



MeiosisOnline: A Manually Curated Database for Tracking and Predicting Genes Associated With Meiosis

Xiaohua Jiang[†], Daren Zhao[†], Asim Ali[†], Bo Xu, Wei Liu, Jie Wen, Huan Zhang^{*}, Qinghua Shi^{*} and Yuanwei Zhang^{*}

First Affiliated Hospital of USTC, Hefei National Laboratory for Physical Sciences at Microscale, School of Basic Medical Sciences, Division of Life Sciences and Medicine, CAS Center for Excellence in Molecular Cell Science, University of Science and Technology of China, Hefei, China

OPEN ACCESS

Edited by:

Liangran Zhang, Shandong University, China

Reviewed by:

Xuejiang Guo, Nanjing Medical University, China Zuping He, Hunan First Normal University, China

*Correspondence:

Huan Zhang zhh1985@ustc.edu.cn Qinghua Shi qshi@ustc.edu.cn Yuanwei Zhang zyuanwei@ustc.edu.cn

[†]These authors have contributed equally to this work

Specialty section:

This article was submitted to Cell Growth and Division, a section of the journal Frontiers in Cell and Developmental Biology

> Received: 26 February 2021 Accepted: 12 July 2021 Published: 13 August 2021

Citation:

Jiang X, Zhao D, Ali A, Xu B, Liu W, Wen J, Zhang H, Shi Q and Zhang Y (2021) MeiosisOnline: A Manually Curated Database for Tracking and Predicting Genes Associated With Meiosis. Front. Cell Dev. Biol. 9:673073. doi: 10.3389/fcell.2021.673073 Meiosis, an essential step in gametogenesis, is the key event in sexually reproducing organisms. Thousands of genes have been reported to be involved in meiosis. Therefore, a specialist database is much needed for scientists to know about the function of these genes quickly and to search for genes with potential roles in meiosis. Here, we developed "MeiosisOnline," a publicly accessible, comprehensive database of known functional genes and potential candidates in meiosis (https://mcg. ustc.edu.cn/bsc/meiosis/index.html). A total of 2,052 meiotic genes were manually curated from literature resource and were classified into different categories. Annotation information was provided for both meiotic genes and predicted candidates, including basic information, function, protein–protein interaction (PPI), and expression data. On the other hand, 165 mouse genes were predicted as potential candidates in meiosis using the "Greed AUC Stepwise" algorithm. Thus, MeiosisOnline provides the most updated and detailed information of experimental verified and predicted genes in meiosis. Furthermore, the searching tools and friendly interface of MeiosisOnline will greatly help researchers in studying meiosis in an easy and efficient way.

Keywords: meiosis, manual curation, data mining, Greed AUC Stepwise, database

BACKGROUND

Meiosis, the process to generate daughter cells with an intact, haploid genome through one round of DNA replication followed by two rounds of cell division, is a basic feature of sexual reproductive organisms (Gerton and Hawley, 2005; Miller et al., 2013; Bolcun-Filas and Handel, 2018; Biswas et al., 2021). Compared with mitosis, meiosis is characterized by homologous chromosome separation, which ensures the genetic integrity of all daughter cells (Sato et al., 2021). A series of biological processes would take place during meiosis prophase I to guarantee the formation and repair of programmed meiotic DNA double-strand breaks (DSBs) and the pair and synapsis between homologous chromosomes, as well as the formation of meiotic crossovers (Handel and Schimenti, 2010; Baudat et al., 2013; Gray and Cohen, 2016; Ranjha et al., 2018; Jiao et al., 2020; Li et al., 2021).

With the development of genomic technologies on model organisms and recent advances of transcriptomics and proteomics, tremendous articles have been published on meiosis from different species, and we get a clearer understanding about the genetic control of key events in

MeiosisOnline Tracking and Predicting Meiotic Genes

meiosis (Watanabe et al., 2001; Wang et al., 2009; Chalmel and Rolland, 2015; Chen et al., 2018). However, information about meiotic genes is widely fragmented, which makes it still difficult to illuminate/highlight genes, molecular complexes, and/or signaling pathways involved in meiosis. What is more, it is still challenging to identify novel meiotic genes, especially in mammalian meiosis, since genetic modification in model organisms is time-consuming and is like a gamble sometimes (Khan et al., 2018, 2020; Huang et al., 2019; Xie et al., 2019; Yousaf et al., 2020). Thus, a specialist database that can provide integrated annotation of meiotic genes and predict novel functional genes is urgently needed.

Here, we report a publicly accessible, comprehensive database, MeiosisOnline.¹ It is the first resource that is not only a wellstructured repository of experimentally verified meiotic genes with detailed annotation, but also a powerful tool to predict genes that may function in meiosis.

MATERIALS AND METHODS

Manual Curation of Literature

To collect the information of meiotic genes, specific keywords were used to search in PubMed (**Supplementary Table 1**). Then all the collected papers were curated manually and genes that had been validated by experiments were deemed as functional meiotic genes.

Gene Expression Data Collection

Gene expression information was retrieved from the ArrayExpress database.² Datasets from Affymetrix GeneChip platform were downloaded and were divided into different categories, including "developmental stages," "gene disturbance," "before and after treatment," and "tissues and cell types" (**Supplementary Table 2**). Gene expression data combined with category information are provided as annotation information in MeiosisOnline and applied for predicting genes that may function in meiosis.

Annotation

Annotation information for each gene in MeiosisOnline contains "basic information," "function annotation and classification," "protein-protein interaction (PPI) and gene expression."

- Basic information: gene name/synonyms, nucleotide sequences, etc., were extracted from GenBank³ and UniProt Knowledgebase.⁴
- (2) Function annotation and classification: detailed functional information is also manually collected from literature reports. (i) Which meiotic stage is the gene involved? (ii) Did the gene function in one sex or both sexes? (iii) Whether deletion or mutation of the gene in

model organism has a phenotype in fertility? (iv) Which protein complex of the gene is involved? (v) The cellular location and expression pattern in tissues or cell lines. (vi) Experimental methods used for functional analysis. (vii) The information of related literature and figures for illustrating the function of protein/gene. (viii) Gene ontology annotation for collected genes.

(3) Protein-protein interaction and gene expression: both verified and predicted PPI information were provided. Gene expression pattern in reproductive system was also provided graphically.

Implementation

To execute more jobs in parallel, a Dell 730 server with LAMP architecture is used to host the MeiosisOnline database. The server is equipped with 128 GB RAM and two 12-core Intel processors (2.2 GHz). The jQuery is used to render the interface and Python and R languages are employed to supply the backend.

RESULTS

The Manual Curation of Meiosis-Related Genes From the Literature

MeiosisOnline is aiming to construct a functional annotation pipeline about meiosis-associated genes from published articles. After keywords querying in PubMed, about 45,000 research articles published before January 1, 2021, were collected. All collected papers were manually curated, and functional meiotic genes are only included as those with functional experimental validation (Supplementary Table 3). In total, 2,052 unique meiotic genes with experimentally verified functions from 84 species were curated along with functional information in MeiosisOnline. We found that the functional meiotic genes are firstly derived from mice, which accounts for 28.74% of the total reported genes, followed by human (5.16%) and rat (5.07%). Furthermore, other species comprise the rest of 61.03% (Supplementary Table 4). To be noted that the genes always have preferably expression profiles, for example, Mlh3 (mg0000873) expresses during both male and female meiosis, Sun5 (mg0000693) only presents in male germ cells, while Bmp15 (mg0000982) is specially expressed in oocytes.

