Follicle Online: an integrated database of follicle assembly, development and ovulation

Juan Hua1,2,†, Bo Xu1,3,†, Yifan Yang4, Rongjun Ban1, Furhan Iqbal1,5,6,7, Howard JCooke1, Yuanwei Zhang1,5,6,* and Qinghua Shi1,5,6,*

1Molecular and Cell Genetics Laboratory, The CAS Key Laboratory of Innate Immunity and Chronic Disease, Hefei National Laboratory for Physical Sciences at Microscale, School of Life Sciences, University of Science and Technology of China, Hefei 230027, China, 2Department of Biochemistry and Molecular Biology, School of Basic Medical Sciences, Anhui Medical University, Hefei 230032, China, 3Center for Reproductive Medicine, Anhui Provincial Hospital Affiliated to Anhui Medical University, Hefei 230001, China, 4Department of Statistics, University of Kentucky, Lexington, KY 40506, USA, 5Collaborative Innovation Center of Genetics and Development, Fudan University, 2005 Songhu Road, Shanghai 200438, China, 6Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei 230031, China and 7Institute of Pure and Applied Biology, Bahauddin Zakariya University Multan, 60800, Pakistan

*Corresponding author: Tel/Fax: +86-0551-63600442, Email: zyuanwei@mail.ustc.edu.cn
*Correspondence may also be addressed to Qinghua Shi. Tel/Fax: +86-0551-63600344, Email: qshi@ustc.edu.cn
†These authors contributed equally to this work.

Citation details: Hua,J., Xu,B., Yang,Y., et al. Follicle online: an integrated database of follicle assembly, development and ovulation. Database (2015) Vol. 2015: article ID bav036; doi:10.1093/database/bav036

Received 15 November 2014; Revised 6 March 2015; Accepted 31 March 2015

Abstract

Folliculogenesis is an important part of ovarian function as it provides the oocytes for female reproductive life. Characterizing genes/proteins involved in folliculogenesis is fundamental for understanding the mechanisms associated with this biological function and to cure the diseases associated with folliculogenesis. A large number of genes/proteins associated with folliculogenesis have been identified from different species. However, no dedicated public resource is currently available for folliculogenesis-related genes/proteins that are validated by experiments. Here, we are reporting a database ‘Follicle Online’ that provides the experimentally validated gene/protein map of the folliculogenesis in a number of species. Follicle Online is a web-based database system for storing and retrieving folliculogenesis-related experimental data. It provides detailed information for 580 genes/proteins (from 23 model organisms, including Homo sapiens, Mus musculus, Rattus norvegicus, Mesocricetus auratus, Bos Taurus, Drosophila and Xenopus laevis) that have been reported to be involved in folliculogenesis, POF (premature ovarian failure) and PCOS (polycystic ovary syndrome). The literature was manually curated from more than 43 000 published articles (till 1 March 2014). The Follicle Online database is implemented in PHP + MySQL + JavaScript and this user-friendly web application provides access to the stored data. In summary, we have developed a centralized
Introduction

Folliculogenesis consists of three main processes including primordial follicle assembly, follicle development and ovulation (1–3). It is the progression of number of small primordial follicles into large preovulatory follicles that enters the menstrual cycle (4, 5). Folliculogenesis begins with primordial follicle assembly (during the perinatal period in human, mouse and rat) during which the oocytes that survive the process of germ cell cluster breakdown are individually surrounded with squamous pre-granulosa cells (6–11). The recruitment (or the initial growth) of primordial follicles is defined by a dramatic enlargement of the oocyte itself, accompanied by proliferation and differentiation of the surrounding granulosa cells (11–14). Ovulation is the last step of folliculogenesis, which involves the breakdown of a mature ovarian follicle to release a fertilizable oocyte (14–18).

