

Supplementary Materials for

**Suppressing estrogen signaling by a Y-linked nuclear receptor
determines sex in *Bombina* toads**

Heiner Kuhl, Leon Obendorf, Christoph Winkler, Sophie Leinweber, Martin Knytl,
Dmitrij Dedukh, Evgeniya Grigoryeva, Sylke Winkler, Thomas Fullerton Brown, Martin
Pippel, Yunjing Goh, Julia Bindl, Susan Mbedi, Markus Harzdorf, Ralph Tiedemann, Sven
Wuertz, Maria Ogielska, Gene Myers, Mark L. Blaxter, Jacek M. Szymura,
Petra Knaus, Matthias Stöck

Corresponding authors: matthias.stoeck@igb-berlin.de , petra.knaus@fu-berlin.de

The PDF file includes:

Materials and Methods
Supplementary Text S1 to S2
Figs. S1 to S11
Tables S1 to S7

References

Other Supplementary Materials for this manuscript include the following:

Data S1 to S3

Materials and Methods

Animals

A genetic family of *Bombina bombina* from Wodzisław Małopolski, Poland, was reproduced in captivity, and we obtained the DNA of mother, father and 90 offspring for ddRADseq. The juveniles were raised until Gosner-stages 44-46, euthanized and sexed using anatomical and/or histological assignment of gonads as ovaries or testes. Offspring comprised 45 females, 32 males and 13 unknown or not unequivocally sexed juvenile toads (Table S1).

Sex-diagnostic molecular markers obtained from *B. bombina*, were tested in samples of three adult *B. orientalis* individuals (1♀, 2♂), as well as in anatomically sexed offspring (10♀♀, 7♂♂) from these three potential *B. orientalis* parents. We also sampled buccal cells using swabs from life adult toads of known sex (secondary sexual characters, proof of reproduction including individual photographic documentation). These comprised the following nominal taxa: *Bombina variegata*, *B. v. scabra*, *B. v. 'kolombatovici'*, *B. pachypus*, *B. orientalis*, *B. maxima*, and *B. microdeladigitata* (for numbers and details Table S1). Of the latter species, we also obtained two adult males from a breeder. For approximate phylogeographic allocation of European *Bombina* individuals of unknown geographic origin from breeders, we Sanger-sequenced mitochondrial DNA, specifically ca. 1.2 kb of the cytochrome *b* gene as described with slight modifications. Briefly, PCRs were carried out using primers L16245F: 5'_CCTTAATGGCACCCAACCT_3' and G106R: 5'_AACTGCCAATAGTGTGTGTCAAAG_3' under PCR-conditions described for *err1Y* at an annealing temperature of 56°C. Products were Sanger-sequenced in both directions, edited and compared to the large dataset from [53].

Double-digest restriction site-associated DNA sequencing (ddRADseq)

DNA of the genetic family of *Bombina bombina* from both parents and 90 offspring was extracted using the DNeasy Tissue Kit or the BioSprint robotic workstation (QIAGEN), following manufacturer's protocols. DdRADseq libraries were prepared according to [54]. In short, digestion of genomic DNA by two restriction enzymes (*SbfI*, *MseI*) and subsequent ligation of fragments to adapters with unique barcodes (4–8 bp) was followed by library-PCR (20 cycles) of fragments with different cut-sites, and a gel-based size selection, isolating the 400–500 bp fraction. To increase marker distribution in the huge genome, RADseq libraries were single-end sequenced (150 bp) on an Illumina NextSeq 500 at the Max Delbrück Center, resulting in an average read number of 5.2 million reads per individual.

