefox, Google Chrome and Apple Safari browsers. Web user interfaces are developed using JSP, HTML5, CSS3 as well as JS. For dynamic data visualization, ECharts with version 4.1.0 is incorporated to generate charts. The web interface is available online at http://www.nwsuaflmz.com/FifBase/.

Results

Data summary

A total of 470 genes associated with fertility through text mining were extracted in 10 species, whereas 1077 genes in 4 species were detected using high-throughput datasets of 473 samples from 14 projects, involving 2910 fertility-gene pairs and 58 fertility-environmental factors. Currently, FifBase hosts various data types including genomes, protein–protein interaction from STRING database, and gene expression level calculated by us based on our collected RNA-Seq datasets. The latest reference genomes and annotated information of different species were downloaded from Ensembl. Figure 1 shows the data types managed by the database and summarizes the statistics. In order to identify gene functions, gene ontology (GO) [35] and pathway and their annotation information were organized and imported into the database. In addition, some commonly used data analysis and visualization tools, such as BLAST [36] and Genome Browser, are included to provide free-accessed data visualization services to users. The basic information about our database, including species, literature number, gene number and so on were list in Table 1.

Fertility-related gene

FifBase provides annotation of fertility-related genes through Ensembl BioMart [37], such as gene ID, gene name, related GO term and pathway ID. In total, 1558 genes with unique 'ENSEMBL' IDs were obtained. Subsequently, we integrated other annotation information for each gene, such as gene structure, literature source, expression of gene and protein–protein interaction information, etc. All these information are shown on one page for easy browsing. This page contains nine parts: summary of gene (Figure 2A), expression of gene (Figure 2B), gene structure (Figure 2C), protein of gene (Figure 2D), pathway information of gene (Figure 2E), subcellular location of gene (Figure 2F), Go term of gene (Figure 2G), literature resource (Figure 2H) and regulation function (Figure 2I).

Functional annotation

GO was used to annotate the detected fertility-related gene functions. Based on GO consortium [35] and BioMart annotation tool in Ensembl [38], FifBase tried to provide more abundant annotation information. We retrieved GO terms based on the criteria that, if one gene is annotated into a GO term, then it should be annotated into its parent GO term [39], and it is achieved with Python script. As results, 1250 genes with 9044 annotations in molecular function, 1258 genes with 16 257 annotations in biological process, 1274 genes with 7677 annotations in cellular component were identified. The GO information is presented in a table, which contained 'GO_term_accession', 'GO_term_name', 'GO_term_definition', 'GO_term_evidence' and 'GO_domain', explaining the function of the gene in detail.

Pathway information of 10 species was also integrated into FifBase, based on KEGG: Kyoto Encyclopedia of Genes and Genomes (https://www.genome.jp/kegg/pathway.html), which was completed with Python scripts. As a result, 895 pathways involving 581 genes related to fertility were identified. Pathway information is also presented in the form of table, including 'Term', 'ID', '*P*-value' and external links, through with users can directly jump into the KEGG Pathway page to browse the information of the gene and its role in the pathway.

Expression

The expression section integrates genome-wide gene expression profiles derived from RNA-Seq datasets about various tissues of 10 species. We collected the RNA-Seq datasets about the 10 species from GEO and SRA database from NCBI. We then used FASTQ-dump to convert the data format to FASTQ, and trim_galore (Version 0.5.0) to process quality control (e.g.

