

## DATABASE

## ASER: Animal Sex Reversal Database



Yangyang Li<sup>1,#</sup>, Zonggui Chen<sup>2,#</sup>, Hairong Liu<sup>3,#</sup>, Qiming Li<sup>1</sup>, Xing Lin<sup>2</sup>, Shuhui Ji<sup>3</sup>,  
 Rui Li<sup>3</sup>, Shaopeng Li<sup>1</sup>, Weiliang Fan<sup>1</sup>, Haiping Zhao<sup>3</sup>, Zuoyan Zhu<sup>2</sup>, Wei Hu<sup>2,4,\*</sup>,  
 Yu Zhou<sup>1,5,\*</sup>, Daji Luo<sup>2,3,4,\*</sup>

<sup>1</sup>State Key Laboratory of Virology, College of Life Sciences, Wuhan University, Wuhan 430072, China

<sup>2</sup>State Key Laboratory of Freshwater Ecology and Biotechnology, Institute of Hydrobiology, Innovation Academy for Seed Design, Chinese Academy of Sciences, Wuhan 430072, China

<sup>3</sup>School of Basic Medical Science, Wuhan University, Wuhan 430071, China

<sup>4</sup>University of Chinese Academy of Sciences, Beijing 100049, China

<sup>5</sup>Frontier Science Center for Immunology and Metabolism, Wuhan University, Wuhan 430072, China

Received 6 January 2021; revised 25 October 2021; accepted 1 November 2021

Available online 25 November 2021

Handled by Zhang Zhang

**Abstract** Sex reversal, representing extraordinary sexual plasticity during the life cycle, not only triggers reproduction in animals but also affects reproductive and endocrine system-related diseases and cancers in humans. Sex reversal has been broadly reported in animals; however, an integrated resource hub of sex reversal information is still lacking. Here, we constructed a comprehensive database named ASER (Animal Sex Reversal) by integrating sex reversal-related data of 18 species from teleostei to mammalia. We systematically collected 40,018 published papers and mined the sex reversal-associated genes (SRGs), including their regulatory networks, from 1611 core papers. We annotated homologous genes and computed conservation scores for whole genomes across the 18 species. Furthermore, we collected available RNA-seq datasets and investigated the expression dynamics of SRGs during sex reversal or sex determination processes. In addition, we manually annotated 550 *in situ* hybridization (ISH), fluorescence *in situ* hybridization (FISH), and immunohistochemistry (IHC) images of SRGs from the literature and described their spatial expression in the gonads. Collectively, ASER provides a unique and integrated resource for researchers to query and reuse organized data to explore the mechanisms and applications of SRGs in animal breeding and human health. The ASER database is publicly available at <http://aser.ihb.ac.cn/>.

**KEYWORDS** Sex reversal; Sex reversal-associated gene; Database; Omics; Conservation

## Introduction

Sex determination mechanisms in animals mainly include genetic sex determination (GSD) and environmental sex

determination (ESD) [1]. In GSD, the primary sex of organisms is determined by genetics during fertilization, while organisms with ESD remain bipotential gonads until they perceive environmental stress to promote sex differentiation during ontogeny [2]. For many years, it was a dogma in vertebrates in the field of sex determination that sex would be fixed for life after primary sex determination. After sex reversal was first reported in *Aplocheilichthys latipes*

\*Corresponding authors.

E-mail: [huwei@ihb.ac.cn](mailto:huwei@ihb.ac.cn) (Hu W), [yu.zhou@whu.edu.cn](mailto:yu.zhou@whu.edu.cn) (Zhou Y),

[luodaji@ihb.ac.cn](mailto:luodaji@ihb.ac.cn) (Luo D).

#Equal contribution.

Peer review under responsibility of Beijing Institute of Genomics, Chinese Academy of Sciences / China National Center for Bioinformation and Genetics Society of China. <https://doi.org/10.1016/j.gpb.2021.10.001>

1672-0229 © 2021 The Authors. Published by Elsevier B.V. and Science Press on behalf of Beijing Institute of Genomics, Chinese Academy of Sciences / China National Center for Bioinformation and Genetics Society of China.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

and natural sex reversal was found in *Monopterus javanensis* [3], it has been widely accepted that sex determination is amazingly plastic in vertebrates, especially in fish. This plasticity shows that sexual fate is not an irreversible process. Indeed, this reversible process leads to sex reversal, a redirection of the sexual phenotype during development [4]. Environmental factors can override genetic factors to redirect sexual fate in fish [5] and reptiles [6]. Sex reversal has been found to be driven by diverse factors, such as genetic factors, hormones, temperature, and social changes [7]. Unlike sex change, which implies a transition from the stabilized sex to the opposite sex, sex reversal occurs during gonadal development, including the initiation phase and maintenance phase of sex determination [4].