The Overall Framework of MeiosisOnline

MeiosisOnline is developed in a user-friendly manner and the major functional modules of the database (Figure 1) include:

Search Page

(1) Advanced search. Users can query up to three keywords and set up different combination by selecting the operators

¹https://mcg.ustc.edu.cn/bsc/meiosis/index.html

²http://www.ebi.ac.uk/arrayexpress/

³http://www.ncbi.nlm.nih.gov/genbank

⁴http://www.uniprot.org/

Users could find their interested genes using the Search page.⁵ Four additional search options were also provided⁶:

⁵http://mcg.ustc.edu.cn/bsc/meiosis/search.html

⁶http://mcg.ustc.edu.cn/bsc/meiosis/advanced_search.html



("and," "or," or "exclude") to find the information more specifically (**Supplementary Figure 1A**).

- (2) BLAST search. After protein sequence (FASTA format) uploading, it could map identical and homologous proteins recorded in MeiosisOnline database (**Supplementary Figure 1B**).
- (3) Orthologous search. Simple orthologous search is specified in finding orthologs for designated genes, while advanced orthologous search is to display all orthologs between two selected species (**Supplementary Figure 1C**).
- (4) Chromosome location. This module could list all genes in a given genome region (**Supplementary Figure 1D**).

Browse Page

In the Browse page, users can browse genes by classifications or species⁷ (**Supplementary Figure 2A**). Users can get all the genes belonging to a certain category in a tabular form. For example, users can browse MeiosisOnline Genes (MGs) collected from knockout mice (**Supplementary Figure 2B**) or MG genes identified in different species, e.g., *Homo sapiens* (**Supplementary Figure 2C**).

Candidates Page

In the Candidates page, MeiosisOnline lists all the predicted functional genes in mouse⁸ (**Supplementary Figure 3A**). Clicking the MG ID, detailed information for a candidate gene can be seen (**Supplementary Figure 3B**).

⁷https://mcg.ustc.edu.cn/bsc/meiosis/browse.html

Feedback Page

Users can submit suggestions about the records integrated in MeiosisOnline or submit novel verified meiotic gene information to our database.⁹

MeiosisOnline Integrates Information of Functional Genes in Meiosis

Besides the general information including gene ID, protein ID, taxonomy ID, and basic descriptions, MeiosisOnline also provides high-quality functional annotation for the collected genes. Based on the function annotation information, the experimental verified genes were classified into different categories. Additionally, figures and/or tables illustrating the function of the collected meiotic genes were also incorporated. Moreover, manually annotated functions, signaling pathways, and associated protein complexes of the collected meiotic genes were provided (**Figure 2**). The functional distribution of these genes in various stages of meiosis and fecundity is also listed (**Supplementary Tables 4, 5**).

For instance, using "*Stra8*" for query, results will be listed in tabular form, including MeiosisOnline Gene ID (MG ID), gene names, UniProt ID, etc. (**Figure 2A**). Once clicking the MG ID (MG0001089), detailed information for mouse *Stra8* is available that includes the following: (1) basic information (gene name, nucleotide and protein sequences, etc.), (2) functional annotation and classification from related literature (developmental stages, experimental methods, literature abstract, relevant figures, etc.), and (3) PPI and gene expression information (**Figure 2B**).