Downloaded from https://academic.oup.com/bib/advance-article/doi/10.1093/bib/bbaa432/6120284 by University of California, San Francisco user on 26 May 2021

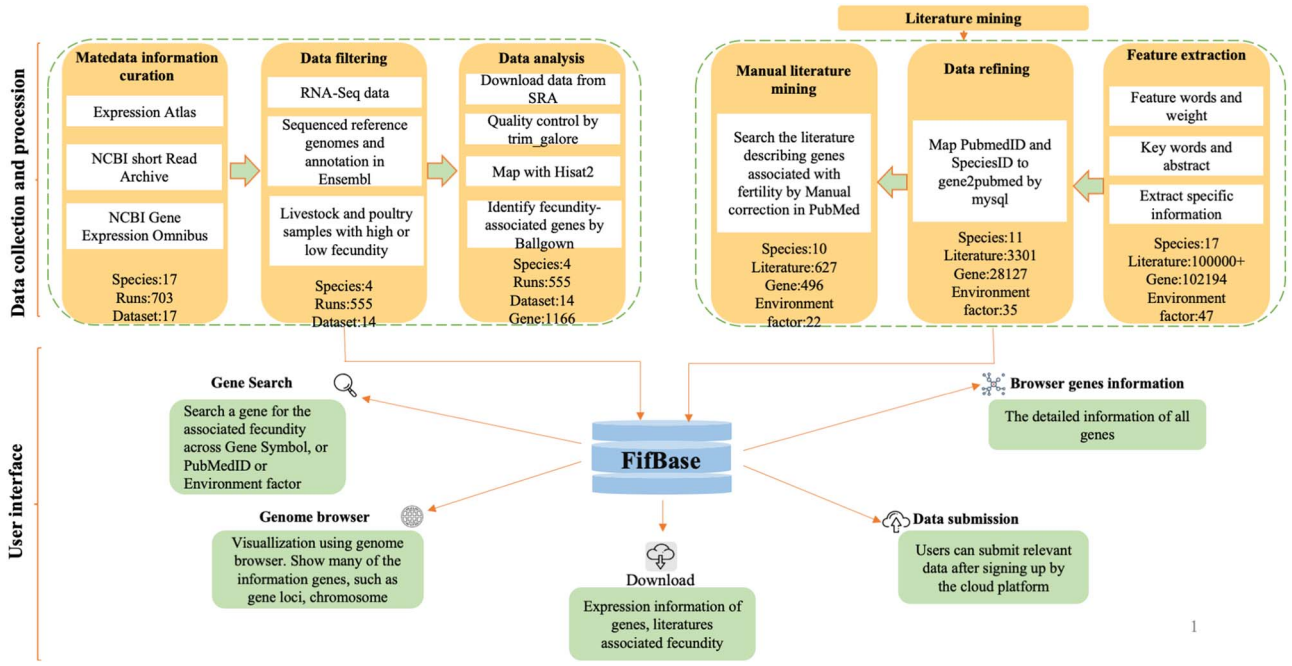

Figure 1. An overview of FifBase database. All datasets were collected from PubMed, GEO, and SRA, and were manually annotated and curated. We collected refined 10 species, all data resources were deposited in FifBase and displayed in web pages (gene search, browse, download, and submit).

Table 1. Basic information about FifBase database

Species	Literature	Gene	Environmental factor	Fertility-gene pair	Fertility-environment
Cat	8	8	1	8	1
Cattle	177	158	14	158	22
Chicken	35	34	5	34	5
Dog	17	19	1	19	1
Duck	8	8	1	8	1
Goat	21	23	3	23	3
Horse	26	24	3	24	3
Pig	143	130	14	130	14
Rabbit	9	9	3	9	4
Sheep	93	80	13	80	15

base quality >20, remove adaptors). Then, we mapped the high-quality reads to the corresponding species reference genome using HISAT2 (v2.2.0) [40], filtered samples and data with low mapping rates (<70%), and adopted FPKM (fragments per kilobase of exon per million fragments mapped) tool to estimate expression levels of genes and transcripts with Ballgown (version 2.16.0) [34]. For data with multiple replicates, we took the average of these replicates to characterize the gene expression. In the end, gene expression profiles of 151 tissues were constructed in10 species. FifBase provides a user-friendly web interface to enable easy access, search and view of gene expression profiles. For a given gene, the expression profile for all tissues will be visualized in a box plot. User can observe the expression of the gene visually in different tissues.