Specifically, sex reversal has been studied in fish, reptiles, birds, amphibians, and even mammals. In fish, types of gonadal differentiation are roughly divided into two groups: hermaphroditic and gonochoristic [5]. Hermaphroditic species undergo sex reversal during their lifetime and include three strategies: female-to-male (protogynous), male-to-female (protandrous), or bidirectional (serial) sex change [8]. Taking *Monopterus albus* as an example, an individual is female from the embryonic stage to first sexual maturity, then enters an intersex state, and later develops into a male [9]. Additionally, some hermaphroditic species undergo socially cued female-to-male sex reversal, whereby the removal of the dominant male induces sex reversal in a resident female, such as *Thalassoma bifasciatum* [10]. Among gonochoristic fish, sex reversal is a synergistic result of both GSD and ESD [11]. For example, *Cynoglossus semilaevis* is a gonochoristic fish with a female heterogametic sex determination system ( $ZW_{\text{♀}}/ZZ_{\text{♂}}$ ) characterized by GSD and temperature-dependent sex determination (TSD; a subclass of ESD) [12]. In many reptiles, including *Trachemys scripta*, gonadal sex is determined by the environmental temperature during egg incubation [13]. However, estrogens, including estradiol-17 $\beta$ , have also been proven to participate in the sex determination of *Trachemys scripta* [14]. Sex reversal in birds, such as *Gallus gallus*, is mainly related to alterations in sex steroid hormone action, especially estrogens [15]. Amphibians also show plasticity in sex determination, influenced by estrogens, androgens [16], and sometimes by temperature [17]. Sex determination in mammals has been reported to depend on three processes: chromosome determination (XX or XY), appropriate pathway of gonadal differentiation, and accurate development of secondary sexual characteristics [18]. Disrupting any of these three steps of gonadal differentiation can lead to aberrant sex determination. In *Homo sapiens*, the frequencies of XX and XY sex reversal are 1/20,000 and 1/3000, respectively, and most of these cases are caused by translocations of *SRY* [19]. Although

sex reversal has been broadly reported among vertebrates, the molecular events underlying sex reversal remain poorly understood, limited by the lack of integrated omics data across species.

Although there are several reproduction-related resources, such as GUDMAP [20], GonadSAGE [21], and ReproGenomics Viewer [22], an integrated and dedicated database for the community studying sex determination and differentiation is missing. The GUDMAP database is a comprehensive gene expression dataset of the developing genitourinary system in mouse with both *in situ* and microarray data. GonadSAGE is a Serial Analysis of Gene Expression (SAGE) database for male embryonic gonad development in mouse. The ReproGenomics Viewer is a cross-species database of omics data (*e.g.*, RNA-seq and ChIP-seq) for tissues related to reproduction, such as gametogenesis, in 9 model organisms. Here, we developed the Animal Sex Reversal (ASER) database, the first functional genomics hub for sex reversal to our best knowledge. The main works of ASER can be roughly divided as follows. 1) We screened 18 important and typical species with sex reversal phenomena from teleostei to mammalia, including *Cyprinus carpio*, *Danio rerio*, *Oryzias latipes*, *Oreochromis niloticus*, *Epinephelus coioides*, *Thalassoma bifasciatum*, *Betta splendens*, *Monopterus albus*, *Lates calcarifer*, *Paralichthys olivaceus*, *Cynoglossus semilaevis*, *Xenopus laevis*, *Trachemys scripta*, *Gallus gallus*, *Homo sapiens*, *Mus musculus*, *Bos taurus*, and *Equus caballus*, and summarized the major inducements of sex reversal or common

**Table 1** Inducements of sex reversal or common approaches used to manipulate sex in 18 species

Species	Gonadal differentiation	Inducement/approach
<i>Cyprinus carpio</i>	Gonochoristic	H/D; Temp
<i>Danio rerio</i>	Gonochoristic	GA; H/D; Temp
<i>Oryzias latipes</i>	Gonochoristic	GA; H/D; Temp
<i>Oreochromis niloticus</i>	Gonochoristic	GA; H/D; Temp
<i>Epinephelus coioides</i>	Hermaphroditic	Nat; GA; H/D; SF
<i>Thalassoma bifasciatum</i>	Hermaphroditic	Nat; SF
<i>Betta splendens</i>	Gonochoristic	H/D
<i>Monopterus albus</i>	Hermaphroditic	Nat
<i>Lates calcarifer</i>	Hermaphroditic	Nat
<i>Paralichthys olivaceus</i>	Gonochoristic	H/D; Temp
<i>Cynoglossus semilaevis</i>	Gonochoristic	Nat; Temp
<i>Xenopus laevis</i>	Gonochoristic	H/D
<i>Trachemys scripta</i>	Gonochoristic	GA; H/D; Temp
<i>Gallus gallus</i>	Gonochoristic	H/D
<i>Homo sapiens</i>	Gonochoristic	GA
<i>Mus musculus</i>	Gonochoristic	GA
<i>Bos taurus</i>	Gonochoristic	GA
<i>Equus caballus</i>	Gonochoristic	GA

Note: Nat, natural; GA, genetic abnormality; H/D, hormone or drug; Temp, temperature; SF, social factor.

approaches used to manipulate sex in these species (Table 1). 2) We compiled a list of the most common genes or drugs related to sex reversal. Then, we collected and analyzed PubMed literature to mine sex reversal-associated genes (SRGs) and obtained their regulatory networks. Meanwhile, we gathered protein–protein interaction (PPI) networks related to SRGs from the STRING database. 3) To facilitate users comparing the homology of SRGs in different species, we collected or assembled the gene annotations for the 18 species, identified homologous genes, computed the basewise conservation scores across these species, and identified conserved motifs for orthologous gene groups. 4) We systematically processed available RNA-seq data and provided gene expression dynamics during sex reversal between females and males or different developmental stages. A user-friendly genome browser was customized to visualize these genome-wide data. 5) We collected and annotated available *in situ* hybridization (ISH), fluorescence *in situ* hybridization (FISH), and immunohistochemistry (IHC) data to display the spatial expression of SRGs in the gonads. In conclusion, our ASER database provides comprehensive and systemic integration of sex reversal-related data, and we believe that this open resource will greatly promote research on the mechanisms of sex reversal.