The ultimate role of the folliculogenesis, a tightly regulated process that lasts for two to five decades in women, is to produce a healthy egg that will transmit genes to the next generation (18–23). Inappropriate coordination of these events will result in serious ovarian pathologies such as POF (premature ovarian failure) and PCOS (polycystic ovary syndrome) (24, 25). In humans, POF is defined as an ovarian defect characterized by absence of menstruation (primary amenorrhea) or premature depletion of ovarian follicles/arrested folliculogenesis before the age of 40 years (secondary amenorrhea, 25). POF affects ~1 in 10,000 women by the age of 20 (0.01%), 1 in 1,000 women by the age of 30 (0.1%) and 1 in 100 women by the age of 40 (1%, 24). PCOS is another common female endocrine disorder that affects ~5-10% of women of reproductive age (12–45 years) (24). These two diseases are considered to be the leading causes of female subfertility and the most frequent endocrine problems in women of reproductive age (24). Currently there are no effective treatments for these ovarian pathologies and in many cases, POF and PCOS are generally diagnosed only when women have already lost their fertility (24, 25).

Although the physiological function of folliculogenesis is well defined, the mechanisms controlling this process are not properly understood. Until recently, the tools available to reproductive biologists to determine the functions of genes involved in mammalian folliculogenesis were based on the study of homologous genes in lower model organisms and the production of genetically modified mice with targeted disruptions of genes of interest (26, 27). These studies not only suggested that which molecules may be involved in the folliculogenesis in human but they were also able to highlight specific candidate molecules that can be used as targets for potential drugs used to treat ovarian pathologies (26, 27). However, information about the genes/proteins involved is folliculogenesis was scattered across thousands of publications and extraction of relevant functional information from various resources is a significant task. Thus, an integrated database of the existing resources regarding folliculogenesis was urgently required.

By searching literature, we manually collected unique proteins from 23 species with experimentally verified functions during folliculogenesis. This led us to the development of an integrated and searchable database, Follicle Online. We present here the database (Follicle Online) of genes experimentally linked with folliculogenesis (also PCOS and POF). This online service is implemented in PHP + MySQL + JavaScript. Follicle Online can provide potentially orthologous genes for these genes/proteins reported in Follicle Online from seven model organisms, including Homo sapiens, Mus musculus, Rattus norvegicus, Macaca mulatta, Danio rerio, Bos taurus and Gallus gallus. Follicle Online will be regularly updated as soon as new genes/proteins regulating folliculogenesis (also PCOS and POF) will be reported. It possess a comprehensive set of features, which allows users to explore different aspects of functionality of the folliculogenesis genes indexed in Follicle Online and provides important information required for further experimentally or bioinformatics analysis of these genes.

Materials and methods

The general process of data collection and annotation for Follicle Online is illustrated in Figure 1.

Data generation

As data quality is critical for any database, we searched PubMed with specific key words for different stages of
foliculogenesis and have collected experimentally identified functional genes/proteins from more than 43,000 published articles (till 1 March 2015). The inclusion criterion was that only those genes/proteins will be represented in database the function of who has been experimentally verified. Genes/Proteins for which only expression information was available were excluded from Follicle Online.

To search for follicle assembly genes/proteins in PubMed, we used the keywords ‘primordial follicle endowment’, ‘primordial follicle formation’ and ‘primordial follicle assembly’. For proteins involve in follicle development, we used the keywords ‘primary follicle’, ‘follicle activation’, ‘follicle transition’, ‘follicle recruitment’, ‘secondary follicle’, ‘preantral follicle’, ‘preantral foliculogenesis’, ‘follicle development’, ‘antral follicle formation’, ‘preovulatory follicle’ and ‘follicle growth’ as queries, as these 11 keywords can cover the entire process of follicular growth. In addition, for genes/proteins associated with ovulation, we used ‘ovulation’ as keyword to search PubMed. To avoid any missing data, we also searched by using the keywords ‘cumulus expansion’ and ‘luteinization’. For PCOS and POF, we used ‘polycystic ovary syndrome’ and ‘premature ovarian failure’ as keywords to search in PubMed, respectively. In total, we collected 578 unique Follicle Online genes/proteins from 23 organisms (Table 1).