Literature

In order to provide a public literature repository and collection of various research studies on fertility of livestock and poultry, we provide literatures section with search navigation bar. Users can quickly locate specific articles by searching 'PubMed ID', 'Gene Symbol', 'Species', 'Title' or other related words. Abstracts of these papers were fetched from PubMed (http://www.ncbi.nlm.nih.gov/pubmed) using NCBI's E-utilities tool. The relationship between given literature and fertility was built according to fecundity index keywords, and then the genes were filtered by manual curation. In total, 1279 papers/books are included and 481 genes linked to them accordingly in 10 species, involving 75 fertility indicators and 47 environmental factors.

Blast online tool

In addition, FifBase incorporated the ViroBLAST tools for online data analysis. We formatted the nucleic acid database and protein database of 10 species including pig, cow and chicken and so on. Users can upload their sequence or sequence file to perform blastn, blastp, blastx, tblastx or tblastn against whole genome sequences, CDS regions or peptides.

Subcellular location

Subcelluar location annotation for all genes in different animals was gathered based on the subcelluar location

Downloaded from https://academic.oup.com/bib/advance-article/doi/10.1093/bib/bbaa432/6120284 by University of California, San Francisco user on 26 May 2021

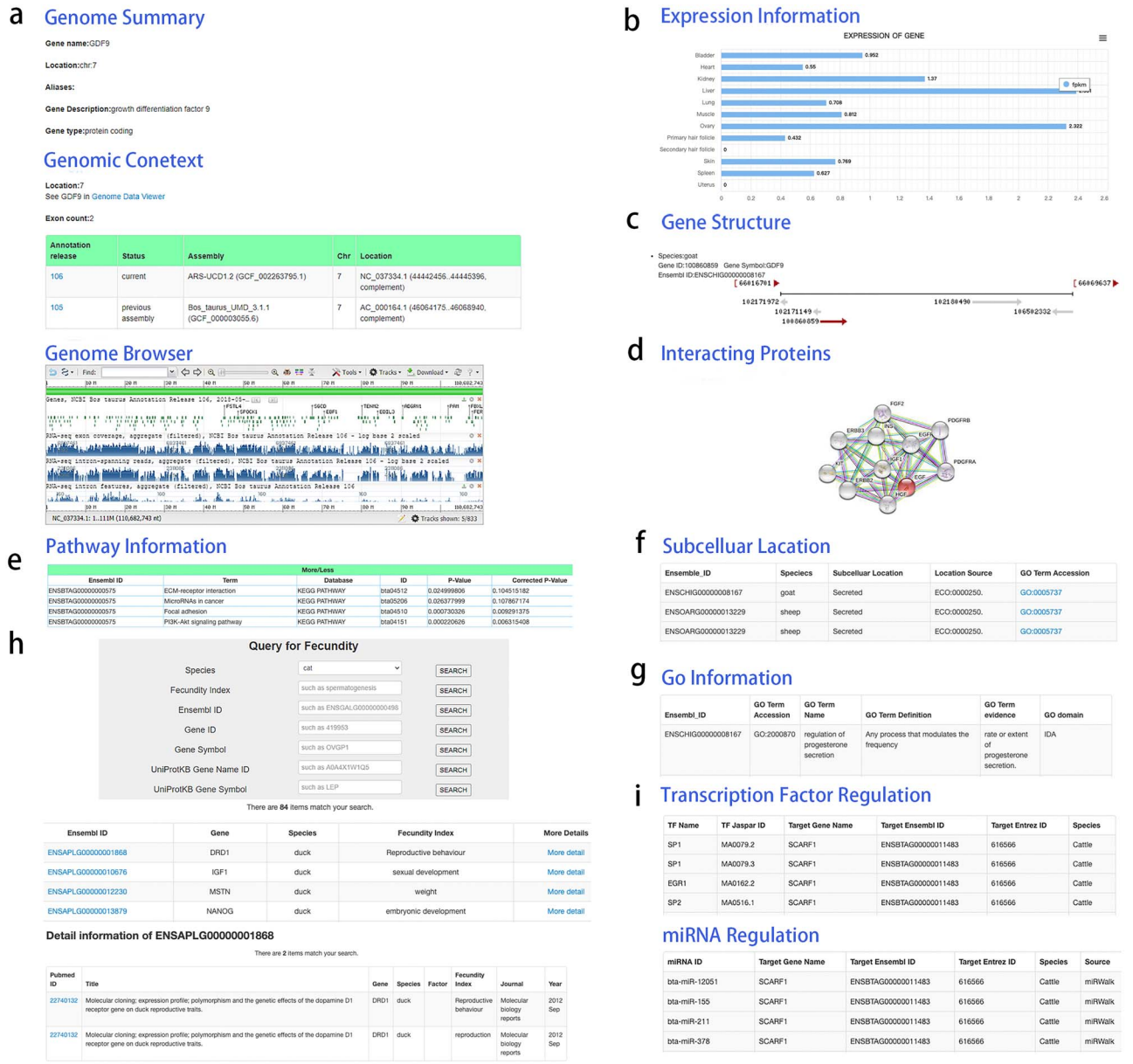

Figure 2. Annotation information of gene. (**A**) Genome summary. (**B**) Expression information. (**C**) Gene structure. (**D**) Interacting proteins. (**E**) KEGG pathway. (**F**) Subcellular location. (**G**) GO. (**H**) Search page and search result. (**I**) Regulation function.

term in UniprotKB database (https://www.uniprot.org/). Raw subcelluar location term was divided into two sections, including subcelluar location (e.g. secreted) and corresponding supporting evidence [e.g. ECO:0000256|PIRNR:PIRNR036893 (type|source, the detail description: https://www.uniprot.org/help/evidences)].

Gene regulation