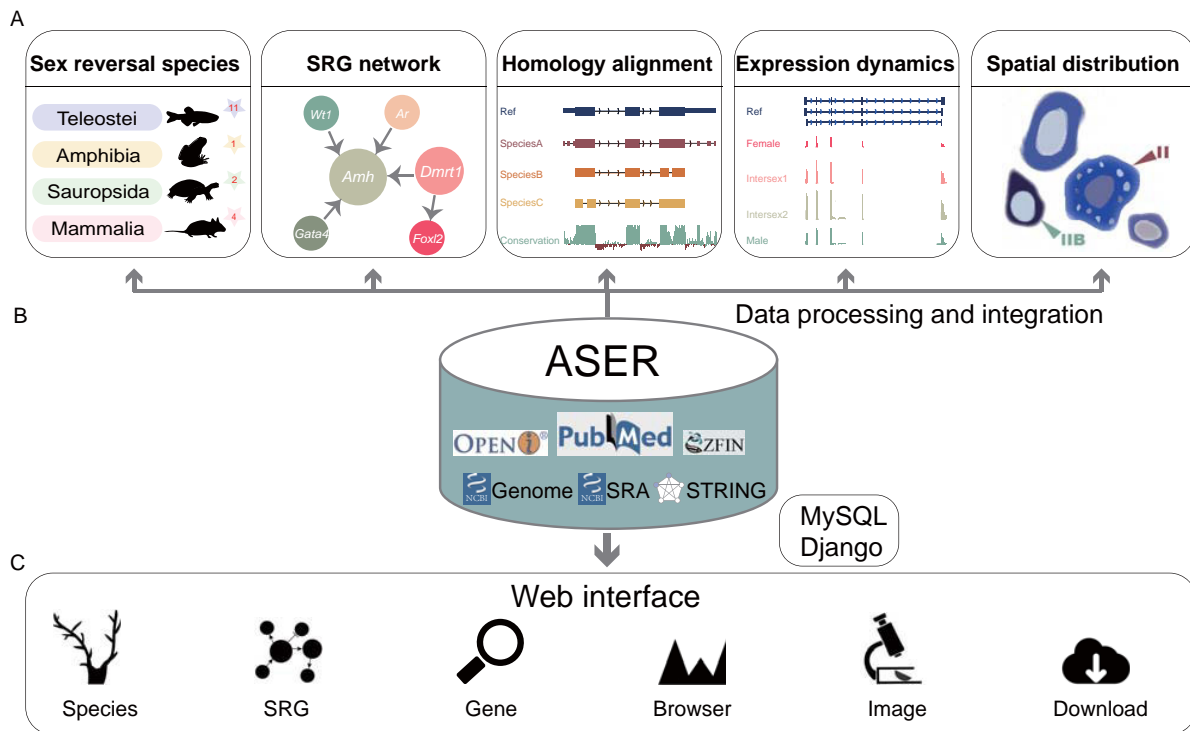
## Data collection and database content

### Framework of ASER

An overview of the ASER database and web server is shown in Figure 1. The ASER database contains five main functional modules: 1) information for 18 sex reversal species, 2) SRG regulatory networks, 3) homology alignment, 4) expression dynamics, and 5) ISH, FISH, and IHC images of SRGs (Figure 1A). The preprocessed data were managed with the MySQL database. Django-based applications were developed to provide a user-friendly interface including an embedded genome browser for visualizing genome-wide data. The key workflows, tools, and processed data are summarized in Figure S1A and B and described in detail below.

### Data sources

SRG information and their regulatory networks were curated from PubMed literature. All genome sequences and species information used in this database were downloaded from NCBI public database. All raw sequencing data were downloaded from the Sequence Read Archive (SRA) of NCBI. The sets of RNA-seq data were organized by species, gonad developmental stages, and temperature (Table S1). In addition,



**Figure 1** Schematic diagram of ASER database

**A.** Five main functional modules in ASER, including 1) information for the 18 sex reversal species, 2) SRGs and their regulatory networks, 3) multiple sequence alignments and conservation scores, 4) gene expression dynamics during sex reversal from RNA-seq data, and 5) spatial distribution of SRGs from ISH, FISH, and IHC images. **B.** Data sources in ASER database. ASER stores all processed data in a MySQL database with additional indexes and uses the Django framework for interactive queries from the web interface to the backend database. **C.** Overview of the ASER web interface. The main functionalities are provided and organized into six modules. ASER, Animal Sex Reversal; SRG, sex reversal-associated gene; ISH, *in situ* hybridization; FISH, fluorescence *in situ* hybridization; IHC, immunohistochemistry.

we retrieved images related to sex reversal from the OPENi (<https://openi.nlm.nih.gov/>) and ZFIN databases [23].