**Figure 1.** Follicle Online Database scheme.

**Table 1.** Follicle Online data types and their corresponding numbers

| Species (number) | A | D | O | PCOS | POF | Multi-process |
|------------------|---|---|---|------|-----|--------------|
| Mus musculus (207) | 12 | 104 | 67 | 2   | 0   | 22           |
| Homo sapiens (175) | 2  | 18  | 7  | 118  | 24  | 6            |
| Rattus norvegicus (66) | 3  | 43  | 8  | 6   | 0   | 6            |
| Bos taurus (33) | 0  | 17  | 13 | 0    | 0   | 3            |
| Sus scrofa (28) | 1  | 11  | 14 | 0    | 0   | 2            |
| Others (71) | 11 | 29  | 22 | 0    | 4   | 5            |
| **Total (580)** | **23** | **222** | **131** | **126** | **28** | **44** |

A, follicle assembly; D, follicle development; O, ovulation. PCOS, polycystic ovary syndrome; POF, premature ovarian failure.
Protein annotation

Genes/Proteins in Follicle Online were annotated as follows: (i) the basic information [e.g. name/synonyms, protein sequences, nucleotide sequences, PI (isoelectric point) and Mw (molecular weight)] of the genes/proteins were annotated referring to GenBank and UniProt Knowledgebase (Release 2015.01); (ii) detail description of the genes/proteins [subcellular location, folliculogenesis functional stage and disease type] were provided from the experimentally verified data, also the result figures and abstracts of literatures (PubMed) about genes/proteins reported in Follicle Online are provided; (iii) the PPI (protein–protein interaction) information, the orthologous of each genes/proteins, functional domain, structural domain and the GO term were also provided. The PPI information was integrated from several major public databases, such as HPRD (Release 9), BioGRID (Version 3.2.119), DIP (Released 20130131), MINT (Downloaded 20130520), IntAct (Downloaded 20140329) and STRING (Version 9.1) (28–31). The redundant PPIs were removed and the cleared PPIs were divided into two parts, the experiment verified and predicted PPIs.

Microarray data collection

To collect the expression information for genes/proteins reported in Follicle Online database, the ArrayExpress database (http://www.ebi.ac.uk/arrayexpress) was used as a resource. We screened the whole ArrayExpress database and downloaded whole transcript microarray experiment datasets related with folliculogenesis from 18 mouse models by using Affymetrix GeneChip Mouse Genome 430 2.0 platform. The mRNA expression information of genes reported in Follicle Online database were extracted from these microarray data and presented in the web page.

Results

User interface—simple and advanced search

Follicle Online provides a search engine for users to search the genes/proteins of their interest. The search option (http://mcg.ustc.edu.cn/sdap1/follicle/follado.php) provides an interface for querying the Follicle Online database with one or more keywords such as genes/proteins names, UniProt ID or Follicle Online ID, functional stage and function etc (Figure 2A).

There is also an ‘Advance search’ option in Follicle Online that allows multiple fields search simultaneously (Figure 2B). Search options includes: (i) advance search where users can use relatively complex or combination of keywords to locate the precise information. The interface of search-engine permits the querying by different annotation fields and the queries are linked by three operators (‘and’, ‘or’ and ‘exclude’); (ii) Browsing, instead of searching for a specific protein, all entries of Follicle Online database could be listed by species name, functional stage. The human reproductive tract diseases are also included in this database as a dropdown menu under the ‘Disease’ section. Users can choose the disease of their interest from this menu and can explore it further; (iii) BLAST search; this option was designed to quickly find all the related information about a gene/protein reported in Follicle Online database. For this purpose, the blast program in NCBI BLAST is included in Follicle Online database. Users can input a protein sequence in FASTA format for searching identical or homologous proteins; (iv) orthologous search, user can specifically browse the orthologous information for a protein by providing the genes/proteins names.

User interface results

The results for each search returns information in tabular format (Figure 3). The results for each search are displayed in three part: (i) the basic information [e.g. gene name, nucleotide and protein sequences, Mw (molecular weight)]; (ii) the literature information [functional stage, disease type (follicle endowment, growth, ovulation, PCOS or POF), regulation type (positive or negative), figures and abstracts from the reference literature, functional description, subcellular location of FAL proteins, tissue specificity of FAL proteins and GO annotation]; (iii) the other database information [Protein domain organization, PPI information and mRNA expression information from folliculogenesis related microarray data].