In order to provide mechanistic insight into fertility of domestic animals, FifBase provides gene regulation function, including transcriptional regulation and miRNA regulation. In case of transcriptional regulation, based on JASPAR [41], we used the multiple EM for motif elicitation (MEME) and Homer algorithm to perform genome-wide identification and annotation on the relationship of TFs and their target gene. In the end, we annotated 205 genes and 512 TFs with 2834 TF-Gene regulation relationship pairs in 8 species about fertility-related genes. In case of miRNA regulation, we integrated miRecords [37], miRTarBase [42], TarBase [43], miRWalk [44], TargetScan [45] and miRDB [46] results to annotate our fertility-related genes. Finally, we obtained more than 100 000 miRNA-target gene pairs among our research species.

Downloads

All the search results can be downloaded as CSV files for customized analysis by clicking the Download button on the top navigation bar. Alternatively, FifBase offers users the RNA-Seq data analysis results in CSV files for each group on the Download page [47].

Downloaded from https://academic.oup.com/bib/advance-article/doi/10.1093/bib/bbaa432/6120284 by University of California, San Francisco user on 26 May 2021

Data submission

FifBase also provides a function to upload data in the navigation bar. After registering and logging in, users can submit the relevant data with table form to us through the submit button. In our current version, FifBase only accepts results from open access RNA-Seq datasets. The submitted data would be added to FifBase after curation and analysis as described in the section of Materials and methods.

Discussion and conclusion

Since the first fertility paper appeared in 1868, more than 100 000 papers have been published and more than 1400 genes associated have been identified. Though some of these genes have been shown to be important regulators or biomarkers in animal fertility [48], these genes scattered in thousands of papers. So, the construction of a systematic database for genes associated fertility would be a valuable and challenging work to the detection, evolution, function and mechanism studies of animal reproduction.

In addition, with the development of high-throughput detection technology, it is possible to detect the potential fertility-related genes. So, we also hunted biomarkers associated with fertility based on RNA-Seq datasets, including gene and protein–protein interaction, which will provide references for mechanism research of animal reproduction.