### SRG mining

We retrieved thousands of articles from PubMed by querying species and functional keywords (*e.g.*, sex reversal). First, the abstracts and full texts of these articles were collected by text crawler technology. The full texts of non-open access papers were obtained through the library portal of Wuhan University, China. Next, we separated a chunk of continuous text into separate words, and carried out word stemming to remove plural and different tenses. Then, we removed stop words, such as “the”, “is”, and “however”. Finally, we counted the frequency of the words from the literature, and manually filtered out some high-frequency but irrelevant words, such as “masculinizing”, “ovotestes”, “pseudomale”, “hermaphrodite”, and “gynogenesis”, into a blacklist until most of the high-frequency words were gene symbols and drug names. The remaining words related to genes and drugs were manually added to the wordlist (Table S2).

We retained 1611 papers that contained the words in the wordlist and manually read them with notations about SRG regulation (Table S3). Finally, we found 258 SRGs, 6 drugs, and 11 hormones, which were validated to be functional in sex reversal in different species, and constructed the regulatory networks of SRGs. We next predicted another 498 genes that were homologous with those SRGs in the 18 species (Figure S1C). Furthermore, PPI networks of SRGs were extracted from the STRING database [24].

### RNA-seq data processing

The data quality of the collected RNA-seq data was assessed using FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>), and the adapters and low-quality bases in raw reads were removed using Trim Galore ([http://www.bioinformatics.babraham.ac.uk/projects/trim\\_galore/](http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/)). Filtered reads were aligned to the genome using STAR [25] in end-to-end mode. The primary alignments were retained through SAMtools [26]. Gene expression quantification in fragments per kilobase of exon per million mapped fragments (FPKM) was computed using StringTie [27]. Differential expression analysis was performed using DESeq2 [28].

### Transcriptome assembly

High-quality reads were *de novo* assembled using StringTie [27] with default parameter settings. The longest open reading frames (ORFs) were predicted in the assembled transcripts using TransDecoder.LongOrfs (<https://help.rc.ufl.edu/doc/TransDecoder>). DIAMOND [29] was used to

collect homologous evidence of identified ORFs from the UniProt database (<https://www.uniprot.org/>). The potential coding regions were further refined by TransDecoder.Predict. Finally, a GFF3 file based on the coding regions of the reference genome was generated through the “`cdna_alignment_orf_to_genome_orf.pl`” function in TransDecoder.

### Homology alignment

Orthologous groups of SRGs were identified among all sex reversal species using BLAST [30] all-v-all algorithm in OrthoFinder [31]. Conserved motifs of orthogroups were predicted using MEME [32]. Comparisons between conserved motifs and known motifs were performed using Tomtom [33]. A species tree was constructed according to the species taxonomy on NCBI. Meanwhile, the evolutionary relationship was verified by OrthoFinder.

For the orthologue tracks in a reference species, the homologous genes in other species were mapped to the reference genome using BLAT [34]. Alignments with sequence identity  $\geq 60\%$  were retained, and the maximum intron size was set to 450,000 bp.

For the conservation track, pairwise alignments between genome sequences were built using LASTZ [35], and MULTIZ [36] was then used to construct multiple alignments, based on which the conservation scores were calculated using phyloP from the PHAST package [37].