Feedback

Users are welcome to review Follicle Online records. They can revise the Follicle Online records as well as can submit novel proteins or proteins with novel function in folliculogenesis or in any disorder associated with folliculogenesis. The contribution from the users is highly encouraged.

Discussion

Folliculogenesis is a multi-faceted and tightly regulated process that includes primordial follicle assembly, follicle development and ovulation (1–6). In recent years, many critical molecules involved in folliculogenesis have been identified, mainly through the use of genetically modified mouse models. The data obtains from these models provide clues about abnormal reproductive conditions in human (7–9). Multiple genes have been reported to
regulate folliculogenesis and fertility in animals and they also play role in human reproduction because of evolutionary conservation between species (10–12). To effectively analyse data generated in various experiments, it is necessary to manage this data in an organized manner.

Follicle Online is a database for storing and retrieving all the folliculogenesis-related experimental data. The core of Follicle Online consists of genes/proteins involved in folliculogenesis with their experimentally confirmed functional role. All these genes/proteins identified in Homo sapiens, Mus musculus and Rattus norvegicus were submitted to a GO enrichment analysis. Using the proteome as the background, we statistically calculated the represented biological processes, molecular functions and cellular components in Follicle Online genes/proteins distribution ($P < 0.05$). The top 10 most enriched GO terms in each category are shown in Supplementary Tables S1–S3. We then enriched the biological pathway of 580 Follicle Online genes/proteins from Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway. The enriched pathway entries in each category are represented in Supplementary Tables S4–S6. This analysis further illustrates the essential roles played by genes/proteins reported in Follicle Online during folliculogenesis and facilitates the users to uncover the undiscovered functions of these genes/proteins from Homo sapiens, Mus musculus and Rattus norvegicus documented in Follicle Online.

Folliculogenesis must be, in part, dictated by networks of genes expressed in the ovary. Thus, PPI information for proteins reported in Follicle Online was critical for understanding the mechanisms of folliculogenesis. Combined with experimentally validated and computationally predicted PPIs, we constructed a potential protein network for folliculogenesis with 111 proteins having 208 interactions (Figure 4). When we carefully analysed the scientific literature, we found that some of our listed interactions have been reported to be involved in folliculogenesis. For example, John et al. (32, 33) showed that the PI3K pathway has a key role in the initiation of follicle growth, while it has been documented that oocyte-specific deletion of Pten in mice resulted in POF (34). Female mice deficient for Pten interaction proteins Akt1 and PDK1 also exhibited significant primordial follicle loss and POF (35, 36). Esr1, the Akt1 interaction proteins also have been shown to take part in folliculogenesis (37).
Moreover, our results also indicated a number of interesting interactions whose function in folliculogenesis is unknown at the moment. Hence, to explore the role of these proteins in folliculogenesis would be very interesting for the groups working in this research field.

Follicle Online is specifically focused on the genes/proteins involved in folliculogenesis and is distinct from the published online public databases like Ovarian Kaleidoscope that is providing about all the proteins expressing in ovary (38). The first differentiating feature is that the data included in Follicle Online are from 23 model organisms including Homo sapiens, Mus musculus, Rattus norvegicus. However, the data in the Kaleidoscope database is only from species such as Homo sapiens, Mus...
In summary, we have developed a database (Follicle Online) that provides a comprehensive platform to get the detailed experimentally validated information regarding the 580 genes/proteins involved in folliculogenesis, POF and PCOS. Follicle Online can be accessed via http://mcg.ustc.edu.cn/sdap1/follicle/index. All the data stored in the database can be accessed without any registration. Follicle Online was also added as unique features of this database.

In summary, we have developed a database (Follicle Online) that provides a comprehensive platform to get the detailed experimentally validated information regarding the 580 genes/proteins involved in folliculogenesis, POF and PCOS. Follicle Online can be accessed via http://mcg.ustc.edu.cn/sdap1/follicle/index. All the data stored in the database can be accessed without any registration. Follicle Online was also added as unique features of this database.