⁸http://mcg.ustc.edu.cn/bsc/meiosis/prediction.php

⁹http://mcg.ustc.edu.cn/bsc/meiosis/feedback.html

ME.	IOSIS OS Browse Search	NL.I Candida	T N ntes	E PG	* * * Ustrc * @ MCG (G) 1,1(C) Documentation						
eiosis Online@											
ey words: Meiosis , Reproductive,	Genes/Proteins				ACTI	DUD	na	CC TO	RT.		+
earch: 😣					WELIC	<i>SIS</i>	UJV		V.L	PDA G	MCG
Search the Meio Search, Orthologous se	arch, please click Advanced Options	to display adva	nced op	tion items.	Home Bro≀ eiosis Online⊛	wse Sea	arch	Candidates	Feedba	KK (AL (G	ocumentation
				Key	words: Meiosis , Reproductive, Genes	Proteins					
Example	Submit			MG	ID 🕆 Status Gene Name	es Uniprot ID	Species	Gender	Development sta	ge Fecundity	Experiment method
					1089 * 5400	STRAS MOUFF	Mus muscolur	Male Fermio	Spermatogonia	Reduced familie	
dvanced Ontions	Click the right ices to expand advanced search	ob ontions		nigooo	3005	STINI6_MOUSE	Mus musculus	Male Penale	Oocyte	Reduced fertility	, KO
mg0001089 and Gene Name Gene Synonym/Allas	s spatind in May 20, 2014 Stead	ŝ	Ŷ	E. L. Anderson, A. E. Balton, H. L. F	Roepers-Gajadien, T. J. Hassold, J	D. G. de Rooij, A. M. van	1 Pelt and D. C. Pag	ge			
mg0001089 and a second	stepiend in May 20, 2014 Stead Standard by retinoic acid gene 8 protein STRAE MOTISE PHOTO Acidotyc		Î	E. L. Anderson, A. E. Baltus, H. L. F. (2008) Stra8 and its inducer, retinoic Proc Natl Acad Sci U S A 105(59): 1	Roepers-Gajadiem, T. J. Hausold, J. acid, regulate meiotic imitiation i 4976-80. PMID: [18799751]	D. G. de Rooij, A. M. van in both spermatogenesis at	a Pelt and D. C. Pa and oogenesis in mi	ge ice.			
mg0001089 and a second	tropienel in May 20, 2014 Straß Simulated by retinoic acid gene 8 protein STRAB_MOUSE P70278 Q496Q5 T27287 BC107046 BC101509 RP107433			E. L. Anderson, A. E. Baltus, H. L. F (2008) Straß and its inducer, retinoic Proc Natl Acad Sci U.S. A 105(39): 1	Roepers-Gajsdien, T. J. Hassold, J. Add, ergulate meiotic initiation i 4976-80. PMID: [18799751]	D. G. de Rooij, A. M. van n both spermatogenesis as	1 Pelt and D. C. Paj nd cogenesis in mi	ge ice.			
mg0001089 Gens Name Gens Synonym/Allas Protein Name Protein Name UnilProt Accession Genbank Protein ID	speined in Ney 30, 3014 Straß Straß Straß Straßkeiter of the strain of t	103592.1		E. L. Anderson, A. E. Baltus, H. L. F. (2003) Straß and its indicace, retinoic <i>Proc Natl Acad Sci U S A</i> 105(39): 1 E. L. Anderson, A. E. Baltus, H. L. F.	Roepers-Gajadien, T. J. Hassold, J. 4976-80. PMID: [18799751] Roepers-Gajadien, T. J. Hassold, J.	D. G. de Rooij, A. M. van n both spermatogenesis as D. G. de Rooii, A. M. van	a Pelt and D. C. Paj nd cogenesis in mi	ge ice.			
mg0001089 200 Gene Name Gene Synonym/Allas Protein Synonym/Allas UniProt Accession Genbank Nucleotide ID Genbank Protein ID Organism	Strail Strail Strail Strail Strail Strail Strail Strail Strail Strail STRAE MOUSE Pr0278 Q46005 Z7327 DC100746 BC103591 C013591 CA4997111 AA1039111 AA1 Mar marchink (Mone) Mar marchink (Mone)	103592.1		E. L. Anderson, A. E. Baltus, H. L. F. (2008) Strað and its inducer, retinoic Proc Natl. Acad Sei U.S.A. 105 (59): 1 E. L. Anderson, A. E. Baltus, H. L. F. (2008) Strað and its inducer, retinois Dens Natl. Jack F. 118 (2007)	Roepers-Gajadien, T. J. Hassold, I day, regulate meloitic imitation i 4976-80. PMID: [18799751] Roepers-Gaiadien, T. J. Hassold, I Basic Information	D. G. de Rooij, A. M. van n both spermatogenesis a D. G. de Rooii, A. M. van Function annotatio	a Pelt and D. C. Paj and oogenesis in mi a Pelt and D. C. Paj an and classifica	ge ice. ation PPI at	nd gene expressio	n	
mg0001089 cm Gene Name Gene Synonyms/Allas Protein Synonyms/Allas UniFrot Accession Genbank Nucleotide ID Genbank Protein ID Organism	speined in May 30, 2014 Sea 8 Sea 8 Stimulated by retanoic acid gene 8 protein STRA8_MOUSE P70278_Q496Q5 Z73287_BC1007046_BC103591 CAA997111_AA1007471_AA1035911_AA1 Mar macrular (Mesue) 10999	103592.1		 E. L. Anderson, A. E. Baltus, H. L. F. (2008) Straß and its inducer, retinoic <i>Proc Natl Acad Sci U S A</i> 105(39): 1 E. L. Anderson, A. E. Baltus, H. L. F. (2008) Straß and its inducer, retinoic <i>Proc Natl Acad Sci U S A</i> 105(39): 1 	Roepers-Gajaden, T. J. Hassold, acid, regulate meiotic imitation i 4976-60. PAIID: [18799751] Roepers-Gaiadien, T. J. Hassold, Basic Information	D. G. de Rooij, A. M. van n both spermatogenesis a D. G. de Rooii. A. M. van Function annotatio	1 Pelt and D. C. Pag and cogenesis in mi 1 Pelt and D. C. Par on and classfic:	ge iice. are ation PPI ar	nd gene expressio	n	
mg0001089 Les Gene Symonyms/Allas Gene Symonyms/Allas Protein Name Tordein Symonyms/Allas UniProt Accession Genbank Nucleotide ID Genbank Nucleotide ID Genbank Protein ID Organism NCEII Taxonomy ID Basic Information	stradium lin May 30, 2014 Straß Straß Straß Straßhold by retinoic acid gene 8 protein STRAB_MOUSE P70278 Q496Q5 Z73287 BC100746 BC103590 BC103591 CAA6971.11 AA100747.1 AA1055911 AA1 Mar mucrular (Monue) 10090 Function annotation and classification	100592.1 PPI and g		 E. L. Anderson, A. E. Baltus, H. L. F. (2008) Straß and its inducer, retinoic Proc Natl Acad Sci U S A 105(39): 1 E. L. Anderson, A. E. Baltus, H. L. F. (2008) Straß and its inducer, retinoic Proc Natl Acad Sci U S A 105(39): 1 	Ropers-Gajaden, T. J. Hasold, I acid, regulate neiotic initiation is 4976-80. PMID. [18799751] Basic Information The verified Protein-Protei	D. G. de Rooij, A. M. van n both spermatogenesis at D. G. de Rooii, A. M. van Function annotatio Click the right arr in	1 Pelt and D. C. Pag and cogenesis in mi 1 Pelt and D. C. Pag n and classfic: row for more info	ge ice. ation PPI at ormation. String dat	nd gene expressio	n	
mg0001089 Gess Name Gess Synonymi/Allas Protein Name Protein Synonymi/Allas UniFrot Accession Genbank Arceledide ID Genbank Protein ID Genbank Protein ID Organism NCBI Taxonomy ID Basic Information	Stand in May 20, 2014 Standard by retinoic acid gene 8 protein Standard by retinoic acid gene 8 protein STRA8_MOUSE P70278 Q496Q5 Z73287 BC100746 BC103590 BC103591 CAA697111 AA1007471 AA1035911 AA1 Mar muculu (Mone) 1000 Function annotation and classification	003592.1 PPI and g		 E. L. Anderson, A. E. Baltus, H. L. F. (2009) Stra3 and its inducer, retinoic Proc Natl Acad Sci U S A 105(39): 1 E. L. Anderson, A. E. Baltus, H. L. F. (2008) Stra3 and its inducer, retinoic Proc Natl Acad Sci U S A 105(39): 1 E. L. Anderson, A. E. Baltus, H. L. F. 	Roepers-Gajadien, T. J. Hassold, J. acid, regulate meiotic initiation i 4976-40. PMID: [18799751] Roepers-Galactien, T. J. Hassold, J. Basic Information The verified Protein-Protei Interaction information	D. G. de Rooij, A. M. van n both spermatogenesis au D. G. de Rooii. A. M. van Function annotatio (Click the right arr UmPre	t Pelt and D. C. Pa and cogenesis in mi Pelt and D. C. Pa n and classfic trow for more infc	ge acc. arc arc arc arc arc arc arc arc arc arc	nd gene expressio inbase O nbol	n	Databases
mg0001089 Constant of the second seco	repetends in May 30, 2014 Stank Stank Stan	1003592.1 PPI and g		 E. L. Anderson, A. E. Baltus, H. L. F. (2005) Straß and its inducer, retimoic Proc Natl Acad Sci U S A 105(39): 1 E. L. Anderson, A. E. Baltus, H. L. F. (2008) Straß and its inducer, retimois Proc Natl Acad Sci U S A 105(39): E. L. Anderson, A. E. Baltus, H. L. (2008) Straß and its inducer, retimois Proc Natl Acad Sci U S A 105(39): 	Roepers-Gajodien, T. J. Hassold, J. Add, regulate meiotic initiation i 4976-50. PMID: [18799751] Roepers-Gaisdien, T. J. Hassold I Basic Information The verified Protein-Proteinteraction information	D. G. de Rooij, A. M. van n both spermatogenesis a D. G. de Rooii, A. M. van Function annotatio Click the right ar UmPre	1 Pett and D. C. Pan md cogenesis in mi 1 Pett and D. C. Pan on and classifier row for more infe ot	ge cice. are are are are are are are are are are	nd gene expressio	n Refi	Databases
mg0001089 Come Name Gene Synonymi/Alias Protein Name Protein Name Protein Synonymi/Alias UniProt Accession Genbank Protein ID Genbank Protein ID Organism NCBI Taxonomy ID Basici Information Development stage Gender		00592.1 PPI and g		 E. L. Anderson, A. E. Baltus, H. L. F. (2003) Straß and 16 inducer, retinoic Proc Natl Acad Sci U S A 105(39): 1 E. L. Anderson, A. E. Baltus, H. L. F (2003) Straß and 16 inducer, retinoic Proc Natl Acad Sci U S A 105(39): 1 E. L. Anderson, A. E. Baltus, H. L. (2003) Straß and 16 inducer, retinoic Proc Natl Acad Sci U S A 105(39): 1 	Reepen-Gajadien, T. J. Hassold, J. Acader, egulate meiotic initiation is 4976-50. PMID: [18799751] Reepens-Gaiadien, T. J. Hassold Basic Information The verified Protein-Prote interaction information	D. G. de Rooij, A. M. van n both spermatogenesis a D. G. de Rooii, A. M. van Function annotatio Click the right arr Click the right arr	1 Pelt and D. C. Paj and cogenesis in mi a Pelt and D. C. Pat and classfic row for more infector ot	ge cice. are ation PPI at Gene Syn ormation. String da	nd gene expressio labase.) nbol labase.)	n	Databases