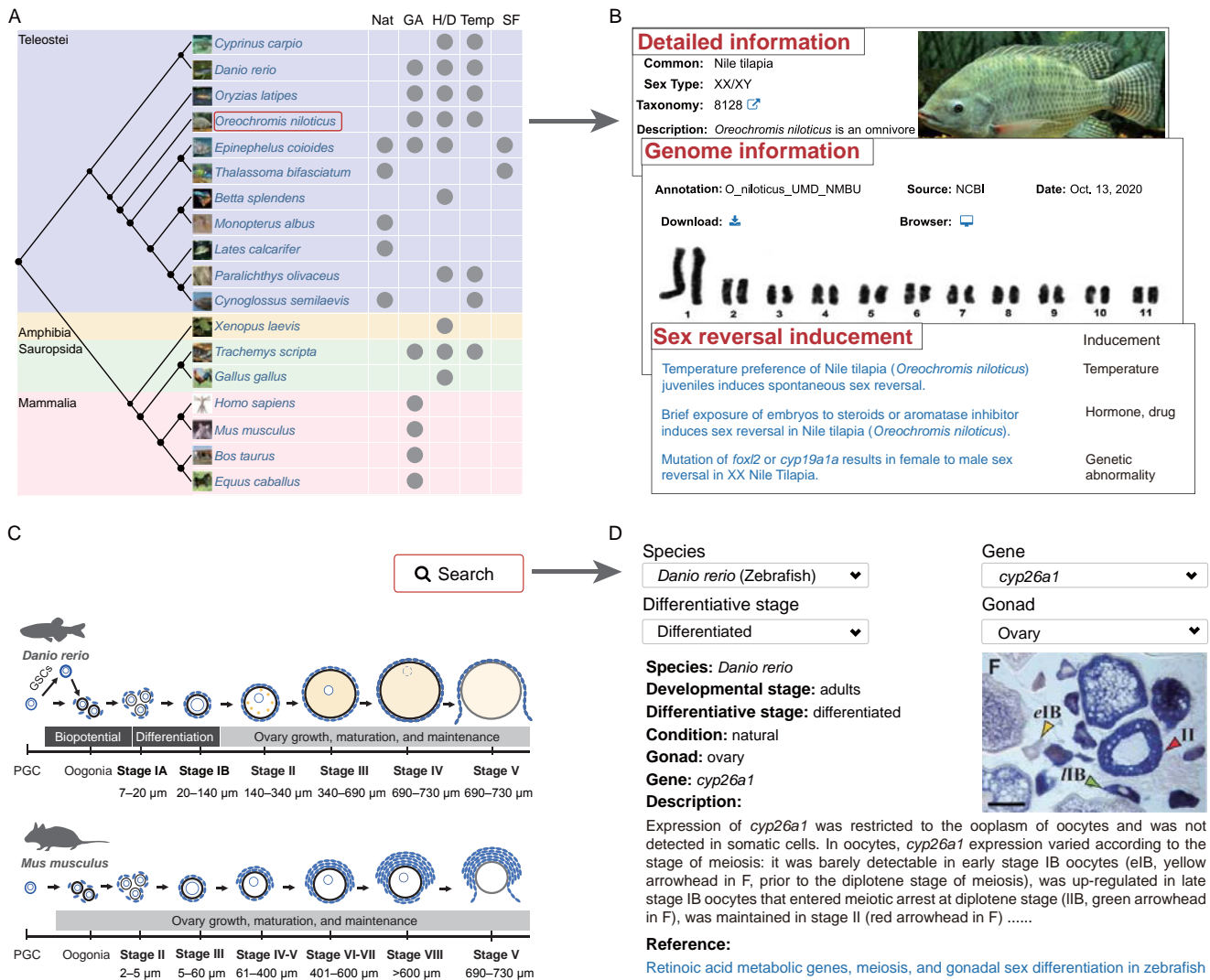
### Image collection and annotation

We collected available ISH, FISH, and IHC data related to SRGs from the OPENi and ZFIN databases. The images were classified by gene, differentiation status, developmental period, and gender. We manually added descriptions for those images based on the original figure legends and articles.

### Web interface and usage

ASER is a user-friendly database, and all the contents are interactive and dynamic. The main functionalities are provided and organized into six modules, including Species, Image, SRG, Gene, Browser, and Download (Figure 1C). In addition, the “Search” module is developed to display and interconnect different kinds of data in other modules.

For the “Species” module, the evolutionary tree constructed for the 18 sex reversal species is displayed on the main page (Figure 2A). The reported inducements of sex reversal or common approaches used to manipulate sex in each species are displayed on this page, including natural processes, genetic abnormality (*e.g.*, *amh* overexpression), administration of exogenous hormones or drugs (*e.g.*, 17 $\alpha$ -methyltestosterone), temperature changes during



**Figure 2** “Species” and “Image” modules in ASER

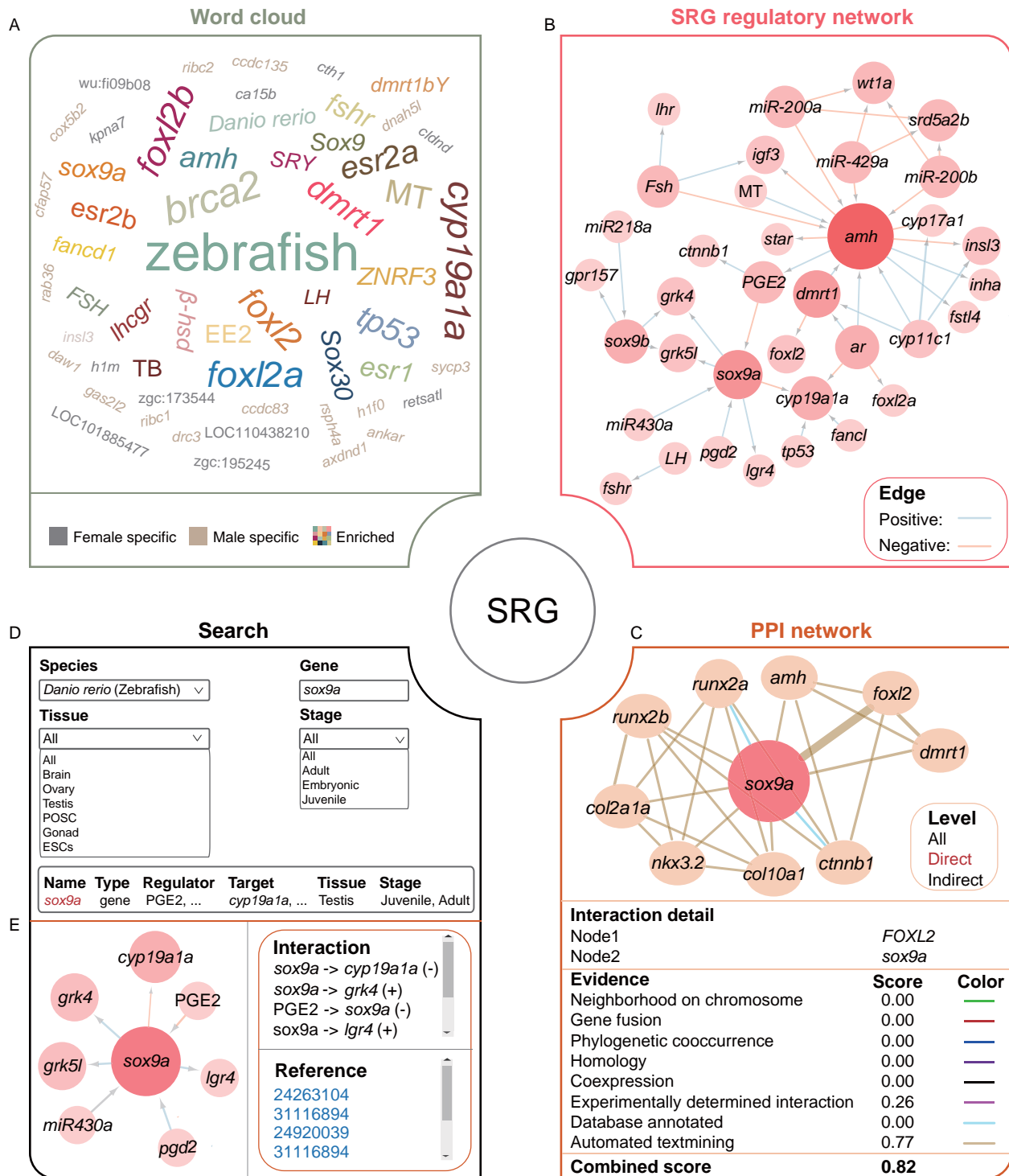
**A.** The main page of “Species” module. Evolutionary relationship and sex reversal inducements or common approaches used to manipulate sex of the 18 species belonging to teleostei, mammalia, sauropsida, and amphibia are shown the main page. **B.** Detailed descriptions, genome information, and references related to sex reversal for each species shown in (A). **C.** The main page of “Image” module showing examples of gonadal morphology at different developmental stages in zebrafish (*Danio rerio*) and mouse (*Mus musculus*). **D.** The search page of “Image” module describing the spatial distribution of SRGs in gonads, with a representative example of *cyp26a1* in zebrafish ovaries. Nat, natural; GA, genetic abnormality; H/D, hormone or drug; Temp, temperature; SF, social factor; GSC, germline stem cell; PGC, primordial germ cell.