Furthermore, animal fertility is affected by various factors, which involve inherent and external factors. With the development of industrialization, the influence of environmental factors on animal fecundity is increasingly prominent. Many studies worldwide showed that the reproductive performance and birth rate of animals lived in contaminated environments were significantly lower than control group [49, 50]. So, detection and collection of environmental factors related with fertility are also an important work.

FifBase is dedicated to a comprehensive resource for animal reproduction. On one hand, we curated available high-quality RNA-Seq datasets about fertility through the same pipeline. On the other hand, information in literatures related to animal fecundity indicators from PubMed were extracted by text mining. After processing and filtration, factors associated with fertility were obtained and stored in FifBase, which provides user-friendly website for researchers.

As far as we know, this is the first database that systematically integrates literature and RNA-Seq information associated with fertility to support functional research and facilitate users to explore the genes and environment factors that they are interested in. As fertility-related research expand rapidly, we will update FifBase regularly by adding more genes associated with fertility when additional RNA-Seq datasets and literatures become available. FifBase will also provide more annotation information in future. Through this work, we expect FifBase will contribute to improving the productivity of poultry and livestock and become a useful resource for fertility.

Key Points

- Manually curated database FifBase, which is of genetic information about domestic animal fertility, was provided.
- FifBase covers results from both published literature and RNA-Seq datasets.
- A total of 1558 genes associated with fertility were filtered in 10 species, of which 1077 and 481 were from RNA-Seq datasets and text mining data, respectively, involving 2910 fertility-gene pairs and 58 fertility-environmental factors in FifBase.
- User-friendly interface of FifBase was provided, through which users where the fertility-related factor information, including gene annotation and environmental factors can be browsed, retrieved and downloaded.

Data Availability

FifBase is an open database, all data are freely available at: http://www.nwsuaflmz.com/FifBase. This statement was also recorded in abstract.

Funding

This work was financially supported by the National Natural Science Foundation of China (grant number: 61772431, 62072377); Program of Shaanxi Province Science and Technology Innovation Team (Grant number: 2019TD-036); China National Basic Research Program (Grant number: 2016YFA0100203); the Mathematical Tianyuan Fund of the National Natural Science Foundation of China (Grant number: 12026414).

References

1. Vander Borght M, Wyns C. Fertility and infertility: definition and epidemiology. *Clin Biochem* 2018;**62**:2–10.
2. Diskin M, Morris D. Embryonic and early foetal losses in cattle and other ruminants. *Reprod Domest Anim* 2008;**43**: 260–7.
3. Santos JEP, Thatcher WW, Chebel RC, *et al*. The effect of embryonic death rates in cattle on the efficacy of estrus synchronization programs. *Anim Reprod Sci* 2004;**82-83**:513–35.
4. Hansen PJ, Block J. Towards an embryocentric world: the current and potential uses of embryo technologies in dairy production. *Reprod Fertil Dev* 2004;**16**:1–14.
5. Diskin MG, Murphy JJ, Sreenan JM. Embryo survival in dairy cows managed under pastoral conditions. *Anim Reprod Sci* 2006;**96**:297–311.
6. Moore SG, Pryce JE, Hayes BJ, *et al*. Differentially expressed genes in endometrium and corpus luteum of Holstein cows selected for high and low fertility are enriched for sequence variants associated with fertility1. *Biol Reprod* 2016;**94**:19.
7. Looney CR, Nelson JS, Schneider HJ, *et al*. Improving fertility in beef cow recipients. *Theriogenology* 2006;**65**:201–9.
8. Ferraz PA, Burnley C, Karanja J, *et al*. Factors affecting the success of a large embryo transfer program in Holstein cattle in a commercial herd in the southeast region of the United States. *Theriogenology* 2016;**86**:1834–41.
9. Roche J, Burke C, Meier S, *et al*. Nutrition reproduction interaction in pasture-based systems: Is nutrition a factor in reproductive failure? *Animal Production Science* 2011;**51**:1045–66.
10. Roche JR, Bell AW, Overton TR, *et al*. Nutritional management of the transition cow in the 21st century – a paradigm shift in thinking. *Animal Production Science* 2013;**53**:1000–23.