Genome assemblies

The genome of *B. bombina* was sequenced following the criteria as described by Rhie et al., 2021 [55]. A combination of PacBio Sequel II continuous long reads (CLR), 10x Chromium linked reads libraries, and Bionano optical mapping has been applied to high molecular weight genomic DNA. These data were complemented by Hi-C chromatin capturing data. All data were integrated following the Vertebrate Genomes Project v1.6 pipeline [55] using FALCON (v1.8.1) and FALCON-Unzip (v1.3.7) [56] to generate a primary and an alternate assembly, followed by duplicate purging with purge_dups (v1.2.3) [57]. The assembly was scaffolded using scaff10X (v4.2a)[58], optical mapping data integrated via Bionano Solve DLS (v3.6.1), and Hi-C chromatin contact data incorporated with SALSA (v2.2) [59]. Base-level polishing was performed with gcpp (v2.0.2)[60] and FreeBayes (v1.3.2) [61], with read alignment handled by Longranger Align

(v2.2.2) [62]. The final assembly (aBomBom1.pri/ GCF_027579735.1) was manually curated using gEVAL and released by the Vertebrate Genomes Project in January 2023.

The genome of *B. variegata* was sequenced to 26x coverage using PacBio HiFi long reads, supplemented with Arima2 Hi-C chromatin contact data, both generated by the Wellcome Sanger Institute Tree of Life Programme. The initial assembly was performed with Hifiasm [63] in trio mode, utilizing k-mers derived from Illumina-sequenced parental samples to phase the assembly into paternal and maternal haplotypes. Each haplotype was scaffolded independently using YaHS [64] with the Hi-C data. The mitochondrial genome was assembled separately with MitoHiFi [65]. Each haplotype assembly was subsequently analysed and manually improved using TreeVal [66] (GCA_964205295.1, paternal / GCA_964205305.1, maternal). Chromosome-scale scaffolds supported by Hi-C data were named in descending order of size. For a statistical overview of all fire- and yellow-bellied toad assemblies: Table S7.

Mapping of ddRADseq markers and re-mapping of the long-read coverage to the *B. bombina* reference genome

To identify sex-diagnostic markers in our ddRADseq data, we initially used the software RADsex [67] without a reference genome but failed. Once, a reference genome of *B. bombina* became available, we re-applied RADsex with less stringent parameters than the default (--disable-correction, --min-depth 1, --signif-threshold 0.001) to accommodate for shortcomings of RADseq fragment representation, stemming from stochasticity and large genome size. Resulting RADseq markers were subsequently mapped to the genome assembly of *B. bombina* (GCF_027579735.1_aBomBom1.pri_genomic.fna.gz) using Minimap2 [68] with parameters for extremely large genomes and short reads (-I 100G -x sr). The marker positions were analyzed in a table calculation software (Excel) and exhibited higher density in males in one chromosomal region compared to the entire rest of the genome. We extracted the 11 markers from the RADsex output using the Linux grep command to check their presence or absence in all samples of the genetic family.

To further analyze this genomic region of enriched RADsex markers, we downloaded the male long-read sequencing data from VGPs GenomeArk repository (Data availability). These data were re-mapped to the *B. bombina* genome (GCF_027579735.1) by Minimap2 (-I 100G -x map-pb). Coverage was analyzed by converting the Minimap2 output to bed-format and applying bedtools' [69] 'genomecov' to create a per-base coverage file. This large file was used for a windowed analysis calculating the median coverage of 40 kbp windows over the genome by bedtools map.

Synteny inference of *Bombina bombina* LGs with those of *Xenopus tropicalis*

The NCBI reference proteins of *X. tropicalis* (GCF_000004195.4_UCB_Xtro_10.0_translated_cds.faa.gz), which included their chromosomal origin in the fasta header, were mapped against those of the *B. bombina* reference (GCF_027579735.1) using Miniprot [70]. Protein matches between chromosomes of both species were counted. Chromosome pairs with more than 200 protein hits were kept for analysis and corresponding chromosomes allocated by majority rule. This assigned *B. bombina* chr2 to *X. tropicalis* chr1 and *B. bombina* chr4 to *X. tropicalis* chr5. For composed chromosomes, the assignment of multiple blocs is shown.