gonadal differentiation, and manipulation of social factors. The literature supporting this information is also provided. Users can click on any species to obtain detailed descriptions and genome information for this species (Figure 2B).

For the “Image” module, we summarized the morphological characteristics of zebrafish (*Danio rerio*) ovary and mouse (*Mus musculus*) ovary at different developmental stages to help users better understand the content of this module (Figure 2C). ISH, FISH, and IHC data related to specific SRGs can be queried in different ways by species, gene, differentiative stage, and gonad. Detailed descriptions of images are shown to help users understand the spatial distribution of SRGs in the gonads (Figure 2D).

The “SRG” module includes word cloud, regulatory and PPI networks, and search pages. The word cloud figure is

dynamically presented by species with hyperlinks on the nodes (Figure 3A). When the user clicks one node, the original references and additional actions for more information will be shown under the figure. For any validated SRG, ASER allows users to obtain its regulators (including genes, hormones, and drugs), targets, and the associated modes of regulation (Figure 3B). At the same time, PPI networks of these SRGs in different species are also displayed (Figure 3C), in which the colors of the edges are used to distinguish known interactions (experimentally determined interaction and database annotated), predicted interactions (neighborhood on chromosome, gene fusion, and phylogenetic co-occurrence), and other types (homology, co-expression, and automated text mining). In addition, the search page provides an interface for a specific SRG to



**Figure 3** “SRG” module in ASER

**A.** Word cloud for genes or drugs mined from the literature as well as the top 50 genes that are specifically expressed in females or males. The size represents word frequency. **B.** Representative regulatory network of SRGs in zebrafish. The size and color gradients represent the node degree in the network. **C.** An interactive view of the PPI network for the SRG query. The colors of the edges represent different evidence for associations. **D.** The search page of “SRG” module by species, tissue, gene, and developmental stage. **E.** Detailed information, regulatory relationships, and literature evidence for the SRG query.

show its regulatory network and more detailed information, such as tissue, developmental stage, and literature evidence (Figure 3D and E).

The “Gene” module provides different kinds of data for

any annotated gene, some of which are linked to the “Browser” module for visualization. The links corresponding to the query gene are shown on the search page by species and gene symbol (Figure 4A). Detailed information

for a specific query gene includes its orthogroup in all species, predicted motifs, and similar known motifs (Figure 4B). The orthologous genes and 18-way conservation scores for the query gene can be inspected in “Browser” tracks (Figure 4C). Detailed alignment information can be obtained and downloaded by clicking on the track. In addition, gene expression quantifications in FPKM across different stages, tissues, and conditions are shown as bar plots and in detail as tables (Figure 4D). The RNA-seq signal profiles are displayed in “Browser”, and the tracks can be customized easily, including color, scale, height, *etc.*

For any species, the available tracks can be dynamically selected or unselected. For example, in *Danio rerio*, a subset of RNA-seq tracks are shown for the *sox9* gene during sex reversal (Figure 4E).

## Discussion

Studies on sex reversal have been especially useful in helping redefine the concept of sex determination. There are diverse master sex-determining genes reported in different species.