In summary, we have developed a database (Follicle Online) that provides a comprehensive platform to get the detailed experimentally validated information regarding the 580 genes/proteins involved in folliculogenesis, POF and PCOS. Follicle Online can be accessed via http://mcg.ustc.edu.cn/sdap1/follicle/index. All the data stored in the database can be accessed without any registration. Follicle Online was also added as unique features of this database.

Supplementary Data

Supplementary data are available at Database Online.

Funding

This work was supported by the National Basic Research Program (2013CB947902, 2012CB944402 and 2013CB945502) of China (973), the National Natural Science Foundation of China (31301227), Major Program of Development Foundation of Hefei Center for Physical Science and Technology (2014FXZY003). Research of BSKY (0108035101) from Anhui Medical University, Natural Science Foundation of Anhui Provincial of China (1308085QH131).

Conflict of interest. None declared.

References

1. Adhikari,D. and Liu,K. (2009) Molecular mechanisms underlying the activation of mammalian primordial follicles. Endocr. Rev., 30, 438–464.
2. McGee,E.A. and Hsueh,A.J. (2000) Initial and cyclic recruitment of ovarian follicles. Endocr. Rev., 21, 200–214.
3. Vanderhyden,B. (2002) Molecular basis of ovarian development and function. Front Biosci., 7, d2006–d2022.
4. DiZerega,G.S., Turner,C.K., Stouffer,R.L. et al. (1981) Suppression of follicle-stimulating hormone-dependent folliculogenesis during the primate ovarian cycle. J. Clin. Endocrinol. Metab., 52, 451–456.
5. Edson,M.A., Nagaraja,A.K., and Matzuk,M.M. (2009) The mammalian ovary from genesis to revelation. Endocr. Rev., 30, 624–712.
6. Pepling,M.E. (2006) From primordial germ cell to primordial follicle: mammalian female germ cell development. Genesis, 44, 622–632.
7. Pepling,M.E. and Spradling,A.C. (2001) Mouse ovarian germ cell cysts undergo programmed breakdown to form primordial follicles. Develop. Biol., 234, 339–351.
8. Baker,T.G. (1963) A quantitative and cytological study of germ cells in human ovaries. Proc. R Soc. Lond. B Biol. Sci., 158, 417–433.
9. Hirshfeld,A.N. (1991) Development of follicles in the mammalian ovary. Int. Rev. Cytol., 124, 43–101.
10. Broekmans,F.J., Knauff,E.A., te Velde,E.R. et al. (2007) Female reproductive ageing: current knowledge and future trends. Trends Endocrinol. Metab. TEM, 18, 58–65.
11. Eppig,J.J. (2001) Oocyte control of ovarian follicular development and function in mammals. Reproduction, 122, 829–838.
12. Faddy,M.J. and Gosden,R.G. (1996) A model conforming the decline in follicle numbers to the age of menopause in women. Hum. Reprod., 11, 1484–1486.
13. Hansen,K.R., Knowlton,N.S., Thyer,A.C. et al. (2008) A new model of reproductive aging: the decline in ovarian non-growing follicle number from birth to menopause. Hum. Reprod., 23, 699–708.
14. Wandji,S.A., Srsen,V., Voss,A.K. et al. (1996) Initiation in vitro of growth of bovine primordial follicles. Biol. Reprod., 55, 942–948.
15. Richards,J.S., Russell,D.L., Ochsner,S. et al. (2002) Novel signaling pathways that control ovarian follicular development, ovulation, and luteinization. Recent Prog. Horm. Res., 57, 195–220.
16. Fathalla, M.F. (1971) Incessant ovulation—a factor in ovarian neoplasia? *Lancet*, 2, 163.

17. Rolaki, A., Drakakis, P., Millingos, S. *et al.* (2005) Novel trends in follicular development, atresia and corpus luteum regression: a role for apoptosis. *Reprod. Biomed. Online*, 11, 93–103.

18. Richards, J.S. (1980) Maturation of ovarian follicles: actions and interactions of pituitary and ovarian hormones on follicular cell differentiation. *Physiol. Rev.*, 60, 51–89.