Whole genome short-read sequencing of *B. maxima* and *B. microdeladigitora*

Genomic DNA of buccal swabs of a female and a male of both 28-chromosome species was extracted using the DNeasy Tissue Kit (QIAGEN), following the manufacturer's protocols. DNA of each sample was sequenced to about 15X genome coverage using DNBseq technology (150 bp paired end reads) according to protocols of the sequencing provider (BGI, Warsaw, Poland). Sequencing data was mapped to the *Bombina bombina* reference genome (GCF_027579735.1) using Minimap2 (-I 100G -x sr) and post-processing with samtools [71].

Reconstruction of *err1Y* in *B. variegata*, *B. orientalis*, *B. maxima* and *B. microdeladigitora*

To predict the *err1Y* ortholog in the *B. variegata* genome, we translated the *B. bombina err1Y* coding sequence (GenBank accession: XM_053700116.1) into its amino acid sequence and then mapped it with Miniprot v0.13-r248 [70] against the *B. variegata* genome (GCA_964205295.1, Jul-31-2024). Miniprot GTF format output was used to extract the corresponding coding and protein sequence of the *B. variegata err1Y* gene. Using the new draft genomes, mappings to *err1Y* and *err1* in *B. maxima* and *B. microdeladigitora* were inspected in IGV [72]. We extracted paired reads mapping to exons of these genes and assembled them by IDBA [73]. The assembled exons for each gene in each sample were concatenated and used to predict their coding sequence by Miniprot [70]. The resulting transcripts were multiple-aligned with the available transcripts from other *Bombina* species by Mafft [74] and manually curated using Bioedit [75]. A phylogenetic tree of *err1*, *err1*-like and *errY* coding sequences was constructed with IQtree2 [76].

Multiplex PCR assays for genetic sexing in multiple Eurasian *Bombina* taxa

To contrast the male-specific (hypothetical Y-chromosomal) steroid receptor (*err1Y*) of *B. bombina* linkage group (LG) 2 with the autosomal estrogen related receptor 1 (= *erra* or *esrra* or *err-alpha*) on linkage group 7 in a molecular assay, we initially used a single primer pair (Y_vs_Auto_F: 5'_GACTTCCAGTWRAGGWGG_3', Y_vs_Auto_R: 5'_ACCACTGCATTCATTGGAKCT_3'), co-amplifying from both loci in males (1 product on the Y, and 2 from both autosomes) but only the autosomal locus in females (PCR conditions below). To improve universal cross-species amplification in multiple Eurasian *Bombina* taxa, we developed a new multiplex-PCR assay, with a Y-specific primer pair (BombYonly_300bp_F: 5'_ACCACTGCATTCATTGGAGCT_3'; BombYonly_300bp_R: 5'_GACTTCCAGTTAAGGAGG_3') that amplified a 300 bp PCR-product exclusively in males, and an autosomal 397 bp PCR-product (Bomb_Auto_397bp_F: 5'_AGTTACCTAGGAGTTGGCCC_3', Bomb_Auto_397bp_R: 5'_CACCTCCATCCAAACGCTTT_3') in both sexes. Both PCR-amplifications were carried out in 25 µL reactions containing 1.25 µL of each primer, 14.5 µL of water, 5 µL of AllTaq buffer (5x), 0.5 µL of AllTaq polymerase (Qiagen), and 2 µL of DNA (concentration 4–20 ng/µL). Temperature regime comprised 2 min at 95 °C (initial denaturation), and 38 cycles of (30 s at 95 °C denaturation, 45 s at 57.7 °C annealing, 30 s at 72 °C extension), and 5 min at 72 °C final extension. PCR products (4 µL) were visualized by electrophoresis on 4% high-resolution Sieve-agarose (Biozym) gels, run for 60 min at 70 V. To verify the homology of the PCR-products across *Bombina* taxa and for phylogenetic analyses, we Sanger-sequenced them from bands cut from preparative agarose gels and tested them by BLAST and phylogenetic analyses (see below).