mg0001089 and a second	spectral to May 30, 2014 Staal Sta	003592.1 PPI and g		 E. L. Anderson, A. E. Baltus, H. L. F. (2003) Straß and its indicace, retinoic Proc Natl Acad Sci US A 105(39): 1 E. L. Anderson, A. E. Baltus, H. L. F (2008) Straß and its indicece, retinois Proc Natl Acad Sci US A 105(39): 1 E. L. Anderson, A. E. Baltus, H. L.: (2008) Straß and its indicece, retinois Proc Natl Acad Sci US A 105(39): 1 	Roepers-Gajadien, T. J. Hassold, J. Ada, regulate meiotic initiation in 4976-80. PMID: [18799751] Basic Information The verified Protein-Protein Interaction information	D. G. de Rooij, A. M. van n both spermatogenesis at Function annotatio Click the right arr Click the right arr	a Pelt and D. C. Pa and cogenesis in mi Pelt and D. C. Pa on and classfic row for more info tow for more info UniProt	ee attion PPI a ormation. String da ormation. String da	nd gene expressio nbase. nbol nabase.	n Refl	Databases
mg0001089 Left Gene Name Gene Synonym/Allas Protein Name Protein Name Protein Synonym/Allas UnilProt Accession Genbank Protein ID Genbank Protein ID Genbank Protein ID Basic Information Development stage Gender Fecundity Experiment Methods	spearade to May 30, 3014 State	103592.1 PPFI and g		 E. L. Anderson, A. E. Baltus, H. L. F. (2003) Straß and its indicace, retinoic <i>Proc Natl Acad Sci U S A</i> 105(39): 1 E. L. Anderson, A. E. Baltus, H. L. F. (2003) Straß and its indicace, retinoic <i>Proc Natl Acad Sci U S A</i> 105(39): 1 E. L. Anderson, A. E. Baltus, H. L.; (2003) Straß and its indicace, retinoic <i>Proc Natl Acad Sci U S A</i> 105(39): 1 E. L. Anderson, A. E. Baltus, H. L.; 	Roepers-Gajalien, T. J. Hassold, J. Aud, regulate meiotic initiation in 4976-80. PMID: [18799751] Roepers-Gaiadien, T. J. Hassold I Basic Information The verified Protein-Protein Interaction Information	D. G. de Rooij, A. M. van n both spermatogenesis at D. G. de Rooii. A. M. van Function annotatio Click the right arr UmiPro Click the right arr	a Pelt and D. C. Paj and cogenesis in mi a Pelt and D. C. Pan on and classificit row for more info ot uniProt E	ge acc. ation PPI at Gene Syn ormation. String da Hugo0	nd gene expressio tabase. nbol tabase. Gene Symbol tal	n 	Databases
mg0001089 Left Gene Name Cene Synonym/Allas Gene Synonym/Allas Protein Name Protein Name Cenbank Nucleotide ID Genbank Nucleotide ID Genbank NetBi Taxonomy ID Basic Information Development stage Gender Fecundity Experiment Methods Protein Complex Taxonomic		1055921 PPI and g		 E. L. Anderson, A. E. Baltus, H. L. F. (2008) Straß and its inducer, retinoic Proc Natl. Acad Sci U S A 105(39): 1 E. L. Anderson, A. E. Baltus, H. L. F. (2008) Straß and its inducer, retinoic Proc Natl. Acad Sci U S A 105(39): 1 E. L. Anderson, A. E. Baltus, H. L. (2008) Straß and its inducer, retinoic Proc Natl Acad Sci U S A 105(39): 1 E. L. Anderson, A. E. Baltus, H. L. (2008) Straß and its inducer, retinoic Proc Natl Acad Sci U S A 105(39): 1 	Repers-Gajatien, T. J. Hassold, I acid, regulate meiotic mitiation is 4976-80. PAID: [18799751] Basic Information The verified Protein-Protei Interaction Information	D. G. de Rooij, A. M. van n both spermatogenesis a D. G. de Rooij, A. M. van Function annotatio Click the right arr UniPro Click the right arr H550A, MOUSE BRX1 MOUSE	1 Pelt and D. C. Paj and cogenesis in mi a Pelt and D. C. Pai n and classifici row for more info tow for more info UniProt E	ge ace ation PPI a Gene Syn Grmation. String dar Huppo. Broto.	nd gene expressio abase. nbol tabase. Gene Symbol aal	n	Databases
mg0001089 Ceae Name Gene Synonyms/Allas Protein Name Protein Name Combark Succession Combark Succession Genbark Succession Organism NCHI Taxonomy ID Basic Information Perviopment stage Gender Facundity Experiment Methods Protein Complex Description	toparnel in May 20, 2014 Start Start Start Start Start Start Start Start Start Start Start Start Start Start Start Start Start Start Start Start Start Start	803592.1 PPI and g		 E. L. Anderson, A. E. Baltus, H. L. F. (2009) Straß and its inducer, retinoic Proc Natl Acad Sci U S A 105(39): 1 E. L. Anderson, A. E. Baltus, H. L. F. (2008) Straß and its inducer, retinoic Proc Natl Acad Sci U S A 105(39): 1 E. L. Anderson, A. E. Baltus, H. L. (2008) Straß and its inducer, retinois Proc Natl Acad Sci U S A 105(39): 1 E. L. Anderson, A. E. Baltus, H. L. (2008) Straß and its inducer, retinois Proc Natl Acad Sci U S A 105(39): 1 	Corpers-Gajolien, T. J. Hausold, J. acid, regulate motoic initiation i 4976-40. PMID: [18799751] Recents-Galodien, T. J. Hausold, J. Basic Information The verified Protein-Protein Interaction Information	D. G. de Rooij, A. M. van n both spermatogenesis at D. G. de Rooii. A. M. van Function annotatio Click the right arr UniPre Click the right arr HS80A_MOUSE STRLI MOUSE STRLI MOUSE	a Pelt and D. C. Pag and cogenesis in mi a Pelt and D. C. Pan and classification tow for more info ot UniProt E	ge ice. are ation PPI at Gene Syn Grens String dat Hap60 Brix1 Sacto	nd gene expressio nbol tabase. Gene Symbol nal	n	Databases Ref Databases
mg0001089 Ceee Name Gene Synonymi/Allas Protein Name Protein Name Cenbank Nucleotide ID Genbank Nucleotide ID Organism NCHI Taxonomy ID Basic Information Development stage Cender Facundity Experiment Methods Protein Complex Description	repennel to May 30, 304 Staral Stara	003992.1		 E. L. Anderson, A. E. Baltus, H. L. F. (2005) Straß and its inducer, retinoic Proc Natl Acad Set U S A 105(39): 1 E. L. Anderson, A. E. Baltus, H. L. F. (2008) Straß and its inducer, retinoic Proc Natl Acad Set U S A 105(39): 1 E. L. Anderson, A. E. Baltus, H. L.: (2003) Straß and its inducer, retinoic Proc Natl Acad Set U S A 105(39): 1 E. L. Anderson, A. E. Baltus, H. L.: (2003) Straß and its inducer, retinoic Proc Natl Acad Set U S A 105(39): 1 	Ropers-Gajolien, T. J. Hassold, J. Add, regulate meiotic initiation i 4976-50. PMID: [18799751] Ropers-Gaistien, T. J. Hassold I Basic Information The verified Protein-Protein Interaction information	D. G. de Rooii, A. M. van n both spermatogenesis a D. G. de Rooii, A. M. van Function annotatio Click the right arr UniPr Click the right arr UniPr Click the right arr HSS0A_MOUSE BRX1_MOUSE STR10_MOUSE ARIAN MOVEN	a Pelt and D. C. Pa and cogenesis in mi a Pelt and D. C. Pa on and classifier row for more info ot UniProt E E	ge ecc. FPPI al ormation. String da Gene Syn Hup900 Brnd. Sta10 Aratas	nd gene expressio tabase. mbol tabase. Gene Symbol aal	n Ref1	Databases Ref Databases
rn g 0001089		003592.1 PPI and g based of the second of th		 E. L. Anderson, A. E. Baltus, H. L. F. (2003) Straß and its inducer, retinoic Proc Natl Acad Sci U S A 105(39): 1 E. L. Anderson, A. E. Baltus, H. L. F (2003) Straß and its inducer, retinoic Proc Natl Acad Sci U S A 105(39): 2 E. L. Anderson, A. E. Baltus, H. L. (2003) Straß and its inducer, retinois Proc Natl Acad Sci U S A 105(39): 1 E. L. Anderson, A. E. Baltus, H. L. (2003) Straß and its inducer, retinois Proc Natl Acad Sci U S A 105(39): 1 E. L. Anderson, A. E. Baltus, H. L. (2003) Straß and its inducer, retinois Proc Natl Acad Sci U S A 105(39): 1 M. Mark, H. Jacobs, M. Oulad-Abd 	Reepers-Gajadien, T. J. Hassold, J. Adv. egulate meiotic initiation in 4976-50. PMID: [18799751] Basic Information The verified Protein-Protein Interaction information Other Protein-Protein Interaction resources	D. G. de Rooij, A. M. van n both spermatogenesis a D. G. de Rooii, A. M. van Function annotatio Click the right arr Click the right arr Click the right arr HIS90A_MOUSE BRX1_MOUSE STR10_MOUSE	a Pett and D. C. Pay and cogenesis in mi a Pett and D. C. Pao on and classific row for more info ot UniProt E E E E E E	ge cice. PPPI at formation. String da Gene Syn Hup90 Brici Sts.10 Sts.10 Sts.10 Sts.10 Sts.10	nd gene expressio nabase. • nabase. • Gene Symbol aal	n Ref 10090 10090 10090 10090	Databases Ref Databases