Downloaded from https://academic.oup.com/bib/advance-article/doi/10.1093/bib/bbaa432/6120284 by University of California, San Francisco user on 26 May 2021

11. Geary TW, Burns GW, Moraes JGN, *et al*. Identification of beef heifers with superior uterine capacity for pregnancy1. *Biol Reprod* 2016;**95**:47.
12. Lucy MC. Reproductive loss in high-producing dairy cattle: where will it end? *J Dairy Sci* 2001;**84**:1277–93.
13. Walsh SW, Williams EJ, Evans ACO. A review of the causes of poor fertility in high milk producing dairy cows. *Anim Reprod Sci* 2011;**123**:127–38.
14. English AM, Byrne CJ, Cormican P, *et al*. Effect of early calf-hood nutrition on the transcriptional regulation of the hypothalamic-pituitary-testicular axis in holstein-friesian bull calves. *Sci Rep* 2018;**8**:16577.
15. Wei Z, Li P, Huang S, *et al*. Identification of key genes and molecular mechanisms associated with low egg production of broiler breeder hens in ad libitum. *BMC Genomics* 2019;**20**:408.
16. Itoh F, Obara Y, Rose MT, *et al*. Insulin and glucagon secretion in lactating cows during heat exposure1. *J Anim Sci* 1998;**76**:2182–9.
17. Yang Q-E, Ozawa M, Zhang K, *et al*. The requirement for protein kinase C delta (PRKCD) during preimplantation bovine embryo development. *Reprod Fertil Dev* 2016;**28**:482–90.
18. Yasuo S, Nakao N, Ohkura S, *et al*. Long-day suppressed expression of type 2 deiodinase gene in the Mediobasal hypothalamus of the Saanen goat, a short-day breeder: implication for seasonal window of thyroid hormone action on reproductive neuroendocrine Axis. *Endocrinology* 2006;**147**:432–40.
19. Bisinotto RS, Ribeiro ES, Santos JEP. Synchronisation of ovulation for management of reproduction in dairy cows. *animal* 2014;**8**:151–9.
20. Butler ST. Nutritional management to optimize fertility of dairy cows in pasture-based systems. *Animal* 2014;**8**:15–26.
21. Abdollahi-Arpanahi R, Carvalho MR, Ribeiro ES, *et al*. Association of lipid-related genes implicated in conceptus elongation with female fertility traits in dairy cattle. *J Dairy Sci* 2019;**102**:10020–9.
22. Janečka JE, Davis BW, Ghosh S, *et al*. Horse Y chromosome assembly displays unique evolutionary features and putative stallion fertility genes. *Nat Commun* 2018;**9**:2945.
23. Zhang W, Chen Z, Zhang D, *et al*. KHDC3L mutation causes recurrent pregnancy loss by inducing genomic instability of human early embryonic cells. *PLoS Biol* 2019;**17**:e3000468.
24. Smarr MM, Sundaram R, Honda M, *et al*. Urinary concentrations of parabens and other antimicrobial chemicals and their association with couples' fecundity. *Environ Health Perspect* 2017;**125**:730–6.
25. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2009;**26**:139–40.
26. Moraes JGN, Behura SK, Geary TW, *et al*. Uterine influences on conceptus development in fertility-classified animals. *Proc Natl Acad Sci U S A* 2018;**115**:E1749–58.
27. Willems E, Guerrero-Bosagna C, Decuypere E, *et al*. Differential expression of genes and DNA methylation associated with prenatal protein undernutrition by albumen removal in an avian model. *Sci Rep* 2016;**6**:20837.
28. Tao Z, Song W, Zhu C, *et al*. Comparative transcriptomic analysis of high and low egg-producing duck ovaries. *Poult Sci* 2017;**96**:4378–88.