FISH-TSA on metaphase chromosome preparations

Using the gene synthesis service of Eurofins Genomics, we synthesized the entire mRNA-sequence of *err1Y* (= *err1-like*, LOC128647330) from the *B. bombina* genome (GCF_027579735.1) using novel primers (BombWholemRNA_F: 5' ATGGGGGAAAGAAGCTGCAATC_3', BombWholemRNA_R: 5' AAAATGCTGCATAACATGTCCAG_3' for amplification. The PPP Master Mix (Top-Bio, Prague, Czech Republic) was used for amplification of this *err1Y* DNA-equivalent according to the suppliers' recommendations (annealing temperature of 58 °C, elongation time 1 min 12 s, 30 cycles). The amplicon was purified [77] and labeled with digoxigenin-11-dUTP (Jena Bioscience, Germany) using the DecaLabel DNA Labeling Kit (Thermo Fisher Scientific). The labeled PCR amplicon was used as a probe for fluorescence *in situ* hybridization with tyramide signal amplification (FISH-TSA). Tissue suspensions (kidney, intestine) were obtained from phenotypically and/or genetically sexed females and males of each of the three species *B. bombina* (two males, two females), *B. variegata* (two males, two females), and *B. orientalis* (one male, two females). Animals were treated with 0.04% colchicine for 16 h (10 µl per gram of body weight) and tissues fixed in methanol-acetic acid (3:1), macerated on a microscope slide [78], followed by FISH. The FISH-TSA followed [79] with minor modifications [80]. The signal was detected by antidigoxigenin-POD, Fab fragments (Roche), and amplified using the TSATM Plus Tetramethylrhodamine System Kit (NEL742001KT, PerkinElmer, Inc., Waltham, MA, USA). Mitotic metaphase spreads were counterstained with ProLongTM Diamond Antifade Mountant with the fluorescent 4',6-diamidino-2-phenylindole (DAPI) stain (ThermoFisher Scientific). Microscopy involved Leica Microsystem (Wetzlar, Germany) and processing of metaphase images in red-blue channels.

Phylogenetic trees of *err1* and *err1Y*

Sanger-sequenced gene fragments of either *err1* or *err1Y* of eight nominal taxa were aligned using the PhyML package [81], and phylogenetic trees were reconstructed using maximum likelihood under the GTR-model of sequence evolution and bootstrapped based on 1000 resampled data sets.

RNA isolation, reverse transcription and cDNA-based PCRs

RNAs of multiple organs (muscle, eye, heart, spleen, liver, testis) of one juvenile, male *Bombina orientalis* (Gosner stage 46, ca. 6 months after metamorphosis; sample name: Bor_male), snout-vent-length ca. 1.5 cm and a female and two males of *B. variegata* (Gosner stage 46 plus 4 months after metamorphosis) were extracted using the TRIzol Reagent (Thermo Fisher Scientific, Waltham, USA) according to the supplier's recommendation and cleaned using the RNeasy Mini Kit (Qiagen). All RNAs included in this study exhibited RIN values >8.0. For the testes RNA of this *B. orientalis* male, library processing and RNAseq were carried out by BGI (Warsaw, Poland) on a NovaSeq 6000 PE 150, generating 9 Gb of raw data per sample.

Using the iScript RT Supermix for RT-qPCR kit (Biorad), 10 µL of RNA (concentration of 80 ng/µL) from the same *B. orientalis* male (Bor_male) as used for RNAseq, was randomly reverse-transcribed mRNA into cDNA according to the manufacturer's recommendations.