mg0001089 Cene Name Cene Synonyms/Allas Protein Name Protein Nucleotide ID Genbank Protein ID Organism NCHI Taxonomy ID Basic Information Development stage Cender Focundity Experiment Methods Protein Complex Description	Separate in May 20, 2014 Standard by retinoic acid gene 8 protein Situated by retinoic acid gene 8 protein CAA97111 AAD07471 AAD07471 AAD07471 AAI Add marcular (Mosure) CAA97111 AAD07471 AAD07471 AAI Add marcular (Mosure) Separategonia Occyte Separategonia Occyte Robused femily Ko Co Add Female Separategonia of oncose embryonic gene cell than the sex chromosome constitution of the gr (14550793] B Add Female Add Female Add Female Add Female Add Female Separategonia of oncose embryonic gene cell than the sex chromosome constitution of the gr (14550793) B Add Female Add Fe	003592.1 PPI and g baseline of factors the senate of factors prem cell. However		 E. L. Anderson, A. E. Baltus, H. L. F. (2003) Straß and its indicace, retinoic Proc Natl Acad Sci U S A 105(39): 1 E. L. Anderson, A. E. Baltus, H. L. F (2003) Straß and its indicec, retinoic Proc Natl Acad Sci U S A 105(39): 1 E. L. Anderson, A. E. Baltus, H. L.: (2003) Straß and its indicec, retinoic Proc Natl Acad Sci U S A 105(39): 1 E. L. Anderson, A. E. Baltus, H. L.: (2003) Straß and its indicec, retinoic Proc Natl Acad Sci U S A 105(39): 1 M. Mark, H. Jacobs, M. Oxlad-Abd B. Ghyneinek (2005) STRAS-deflic Grupprogene conjenention J. Four Sci 	Roopers-Gajadien, T. J. Hassold, J. Agot, regulate meiodic initiation in 4976-50. PMID: [18799751] Roopers-Gajadien, T. J. Hausold, J. Basic Information The verified Protein-Protein Interaction Information Other Protein-Protein Interaction resources	D. G. de Rooij, A. M. van n both spermatogenesis as D. G. de Rooii, A. M. van Function annotatio Click the right arr UniPro Click the right arr UniPro Click the right arr UniPro STK1_MOUSE	a Pett and D. C. Paj and cogenesis in mi a Pett and D. C. Pa on and classfic: row for more info to to to the page of the page to the page of the page of the page to the page of the page of the page of the page to the page of the page of the page of the page to the page of the page of the page of the page of the page to the page of the page of the page of the page of the page to the page of the page	ee eice. attion PPI at Gene Syn ormation. String da ormation. String da Brixd Brixd Brixd Sk10 An484 Sk20 Carebo	nd gene expressio nabase. O nabol Gene Symbol aal	n Ref1 10090 10090 10090 10090 10090	Databases Ref Databases
mg0001089 Cens Name Gens Synonyms/Allas Protein Name Protein Name Protein Synonyms/Allas UniProt Accession Genbank Protein ID Genbank Protein ID Genbank Protein ID Genbank Protein ID Genbank Protein ID Genbank Protein ID Basic Information Development stage Gender Facundity Exprotein Complex Description		103592.1		 E. L. Anderson, A. E. Baltus, H. L. F. (2003) Straß and its indicace, retinoi: <i>Proc Natl Acad Sci U S A</i> 105(39): 1 E. L. Anderson, A. E. Baltus, H. L. F. (2003) Straß and its indicace, retinoi: <i>Proc Natl Acad Sci U S A</i> 105(39): 1 E. L. Anderson, A. E. Baltus, H. L.: (2003) Straß and its indicace, retinoi: <i>Proc Natl Acad Sci U S A</i> 105(39): 1 E. L. Anderson, A. E. Baltus, H. L.: (2003) Straß and its indicec, retinoi: <i>Proc Natl Acad Sci U S A</i> 105(39): 1 E. L. Anderson, A. E. Baltus, H. L.: (2003) Straß and its indicec, retinoi: <i>Proc Natl Acad Sci U S A</i> 105(39): 1 M. Mark, H. Jacobs, M. Oualad-Abd B. Ghyselinck (2008) STRA8-defic chromosome condensation. <i>J Cell S</i> 	Repers Gajstien, T. J. Hassold, 1 acid, regulate neiotic imitation in Arrow of the Analysis of the Analysis of the Analysis Basic Information Basic Information The verified Protein-Protein Interaction Information Other Protein-Protein Interaction resources	D. G. de Rooij, A. M. van hoth spermatogenesis as D. G. de Rooii. A. M. van Function annotatio Click the right arr in Click the right arr INSUMATION Click the right arr INSUMATION RRX1_MOUSE STR10_MOUSE	a Pelt and D. C. Pa and cogenesis in mi a Pelt and D. C. Pa n and classificit row for more info to to to more info to for more info to to to more info to to to more info to to to the pelt of the pelt to the pelt of the pelt of the pelt to the pelt of the pelt of the pelt of the test of the pelt of the pelt of the pelt of the test of the pelt of the pelt of the pelt of the pelt of the test of the pelt of the pelt of the pelt of the pelt of the test of the pelt of the test of the pelt of the pel	ere are are are are are are are are are	nd gene expressio nbase. Inbase. Gene Symbol aal	n Ref 1 10090 10090 10090 10090 10090 10090	Databases
mg0001089 Gene Name Gene Synonymi/Allas Protein Name Protein Name Genbark Nachedrig Genbark Nachedrig Genbark Nachedrig Genbark Nachedrig Basic Information Perodopment stage Gender Facundity Experiment Methods Protein Complex Description		103592.1 PPI and g has make of four promo cell. However in Gene Catherine		 E. L. Anderson, A. E. Baltus, H. L. F. (2009) Straß and its inducer, retimoic <i>Proc Natl Acad Sci U S A</i> 105(39): 1 E. L. Anderson, A. E. Baltus, H. L. F. (2003) Straß and its inducer, retimoic <i>Proc Natl Acad Sci U S A</i> 105(39): 1 E. L. Anderson, A. E. Baltus, H. L. (2003) Straß and its inducer, retimoir <i>Proc Natl Acad Sci U S A</i> 105(39): 1 E. L. Anderson, A. E. Baltus, H. L. (2003) Straß and its inducer, retimoir <i>Proc Natl Acad Sci U S A</i> 105(39): 1 E. L. Anderson, A. E. Baltus, H. L. (2003) Straß and its inducer, retimoir <i>Proc Natl Acad Sci U S A</i> 105(39): 1 M. Mark, H. Jacobs, M. Oulad-Abd B. Ghyvelinck (2005) STRA8-defic chromosome condensation. <i>J Cell S</i> 	Stepper-Gajadien, T. J. Hassold, I. acid, regulate meiotic initiation i 4976-40. PAID: [18799751] Basic Information Basic Information The verified Protein-Protein Interaction Information Other Protein-Protein Interaction resources	D. G. de Rooij, A. M. van hoth spermatogenesis as D. G. de Rooii, A. M. van Function annotatio Click the right arr UniPro Click the right arr UniPro Click the right arr HSS0A_MOUSE STR1_MOUSE STR1_MOUSE STD8_MOUSE ARHB_MOUSE ARHB_MOUSE	a Pelt and D. C. Pag and cogenesis in mi a Pelt and D. C. Paia on and classificit row for more info to for more info UniProt E E E E E E E	ge ice. PPI at cornation String da cornation. String da cornation. String da string cornation. String da cornation. String da da da da da da da da da da da da da	nd gene expressio tabase. O tabase. O Gene Symbol aal	n Ref1 10090 10090 10090 10090 10090 10090 10090 10090 10090	Databases Ref Databases
mg0001089 Ceneo Name Geneo Synonymi/Allas Protein Name Protein Name Combark Xectoris ID Genbank Xectoris ID Organism NCBI Taxonomy ID Basic Information Development stage Gender Facundity Experiment Methods Protein Complex Description	Imputed is May 20, 2014	1035921 PPI and g mail or final prime of However the as male or final prime of However an Grass Cellection		 E. L. Anderson, A. E. Baltus, H. L. F. (2009) Straß and its inducer, retinoic <i>Proc Natl Acad Set U S A</i> 105(39): 1 E. L. Anderson, A. E. Baltus, H. L. F. (2008) Straß and its inducer, retinoic <i>Proc Natl Acad Set U S A</i> 105(39): 1 E. L. Anderson, A. E. Baltus, H. L. (2009) Straß and its inducer, retinoic <i>Proc Natl Acad Set U S A</i> 105(39): 1 E. L. Anderson, A. E. Baltus, H. L. (2009) Straß and its inducer, retinoic <i>Proc Natl Acad Set U S A</i> 105(39): 1 E. L. Anderson, A. E. Baltus, H. L. (2009) Straß and its inducer, retinoic <i>Proc Natl Acad Set U S A</i> 105(39): 1 M. Mark, H. Jacoba, M. Oulad-Abd M. Mark, H. Jacoba, M. Oulad-Abd 	Corpers Gajolien, T. J. Hassold, I. acid, regulate meloic initiation i 4076-40. PMID: [18799751] Recents Gaiadien, T. J. Hassold, I. Basic Information The verified Protein-Protein Interaction information Other Protein-Protein Interaction resources View Microarray data	D. G. de Rooij, A. M. van n both spermatogenesis at D. G. de Rooii. A. M. van Function annotatio Click the right ar UniPre Click the right ar UniPre Click the right ar UniPre Click the right ar HS80A_MOUSE BRX1_MOUSE BRX1_MOUSE SETD#_MOUSE ARH#_MOUSE SETD#_MOUSE ARH#_MOUSE Click to view expre	s Pelt and D. C. Pag and cogenesis in mi si Pelt and D. C. Pau is Pelt and D. C. Pau tow for more info ot UniProt E E E E E E E E E E E	ee ation PPI a ormation. String da formation. String da bring da b	nd gene expressio tabase. • nbol Gene Symbol aal	n Ref1 10090 10090 10090 10090 10090 10090 10090 10090	Databases Ref Databases