**Figure 4** “Gene” and “Browser” modules in ASER

**A.** The search page of “Gene” module for any gene in the 18 species. **B.** Homologous genes in the 18 species (top) and motifs corresponding to the query gene (bottom). The conserved motif for these homologous genes is computationally identified by MEME, and known motifs similar to this motif are also presented. **C.** Genome browser tracks of orthologue and conservation scores (18-way). **D.** Expression dynamics of representative query gene. **E.** Genome browser view of processed RNA-seq signals for a representative query gene (*sox9a* in zebrafish). FPKM, fragments per kilobase of exon per million mapped fragments.

In addition, genes previously known to be involved in sex determination or differentiation are emerging as potential key components of sex reversal in other vertebrates [38]. Therefore, studies in different species continue to reveal genes with unexpected roles in sex reversal, and their homologues in other vertebrates also deserve investigation. ASER fills the gap of the sex reversal database by integrating diverse information at different levels for the 18 species with sex reversal phenomena, including curated SRGs, RNA-seq data, image data, and conservation data.

For any collected sex reversal species, users can obtain major inducements of sex reversal in this species in the “Species” module. For any SRG, users can obtain its regulators and targets during sex determination in the “SRG” module, and spatial distribution in different stages in the “Image” module. For any annotated gene, users can obtain its homologous genes and conserved motifs in 18 species in the “Gene” module. Furthermore, users can also explore and visualize expression dynamics across different conditions in the “Gene” or “Browser” module.

In the future, we will continuously select important and typical sex reversal species as their complete genome and omics data from both “female” and “male” samples become available. Hermaphroditic fish such as *Synbranchus marmoratus* and *Amphiprion perideraion* and invertebrates such as *Macrobrachium rosenbergii* and *Venus mercenaria* are candidates. In addition, we will add more omics data, such as sRNA-seq, BS-seq, and ChIP-seq data. We expect that the resources in ASER will promote further studies to decode the molecular mechanisms of sex reversal.

## Data availability

ASER is publicly available at <http://aser.ihb.ac.cn/>.

## CRedit author statement

**Yangyang Li:** Software, Formal analysis, Data curation, Writing - original draft, Writing - review & editing. **Zonggui Chen:** Software, Formal analysis, Visualization, Writing - review & editing. **Hairong Liu:** Investigation, Data curation. Writing - review & editing. **Qiming Li:** Formal analysis. **Xing Lin:** Data curation. **Shuhui Ji:** Data curation. **Rui Li:** Data curation. **Shaopeng Li:** Data curation. **Weiliang Fan:** Visualization. **Haiping Zhao:** Data curation. **Zuoyan Zhu:** Writing - review & editing. **Wei Hu:** Conceptualization, Funding acquisition, Writing - review & editing. **Yu Zhou:** Conceptualization, Project administration, Funding acquisition, Methodology, Writing - review & editing. **Daji Luo:** Conceptualization, Project administration, Funding acquisition, Methodology, Writing - review & editing.

All authors have read and approved the final manuscript.

## Competing interests

The authors declare no conflict of interest.

## Acknowledgments

This work was supported by grants from the National Natural Science Foundation of China (Grant Nos. 31922085 and 31872191 to DL; Grant No. 31922039 to YZ), the Strategic Priority Research Program of Chinese Academy of Sciences (Grant No. XDA24010108 to WH and DL), the China Agriculture Research System of Ministry of Finance (MOF) and Ministry of Agriculture and Rural Affairs (MARA) (Grant No. CARS-46 to WH), and the Natural Science Foundation of Hubei Province, China (Grant No. 2020CFA056 to DL; Grant No. 2020CFA057 to YZ). Part of the computation of this work was done in the Supercomputing Center of Wuhan University and Hydrobiological Data Analysis Center, China.

## Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.gpb.2021.10.001>.

## ORCID

0000-0002-8441-8484 (Yangyang Li)  
 0000-0003-3346-2765 (Zonggui Chen)  
 0000-0002-6458-1786 (Hairong Liu)  
 0000-0003-4183-9913 (Qiming Li)  
 0000-0002-8471-6684 (Xing Lin)  
 0000-0002-4696-9330 (Shuhui Ji)  
 0000-0002-4434-2120 (Rui Li)  
 0000-0001-6415-2562 (Shaopeng Li)  
 0000-0003-4570-7686 (Weiliang Fan)  
 0000-0002-3982-3624 (Haiping Zhao)  
 0000-0002-3337-5912 (Zuoyan Zhu)  
 0000-0002-6475-6427 (Wei Hu)  
 0000-0002-2102-9377 (Yu Zhou)  
 0000-0002-8051-9624 (Daji Luo)

## References

- [1] Shao C, Li Q, Chen S, Zhang P, Lian J, Hu Q, et al. Epigenetic modification and inheritance in sexual reversal of fish. *Genome Res* 2014;24:604–15.
- [2] Budd AM, Banh QQ, Domingos JA, Jerry DR. Sex control in fish: approaches, challenges and opportunities for aquaculture. *J Mar Sci Eng* 2015;3:329–55.
- [3] Mei J, Gui JF. Genetic basis and biotechnological manipulation of sexual dimorphism and sex determination in fish. *Sci China Life Sci* 2015;58:124–36.