To test organ-specific autosomal *err1* vs. *err1Y* gene expression, we designed primers using mRNA-sequences from the *B. bombina* reference. One primer for each PCR-product crosses the boundary of adjacent exons, preventing amplification from gDNA contamination in cDNA. For *err1* we used Bomb_Auto_RNA_F: 5' GTGACAGACGCCAGATCT_3' and Bomb_Auto_RNA_R: 5' GATATTACCTTGGATAGTGCGCTTGA_3 (Ta = 58.4°C) yielding

a product of ca. 300 bp at the 5' end and Bomb_cDNA_Auto4F: 5'_CATGGCTGCTCTGCTAAATA_3' with Bomb_cDNA_Auto4R: 5'_TTTCCGTAAACCTACCTTTATC_3' (Ta = 57.1°C) amplifying 100 bp at the 3' end of the autosomal gene. For *err1Y*, primers Bomb_Y_RNA_F: 5'_GGCATGAAGAAAGAATGTGTCA_3' and Bomb_Y_RNA_R: 5'_GACTCTGTGGATTATCAAGCACA_3' (Ta = 58°C) produced ca. 100 bp at the 5'-end and Bomb_ERR1Y_cDNA_F: 5'_CATCCGAGGTTATATTGAGTACA_3' with Bomb_ERR1Y_cDNA_R: 5'_GCTCGGCAAACATTTCTG_3' (Ta = 57.7°C) generated 79 bp at the 3' end of the gene. PCR conditions were the same as for multiplex sexing (see above) but with the annealing temperatures (Ta) given above for each primer pair.

In a genetic male of *B. variegata* (Bvarjuv2M) and a female (Bvarjuv5F), both in Gosner stage 46 (plus 3-4 months after metamorphosis), one testis and one ovary, respectively, were prepared for RNAseq (carried out by BGI Hongkong, China) using the single cell technique DNBSEQ SMART seq2 approach [82] for the limited testis tissue available (Bvarjuv2M) with ca. 30 million 100 bp paired end clean reads. For the larger ovary (Bvarjuv5F), we used the standard DNBSEQ with ca. 20 million 150 bp paired end clean reads. From each of the latter two *B. variegata*, the second gonad was processed for HCR (see below), allowing integrative evidence from the same individual male and female.

Gonadal transcriptome assemblies to screen for the *err1Y* gene

Transcriptomes of *B. orientalis* (Bor_male: "BO_tranom", Gosner stage 46, ca. 6 months post-metamorphosis) and *B. variegata* testis (Bvarjuv2M: "RNAB.var2M", Gosner stage 46, 3-4 months post-metamorphosis) and ovary (Bvarjuv5F, Gosner stage 46, 3-4 months post-metamorphosis) were assembled by IDBA_tran [83]. The resulting male assemblies were screened for transcripts matching the *B. bombina err1Y* coding sequence (accession: XM_053700116.1) using minimap2 [68]. All available transcript data of *err1Y* and *err1* were multiple aligned by mafft [74] and the most conserved regions between *err1Y* transcripts with the most differences to *err1* autosomal transcripts were chosen manually for designing HCR probes.

***In situ* expression analysis by Hybrid Chain Reaction**

Spatial expression in *B. variegata* gonads was assessed by multiplex Hybrid Chain Reaction (HCR). We prepared 3 females (Bvarjuv 7F, 19F and 20F) and 4 males (Bvarjuv14M, 16M, 21M and 22M) of genetically sexed *B. variegata* (Gosner[84] stage 46, at metamorphosis), at the beginning of anatomical differentiation of the primordial gonads, with early ovaries stretching almost the entire length of the mesonephroi and lobed external appearance, while early testes had shortened or exhibited already a bean-shaped anatomy. HCR was performed as described [85], with minor modifications. Briefly, for each gene, 10-15 pairs of split-initiator oligos (Table S4) were designed using the *B. bombina* mRNAs using HCRProbeMakerCL (<https://github.com/rwnull/HCRProbeMakerCL>). For *err1Y*, we used the mRNA of *B. variegata* to design specific oligonucleotides (Table S4). Oligos were ordered from Integrated DNA Technologies (IDT Singapore). Corresponding hairpin amplifier oligos and HCR buffer kits were purchased from Molecular Instruments. Gonad tissues were fixed in 4% paraformaldehyde (PFA) in phosphate buffered saline (PBS) containing 0.1% Tween 20 (PBST) for 16 hours at 4°C and then washed three times for five minutes in PBST. Samples were dehydrated in 25%, 50%, 75% and 100% MeOH/PBST for 10 minutes each and then stored at -20°C. After rehydration, gonad tissues were incubated in acetone for 12 min, washed in PBST and then incubated in 10 µg/mL