FIGURE 2 | The search function of MeiosisOnline. (A) Users can simply input gene "*Stra8*" for querying. The results are shown in a tabular format. Users can visualize the detailed information by clicking on the MeiosisOnline ID (MG0001089). (B) The detailed information for mouse gene "*Stra8*." The information presented has been manual checked and will be updated based on newly published data.

MeiosisOnline Facilitates the Discovery of Functional Meiotic Genes

To expand the utilization of our MeiosisOnline database, a prediction model was constructed and used to predict the candidate functional meiotic genes. As mouse is one of the best studied animal models, the GAS algorithm (Zhang et al., 2013) was used to predict potential meiotic functional genes from *Mus musculus* (Supplementary Figure 4).

To verify the efficiency of GAS, we randomly separated the training data into two equal parts: one as a new training dataset and the other as a testing dataset. The model was split into three stages: stage 1 was constructed with features of the category "developmental stages," stage 2 included features from the categories "tissues and cell types" and "before and after treatment," and stage 3 is the features of the category "gene disturbance." As shown in **Supplementary Figure 5**, the performances of GAS models were better when more features were added to them. Then, based on the experimentally verified meiotic genes and gene expression data, 590 mouse genes with experimentally verified function in meiosis were used as the positive training dataset. The negative training dataset contained 5,868 genes from MGI (Mouse Genome Informatics),¹⁰

¹⁰ http://www.informatics.jax.org/



of which knockout mice did not have any abnormalities in the reproductive system. Three hundred and ninety-four features used for GAS construction and prediction were extracted from the 85 microarray data (**Supplementary Table 2**).

Ultimately, 165 candidate genes (GAS probability > 0.5) having potential role were sorted out and annotated in MeiosisOnline (see text footnote 8). For the candidate genes, information that implicate their function in meiosis, including gene expression, protein localization, structure, and protein interactions, are included in MeiosisOnline.