- [4] Weber C, Capel B. Sex reversal. *Curr Biol* 2018;28:R1234–6.
- [5] Devlin RH, Nagahama Y. Sex determination and sex differentiation in fish: an overview of genetic, physiological, and environmental influences. *Aquaculture* 2002;208:191–364.
- [6] Holleley CE, O'Meally D, Sarre SD, Marshall Graves JA, Ezaz T, Matsubara K, et al. Sex reversal triggers the rapid transition from genetic to temperature-dependent sex. *Nature* 2015;523:79–82.
- [7] Baroiller JF, D'Cotta H. Environment and sex determination in farmed fish. *Comp Biochem Physiol Part C Toxicol Pharmacol* 2001;130:399–409.
- [8] Todd EV, Liu H, Muncaster S, Gemmell NJ. Bending genders: the biology of natural sex change in fish. *Sex Dev* 2016;10:223–41.
- [9] Chen H, Liu H, Li R, Lin X, Luo H, Ji S, et al. Blood cell identification and hematological analysis during natural sex reversal in rice field eel (*Monopterus albus*). *Aquaculture* 2021;538:736543.
- [10] Thomas JT, Todd EV, Muncaster S, Lokman PM, Damsteegt EL, Liu H, et al. Conservation and diversity in expression of candidate genes regulating socially-induced female-male sex change in wrasses. *PeerJ* 2019;7:e7032.
- [11] Capel B. Vertebrate sex determination: evolutionary plasticity of a fundamental switch. *Nat Rev Genet* 2017;18:675–89.
- [12] Liu J, Liu X, Jin C, Du X, He Y, Zhang Q. Transcriptome profiling insights the feature of sex reversal induced by high temperature in tongue sole *cynoglossus semilaevis*. *Front Genet* 2019;10:522.
- [13] Holleley CE, Sarre SD, O'Meally D, Georges A. Sex reversal in reptiles: reproductive oddity or powerful driver of evolutionary change? *Sex Dev* 2016;10:279–87.
- [14] Crews D, Cantú AR, Rhen T, Vohra R. The relative effectiveness of estrone, estradiol-17 $\beta$ , and estrin in sex reversal in the red-eared slider (*Trachemys scripta*), a turtle with temperature-dependent sex determination. *Gen Comp Endocrinol* 1996;102:317–26.
- [15] Major AT, Smith CA. Sex reversal in birds. *Sex Dev* 2016;10:288–300.
- [16] Nakamura M. Is a sex-determining gene(s) necessary for sex-determination in amphibians? Steroid hormones may be the key factor. *Sex Dev* 2013;7:104–14.
- [17] Wallace H, Badawy GM, Wallace BM. Amphibian sex determination and sex reversal. *Cell Mol Life Sci* 1999;55:901–9.
- [18] Quinn A, Koopman P. The molecular genetics of sex determination and sex reversal in mammals. *Semin Reprod Med* 2012;30:351–63.
- [19] Camerino G, Parma P, Radi O, Valentini S. Sex determination and sex reversal. *Curr Opin Genet Dev* 2006;16:289–92.
- [20] Harding SD, Armit C, Armstrong J, Brennan J, Cheng Y, Haggarty B, et al. The GUDMAP database – an online resource for genitourinary research. *Development* 2011;138:2845–53.
- [21] Lee TL, Li YM, Cheung HH, Claus J, Singh S, Sastry C, et al. GonadSAGE: a comprehensive SAGE database for transcript discovery on male embryonic gonad development. *Bioinformatics* 2010;26:585–6.
- [22] Darde TA, Lecluze E, Lardenois A, Stévant I, Alary N, Tüttelmann F, et al. The ReproGenomics Viewer: a multi-omics and cross-species resource compatible with single-cell studies for the reproductive science community. *Bioinformatics* 2019;35:3133–9.
- [23] Bradford Y, Conlin T, Dunn N, Fashena D, Frazer K, Howe DG, et al. ZFIN: enhancements and updates to the Zebrafish Model Organism Database. *Nucleic Acids Res* 2011;39:D822–9.
- [24] Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, et al. The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. *Nucleic Acids Res* 2017;45:D362–8.
- [25] Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 2013;29:15–21.
- [26] Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009;25:2078–9.
- [27] Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat Protoc* 2016;11:1650–67.
- [28] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;15:550.
- [29] Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 2015;12:59–60.
- [30] Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. *BMC Bioinformatics* 2009;10:421.
- [31] Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol* 2019;20:238.
- [32] Bailey TL, Johnson J, Grant CE, Noble WS. The MEME Suite. *Nucleic Acids Res* 2015;43:W39–49.
- [33] Gupta S, Stamatoyannopoulos JA, Bailey TL, Noble WS. Quantifying similarity between motifs. *Genome Biol* 2007;8:R24.
- [34] Kent WJ. BLAT—the BLAST-like alignment tool. *Genome Res* 2002;12:656–64.
- [35] Harris R. Improved pairwise alignment of genomic DNA. The Pennsylvania State University 2007.
- [36] Blanchette M, Kent WJ, Riemer C, Elnitski L, Smit AFA, Roskin KM, et al. Aligning multiple genomic sequences with the threaded blockset aligner. *Genome Res* 2004;14:708–15.
- [37] Ramani R, Krumholz K, Huang YF, Siepel A. PhastWeb: a web interface for evolutionary conservation scoring of multiple sequence alignments using phastCons and phyloP. *Bioinformatics* 2019;35:2320–2.
- [38] Ortega-Recalde O, Goikoetxea A, Hore TA, Todd EV, Gemmell NJ. The genetics and epigenetics of sex change in fish. *Annu Rev Anim Biosci* 2020;8:47–69.