29. Zhang Y, Zhong L, Xu B, *et al*. SpermatogenesisOnline 1.0: a resource for spermatogenesis based on manual literature curation and genome-wide data mining. *Nucleic Acids Res* 2012;**41**:D1055–62.
30. Che D, Wang Y, Bai W, *et al*. Dynamic and modular gene regulatory networks drive the development of gametogenesis. *Brief Bioinform* 2016;**18**:712–21.
31. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* 2015;**12**:357–60.
32. Li H, Handsaker B, Wysoker A, *et al*. The sequence alignment/map format and SAMtools. *Bioinformatics* 2009;**25**: 2078–9.
33. Pertea M, Pertea GM, Antonescu CM, *et al*. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol* 2015;**33**:290–5.
34. Frazee AC, Pertea G, Jaffe AE, *et al*. Ballgown bridges the gap between transcriptome assembly and expression analysis. *Nat Biotechnol* 2015;**33**:243–6.
35. The Gene Ontology Consortium. Expansion of the gene ontology knowledgebase and resources. *Nucleic Acids Res* 2016;**45**:D331–8.
36. Deng W, Nickle DC, Learn GH, *et al*. ViroBLAST: a stand-alone BLAST web server for flexible queries of multiple databases and user's datasets. *Bioinformatics* 2007;**23**:2334–6.
37. Kinsella RJ, Kähäri A, Haider S, *et al*. Ensembl BioMarts: a hub for data retrieval across taxonomic space. *Database* 2011;**2011**.
38. Tang B, Zhou Q, Dong L, *et al*. iDog: an integrated resource for domestic dogs and wild canids. *Nucleic Acids Res* 2018;**47**:D793–800.
39. Kim D, Paggi JM, Park C, *et al*. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* 2019;**37**:907–15.
40. Fornes O, Castro-Mondragon JA, Khan A, *et al*. JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res* 2020;**48**:D87–92.
41. Xiao F, Zuo Z, Cai G, *et al*. miRecords: an integrated resource for microRNA-target interactions. *Nucleic Acids Res* 2009;**37**:D105–110.
42. Chou CH, Shrestha S, Yang CD, *et al*. miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. *Nucleic Acids Res* 2018;**46**:D296–302.
43. Karagkouni D, Paraskevopoulou MD, Chatzopoulos S, *et al*. DIANA-TarBase v8: a decade-long collection of experimentally supported miRNA-gene interactions. *Nucleic Acids Res* 2018;**46**:D239–45.
44. Sticht C, De La Torre C, Parveen A, *et al*. miRWalk: an online resource for prediction of microRNA binding sites. *PLoS One* 2018;**13**:e0206239.
45. Agarwal V, Bell GW, Nam JW, *et al*. Predicting effective microRNA target sites in mammalian mRNAs. *Elife* 2015;**4**:e05005.
46. Chen Y, Wang X. miRDB: an online database for prediction of functional microRNA targets. *Nucleic Acids Res* 2020;**48**: D127–31.
47. Shi M-W, Zhang N-A, Shi C-P, *et al*. SAGD: a comprehensive sex-associated gene database from transcriptomes. *Nucleic Acids Res* 2018;**47**:D835–40.
48. Wang D, Yang L, Zhang P, *et al*. AAgAtlas 1.0: a human autoantigen database. *Nucleic Acids Res* 2016;**45**:D769–76.
49. Hani YMI, Turies C, Palluel O, *et al*. Effects of a chronic exposure to different water temperatures and/or to an environmental cadmium concentration on the reproduction of the threespine stickleback (Gasterosteus aculeatus). *Ecotoxicol Environ Saf* 2019;**174**:48–57.
50. Bjørklund G, Aaseth J, Dadar M, *et al*. Exposure to environmental organic mercury and impairments in human fertility. *J Reproduct Infer* 2019;**20**:195–7.

Downloaded from https://academic.oup.com/bib/advance-article/doi/10.1093/bib/bbaa432/6120284 by University of California, San Francisco user on 26 May 2021