We also performed GO enrichment analysis on both literature-reported genes in *M. musculus* (Supplementary Table 6) and GAS-predicted candidate genes (Supplementary Table 7). Compared with whole genome data, we statistically calculated the distribution of MGs in cellular components, biological processes, and molecular functions by R (hypergeometric distribution, p < 0.05) (Yu et al., 2012). Among all GO terms, 118 in biological processes, 1 in molecular functions, and 32 in cellular components are overlapped in both sets of reported and predicted genes (Figure 3A). Considering that predicted genes enriched by overlapped GO term have more potential in regulating meiosis, interestingly, meiotic cell cycle (GO: 0051321) was enriched from both sets of genes (reported and predicted) (Supplementary Table 8).

Furthermore, as genes are mostly regulated through network structure in meiosis, we mapped out the PPIs among all of the genes and constructed a potential meiosis network with 1,083,566 reported PPIs among 26,569 proteins in MeiosisOnline. For example, *Cct5* and *Sf3a3*, which have not been reported, show

very high connectivity with reported genes (**Figure 3B**). Further investigation of *Cct5* and *Sf3a3* would disclose how these two genes interacted with reported genes and what is the function of these interactions. Moreover, we also found that reported genes like *Ccna2* interacted wildly with predicted genes (**Figure 3C**).

DISCUSSION

Studies on animal models, especially genetically modified mice, have revealed many critical regulators involved in meiosis (Marston and Amon, 2004; Handel and Schimenti, 2010; Robert et al., 2016; Jiang et al., 2017, 2018); however, the information for these meiotic genes are scattered among thousands of papers. Thus, it is difficult to collect and compare the information of meiotic genes among different species from papers. Here, based on manual curation of meiosis-related genes from the literature, the first comprehensive database, MeiosisOnline, focusing on meiosis was developed.

As the fundamental process of gametogenesis, systematical annotation for meiotic genes is important to conduct further experiment study. Currently, only a few databases provided information related to meiosis, with limited features. Some databases were only repositories of gene expression data such as GermOnline 4.0 (Lardenois et al., 2010), SpPress (Vibranovski et al., 2009), and GermSAGE (Lee et al., 2009), and the utilization of these databases to obtain valuable information regarding experimentally verified function is not satisfactory. Some of those are limited to a specific species or a certain biological process, such as SpPress that focused only on spermatogenesis in *Drosophila* (Vibranovski et al., 2009), ReCGiP that focused on reproduction in pig (Yang et al., 2010), and MeioBase that focused on meiotic genes in plant (Li et al., 2014). In our study, MeiosisOnline provided detailed and comprehensive information and annotation of the meiotic genes, including basic information, functional annotation and classification, PPI, and gene expression, etc. With these diverse information, users can easily access the detailed functional information of meiosis-associated genes.

Additionally, besides the proven functional meiotic genes, among the 2,300 genes that are predominantly expressed in the testis (Schultz et al., 2003), the function of most genes in meiosis are still unclear. In MeiosisOnline, based on literaturereported genes and genome-wide transcriptional data from ArrayExpress analysis, 165 genes (GAS probability > 0.5) in mouse are predicted to be involved in meiosis. As we know that homologous recombination is the basic feature of meiosis (Zickler and Kleckner, 2015), when we perform GO terms for the predicted genes by MeiosisOnline, double-strand break repair *via* homologous recombination (GO: 0000724) was one of the most enriched GO (**Supplementary Table 8**), which implies that these predicted genes may function during meiosis and further functional study by animal models should be conducted.

What is more, MeiosisOnline could conduct the study focus on complex molecular and/or signaling networks in meiosis. During meiosis, some genes are regulated through the network structure; for example, the deletion of *Hadc1* or *Ddac2* alone did not affect meiosis, while their combined deletion resulted in meiotic arrest and subsequent oocyte depletion (Ma et al., 2012). Hence, mapping out PPIs among known and predicted genes is useful in uncovering the novel regulating mechanism of meiosis.

In summary, MeiosisOnline is the first specialist database on meiosis. It not only provides comprehensive information for experimental verified meiotic genes but can also predict genes that may function in meiosis. It would be a helpful resource for researchers to gain a new insight in meiosis.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

YZ and QS conceived and supervised the project. XJ, HZ, AA, and WL collected the data from the literature and

REFERENCES

Baudat, F., Imai, Y., and De Massy, B. (2013). Meiotic recombination in mammals: localization and regulation. *Nat. Rev. Genet.* 14, 794–806. doi: 10.1038/nrg 3573 ArrayExpress, as well as positive and negative training data for the prediction of potential meiotic functional genes. DZ developed the web interface. XJ, DZ, YZ, and QS wrote the manuscript. JW reviewed the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This project was supported by the National Key Research and Developmental Program of China (2017YFC1001500, 2018YFC1003700, 2016YFC1000600, and 2018YFC1004700), the National Natural Science Foundation of China (31890780, 31630050, 31871514, and 82071709), and the Fundamental Research Funds for the Central Universities (YD2070002006 and YD2070002020).

ACKNOWLEDGMENTS

The authors thank the Bioinformatics Center of the University of Science and Technology of China, School of Life Science, for providing supercomputing resources. The authors also thank Dr. Xin Du in the 901st Hospital, Hefei, China, for his help in the data collection and analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2021. 673073/full#supplementary-material

Supplementary Figure 1 | The advanced options of MeiosisOnline. (A) Advanced search which allows users to simultaneously input three terms for querying; (B) BLAST search used for protein sequence querying; (C) orthologous search. Browse orthologs for a gene in all species by simple search or browse orthologs of all genes in two species by advanced search. An example of pairwise orthologous browsing in human and mice were showed; (D) searching the MG genes based on chromosome and/or chromosomal location.

Supplementary Figure 2 | (A) All entries of MeiosisOnline can be browsed through taxonomic hierarchy and manual classifications; (B) by MG genes collected from KO mouse model, and (C) by MG genes identified in *Homo sapiens*.

Supplementary Figure 3 | (A) The candidate genes that are predicted by GAS model; (B) the detailed information of the predicted example gene "mg0002027."

Supplementary Figure 4 | The pseudo code illustrating the GAS algorithm.

Supplementary Figure 5 | The performance of GAS regarding its sensitivity (*Sn*), specificity (*Sp*), accuracy (*Ac*), Matthews correlation coefficient (MCC), and area under cover (AUC). Increase in the number of selected features resulted in improved AUC value.