19. Assou, S., Anahory, T., Pantesco, V. *et al.* (2006) The human cumulus—oocyte complex gene-expression profile. *Hum. Reprod.*, 21, 1705–1719.

20. Devoto, L., Fuentes, A., Kohen, P. *et al.* (2009) The human corpus luteum: life cycle and function in natural cycles. *Fertil. Steril.*, 92, 1067–1079.

21. Edwards, R.G. (1965) Maturation in vitro of mouse, sheep, cow, pig, rhesus monkey and human ovarian oocytes. *Nature*, 208, 349–351.

22. Fan, H.Y., Liu, Z., Shimada, M. *et al.* (2009) MAPK3/1 (ERK1/2) in ovarian granulosa cells are essential for female fertility. *Science*, 324, 938–941.

23. Mehllmann, L.M. (2005) Stops and starts in mammalian oocytes: recent advances in understanding the regulation of meiotic arrest and oocyte maturation. *Reproduction*, 130, 791–799.

24. Goodarzi, M.O., Dumesic, D.A., Chazenbalk, G. *et al.* (2011) Polycystic ovary syndrome: etiology, pathogenesis and diagnosis. *Nat. Rev. Endocrinol.*, 7, 606–619.

25. Jagarlamudi, K., Liu, L., Adhikari, D. *et al.* (2009) Oocyte-specific deletion of Pten in mice reveals a stage-specific function of PTEN/PI3K signaling in oocytes in controlling follicular activation. *PLoS One*, 4, e6186.

26. Reddy, P., Adhikari, D., Zheng, W. *et al.* (2009) PDK1 signaling in oocytes controls reproductive aging and lifespan by manipulating the survival of primordial follicles. *Hum. Mol. Genet.*, 18, 2813–2824.

27. Matzuk, M.M. and Lamb, D.J. (2008) The biology of infertility: research advances and clinical challenges. *Nat. Med.*, 14, 1197–1213.

28. Stark, C., Breierkrautz, B.J., Chatr-Aryamontri, A. *et al.* (2011) The BioGRID Interaction Database: 2011 update. *Nucleic Acids Res.*, 39, D698–D704.

29. Xenarios, I., Rice, D.W., Salwinski, L. *et al.* (2000) DIP: the database of interacting proteins. *Nucleic Acids Res.*, 28, 289–291.

30. Zanzoni, A., Montecchi-Palazzi, L., Quondam, M. *et al.* (2002) MINT: a Molecular INTeraction database. *FEBS Lett.*, 513, 135–140.

31. Kerrien, S., Aranda, B., Breuza, L. *et al.* (2012) The IntAct molecular interaction database in 2012. *Nucleic Acids Res.*, 40, D841–D846.

32. Emeljan, J.A., Luo, J., and Cantley, L.C. (2006) The evolution of phosphatidylinositol 3-kinases as regulators of growth and metabolism. *Nat. Rev. Genet.*, 7, 606–619.

33. Jagarlamudi, K., Liu, L., Adhikari, D. *et al.* (2009) Oocyte-specific deletion of Pten in mice reveals a stage-specific function of PTEN/PI3K signaling in oocytes in controlling follicular activation. *PLoS One*, 4, e6186.

34. Reddy, P., Adhikari, D., Zheng, W. *et al.* (2009) PDK1 signaling in oocytes controls reproductive aging and lifespan by manipulating the survival of primordial follicles. *Hum. Mol. Genet.*, 18, 2813–2824.

35. Brown, C., LaRocca, J., Pietruska, J. *et al.* (2010) Subfertility caused by altered follicular development and oocyte growth in female mice lacking PKB alpha/Akt1. *Biol. Reprod.*, 82, 246–256.

36. Kim, S., Pyun, J.A., Kang, H. *et al.* (2011) Epistasis between CYP19A1 and ESR1 polymorphisms is associated with premature ovarian failure. *Fertil. Steril.*, 95, 353–356.

37. Le, C.P., Vitt, U.A., and Hsueh, A.J. (2000) The Ovarian Kaleidoscope database: an online resource for the ovarian research community. *Endocrinology*, 141, 3052–3054.