proteinase K in PBST for 16 minutes at room temperature. After several PBST washes, samples were refixed in 4% PFA/PBST for 20 minutes, washed and then prehybridized in hybridization buffer (Molecular Instruments) for 30 minutes at 37°C. Samples were then hybridized with 50 pM split-initiator oligo each overnight at 37°C. After washing with probe wash buffer (Molecular Instruments) and 5x saline-sodium citrate buffer-Tween, denatured amplifier hairpins were added and samples incubated overnight at 25°C in the dark. After several washes in 5x SSCTween and PBST, samples were mounted in glass-bottom dishes containing 1% low-melting point agarose. Confocal imaging was carried out with an Olympus Fluoview FV3000 microscope using 546 and 647 nm laser lines. Imaging data was processed using ImageJ (v1.54p).

Cell culture for luciferase reporter gene assays

HEK293T were procured from German Collection of Microorganisms and Cell Cultures (DSMZ) and maintained in low-glucose, phenol-red free Dulbecco's Modified Eagle's Medium (DMEM) (Gibco) containing 10% FCS, 2 mM L-glutamine, and penicillin (100 units/mL)/streptomycin (100 µg/mL). Cells were cultured at 37 °C and 5% CO₂ (v/v) in a humidified incubator.

Identification of putative ERE motifs in the *B. bombina* *foxl2* promoter region

The putative promoter sequence 1 kbp upstream of the *foxl2* transcription start site was extracted from the genome (NC_069502.1:240,030,455-240,031,455) and analysed for putative estrogen response elements (ERE) binding sites using TFBIND (<https://tfbind.hgc.jp/>). Seven putative EREs were identified, the highest scoring one (Transfac score = 0.788023) was used for experiments (NC_069502.1:240,031,423-240,031,441; located in the closest proximity to the *foxl2* gene-model (XM_053708975.1)).

Firefly luciferase reporter construct cloning

Firefly luciferase reporter plasmids were generated by cloning estrogen response elements (ERE) into a minimal promoter reporter backbone (In-house variant as described [86]). For this, forward and reverse primers encoding the binding motifs were first incubated in 1x T4 ligation buffer (NEB) at 95 °C for 15 min and cooled to room temperature to allow annealing. The annealed oligos were phosphorylated using T4 polynucleotide kinase (NEB). All primers used for PCR are listed in (Table S5). PCRs for cloning were carried out on a Peltier Thermal Cycler PTC-200 with an elongation time of roughly 1 kb/min.

The minimal luciferase promoter vector was digested with KpnI and XhoI restriction endonucleases (NEB), and the linearized backbone purified by gel extraction from a 1% agarose gel. To avoid vector religation, the digested vector was treated with Antarctic Phosphatase (NEB). Annealed oligonucleotides were ligated into the vector using T4 ligase (NEB) at room temperature overnight. Following ligation, reactions were heat-inactivated at 65 °C for 10 min and transformed into in-house prepared chemically competent DH5α *Escherichia coli*. Plasmid DNA was purified using the NucleoSpin Plasmid EasyPure kit (Macherey–Nagel) according to the manufacturer's instructions and verified by Sanger sequencing.