- Biswas, L., Tyc, K., El Yakoubi, W., Morgan, K., Xing, J., and Schindler, K. (2021). Meiosis interrupted: the genetics of female infertility via meiotic failure. *Reproduction* 161, R13–R35.
- Bolcun-Filas, E., and Handel, M. A. (2018). Meiosis: the chromosomal foundation of reproduction. *Biol. Reprod.* 99, 112–126. doi: 10.1093/biolre/ioy021

- Chalmel, F., and Rolland, A. D. (2015). Linking transcriptomics and proteomics in spermatogenesis. *Reproduction* 150, R149–R157.
- Chen, Y., Zheng, Y., Gao, Y., Lin, Z., Yang, S., Wang, T., et al. (2018). Single-cell RNA-seq uncovers dynamic processes and critical regulators in mouse spermatogenesis. *Cell Res.* 28, 879–896. doi: 10.1038/s41422-018-0074-y
- Gerton, J. L., and Hawley, R. S. (2005). Homologous chromosome interactions in meiosis: diversity amidst conservation. *Nat. Rev. Genet.* 6, 477–487. doi: 10.1038/nrg1614
- Gray, S., and Cohen, P. E. (2016). Control of meiotic crossovers: from doublestrand break formation to designation. Annu. Rev. Genet. 50, 175–210. doi: 10.1146/annurev-genet-120215-035111
- Handel, M. A., and Schimenti, J. C. (2010). Genetics of mammalian meiosis: regulation, dynamics and impact on fertility. *Nat. Rev. Genet.* 11, 124–136. doi: 10.1038/nrg2723
- Huang, Z., Khan, M., Xu, J., Khan, T., Ma, H., Khan, R., et al. (2019). The deubiquitinating gene Usp29 is dispensable for fertility in male mice. *Sci. China Life Sci.* 62, 544–552. doi: 10.1007/s11427-018-9469-4
- Jiang, H., Gao, Q., Zheng, W., Yin, S., Wang, L., Zhong, L., et al. (2018). MOF influences meiotic expansion of H2AX phosphorylation and spermatogenesis in mice. *PLoS Genet.* 14:e1007300. doi: 10.1371/journal.pgen.1007300
- Jiang, L., Li, T., Zhang, X., Zhang, B., Yu, C., Li, Y., et al. (2017). rpl10l is required for male meiotic division by compensating for RPL10 during meiotic sex chromosome inactivation in mice. *Curr. Biol.* 27, 1498–1505.e1496.
- Jiao, Y., Fan, S., Jabeen, N., Zhang, H., Khan, R., Murtaza, G., et al. (2020). A TOP6BL mutation abolishes meiotic DNA double-strand break formation and causes human infertility. *Sci. Bull* 65, 2120–2129.
- Khan, M., Jabeen, N., Khan, T., Hussain, H. M. J., Ali, A., Khan, R., et al. (2018). The evolutionarily conserved genes: Tex37, Ccdc73, Prss55 and Nxt2 are dispensable for fertility in mice. *Sci. Rep.* 8:4975.
- Khan, R., Ye, J., Yousaf, A., Shah, W., Aftab, A., Shah, B., et al. (2020). Evolutionarily conserved and testis-specific gene, 4930524B15Rik, is not essential for mouse spermatogenesis and fertility. *Mol. Biol. Rep.* 47, 5207–5213. doi: 10.1007/s11033-020-05595-0
- Lardenois, A., Gattiker, A., Collin, O., Chalmel, F., and Primig, M. (2010). GermOnline 4.0 is a genomics gateway for germline development, meiosis and the mitotic cell cycle. *Database* 2010:baq030. doi: 10.1093/database/baq030
- Lee, T. L., Cheung, H. H., Claus, J., Sastry, C., Singh, S., Vu, L., et al. (2009). GermSAGE: a comprehensive SAGE database for transcript discovery on male germ cell development. *Nucleic Acids Res.* 37, D891–D897.
- Li, H., Meng, F., Guo, C., Wang, Y., Xie, X., Zhu, T., et al. (2014). MeioBase: a comprehensive database for meiosis. *Front. Plant Sci.* 5:728. doi: 10.3389/fpls. 2014.00728
- Li, Y., Wu, Y. F., Jiang, H. W., Khan, R., Han, Q. Q., Iqbal, F., et al. (2021). The molecular control of meiotic double-strand break (DSB) formation and its significance in human infertility. *Asian J. Androl.*
- Ma, P., Pan, H., Montgomery, R. L., Olson, E. N., and Schultz, R. M. (2012). Compensatory functions of histone deacetylase 1 (HDAC1) and HDAC2 regulate transcription and apoptosis during mouse oocyte development. *Proc. Natl. Acad. Sci. U.S.A.* 109, E481–E489.
- Marston, A. L., and Amon, A. (2004). Meiosis: cell-cycle controls shuffle and deal. Nat. Rev. Mol. Cell Biol. 5, 983–997. doi: 10.1038/nrm1526
- Miller, M. P., Amon, A., and Unal, E. (2013). Meiosis I: when chromosomes undergo extreme makeover. *Curr. Opin. Cell Biol.* 25, 687–696. doi: 10.1016/ j.ceb.2013.07.009
- Ranjha, L., Howard, S. M., and Cejka, P. (2018). Main steps in DNA doublestrand break repair: an introduction to homologous recombination and related processes. *Chromosoma* 127, 187–214. doi: 10.1007/s00412-017-0658-1

- Robert, T., Nore, A., Brun, C., Maffre, C., Crimi, B., Bourbon, H. M., et al. (2016). The TopoVIB-Like protein family is required for meiotic DNA double-strand break formation. *Science* 351, 943–949. doi: 10.1126/science.aad5309
- Sato, M., Kakui, Y., and Toya, M. (2021). Tell the difference between mitosis and meiosis: interplay between chromosomes, cytoskeleton, and cell cycle regulation. *Front. Cell Dev. Biol.* 9:660322. doi: 10.3389/fcell.2021.66 0322
- Schultz, N., Hamra, F. K., and Garbers, D. L. (2003). A multitude of genes expressed solely in meiotic or postmeiotic spermatogenic cells offers a myriad of contraceptive targets. *Proc. Natl. Acad. Sci. U.S.A.* 100, 12201–12206. doi: 10.1073/pnas.1635054100
- Vibranovski, M. D., Lopes, H. F., Karr, T. L., and Long, M. (2009). Stage-specific expression profiling of Drosophila spermatogenesis suggests that meiotic sex chromosome inactivation drives genomic relocation of testis-expressed genes. *PLoS Genet.* 5:e1000731. doi: 10.1371/journal.pgen.1000731
- Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* 10, 57–63. doi: 10.1038/nrg 2484
- Watanabe, Y., Yokobayashi, S., Yamamoto, M., and Nurse, P. (2001). Pre-meiotic S phase is linked to reductional chromosome segregation and recombination. *Nature* 409, 359–363. doi: 10.1038/35053103
- Xie, Y., Khan, R., Wahab, F., Hussain, H. M. J., Ali, A., Ma, H., et al. (2019). The testis-specifically expressed Dpep3 is not essential for male fertility in mice. *Gene* 711:143925. doi: 10.1016/j.gene.2019.06.015
- Yang, L., Zhang, X., Chen, J., Wang, Q., Wang, L., Jiang, Y., et al. (2010). ReCGiP, a database of reproduction candidate genes in pigs based on bibliomics. *Reprod. Biol. Endocrinol.* 8:96. doi: 10.1186/1477-7827-8-96
- Yousaf, A., Wu, Y., Khan, R., Shah, W., Khan, I., Shi, Q., et al. (2020). Normal spermatogenesis and fertility in Ddi1 (DNA damage inducible 1) mutant mice. *Reprod. Biol.* 20, 520–524. doi: 10.1016/j.repbio.2020.08.006
- Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16, 284–287. doi: 10.1089/omi.2011.0118
- Zhang, Y., Zhong, L., Xu, B., Yang, Y., Ban, R., Zhu, J., et al. (2013). SpermatogenesisOnline 1.0: a resource for spermatogenesis based on manual literature curation and genome-wide data mining. *Nucleic Acids Res.* 41, D1055–D1062.
- Zickler, D., and Kleckner, N. (2015). Recombination, pairing, and synapsis of homologs during meiosis. *Cold Spring Harb. Perspect. Biol.* 7:a016626. doi: 10.1101/cshperspect.a016626

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Jiang, Zhao, Ali, Xu, Liu, Wen, Zhang, Shi and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.