Receptor overexpression construct cloning

Plasmids encoding receptor overexpression constructs were generated by extracting the receptor protein sequences of *Bombina variegata* via miniprot mapping of *B. bombina* full-length *err1Y* to the *B. variegata* genome or using the truncated *err1Y* isoform sequence obtained from *de novo* assembly of RNAseq data of male *B. variegata* (Bvarjuv2M; Fig. S10). The constructs were

codon-optimized and subsequently synthesized by “Genewiz.com” including short N-terminal peptide tags (HA-Err1Y, flag-Er α , and myc-Err1, sequences provided in Table S6). The designed inserts were assembled into pcDNA3.1(-) backbone using NheI and NotI. The vectors were transformed into DH5 α *E. coli*, followed by plasmid purification as described above. The truncated HA-Err1Y isoform was generated by removing the DNA-binding domain by restriction digestion of the Err1Y plasmid with BamHI (sequence provided in Table S6). The appropriately sized fragments (vector backbone and receptor fragment containing the LBD domain) were purified from a 1% agarose gel. Ligation of the fragments and transformation were performed as described above. All constructs are confirmed via Sanger sequencing, followed by whole plasmid sequencing. The expression was confirmed using Western blot.

Dual luciferase reporter assay

HEK293T cells were seeded in 96-well plates at a density of 50,000 cells per well in DMEM full medium. The following day, cells were transfected with flag-Er α , myc-Err1, HA-Err1Y or truncated HA-Err1Y receptor constructs together with luciferase reporter plasmids in the indicated combinations. A constitutively expressing Renilla luciferase plasmid (RL-TK; Promega) was included in all transfections as an internal control. Transfections were conducted in Opti-MEM using polyethylenimine (PEI) (Sigma Aldrich). After 22-24 h, the transfection medium was removed and cells were incubated in serum-free medium for 24 h. Following starvation, cells were lysed using a passive lysis buffer (Promega). Luciferase activity was quantified with a TECAN Spark Luminometer.

Statistical analysis

Statistical tests for luciferase assay were performed using custom R scripts in RStudio [87]. First, technical replicate outliers were removed using a threshold of three times the median absolute deviation. Next, the distribution of the data across biological replicates was evaluated using the Shapiro–Wilk normality test on raw, non-normalised values. In most cases, the data did not deviate from normality ($p > 0.05$), with the exception of the er1 α condition at 50 ng ($p = 0.0409$), which is likely attributable to normal biological variability. For data visualization, values were normalised to the maximum signal observed in “Er α only”/ Err1 only” condition within each biological replicate. For statistical analysis, data were instead normalised within each biological replicate to a “1 : 1, Er α : Err1Y/truncated ErrY” or “1 : 1, Err1 : Err1Y/truncated Err1Y” conditions. Subsequently, an ANOVA was performed on all conditions with the exclusion of the 1 : 1 conditions, the “Err1Y only”/“truncated Err1Y only” conditions, and control conditions lacking receptor expression. Where ANOVA revealed significant differences, a Tukey post hoc test was performed to identify the specific pairs responsible for these differences, and statistical significance was assigned using an alpha threshold of $p < 0.05$.

Western Blot Analysis of Receptor Expression

To verify successful protein expression levels of receptor constructs, Western blotting was performed. Constructs included epitope tags as follows: flag-Er α , myc-Err1, and HA-Err1Y / truncated HA-Err1Y. Samples were collected from the same 96-well plates used for the dual luciferase assays following cell lysis. Lysates were denatured at 95°C in Laemmli buffer for 10 minutes and separated by SDS-PAGE using a 4% stacking and 10% resolving gel. Proteins were transferred onto a PVDF membrane (Merck) via Western blotting. The membranes were blocked for 1 hour at room temperature in 0.1% TBS-T containing 3% (w/v) BSA and then incubated

overnight at 4 °C with the primary antibodies: rabbit anti-FLAG (Cell Signaling Technology), mouse anti-myc (in-house produced), mouse anti-HA (Sigma Aldrich) at 1:1000 ratio in 0.1% TBS-T / 3% (w/v) BSA. The following day, membranes were washed in TBS-T and incubated for 1 h at room temperature with horse-raddish peroxidase (HRP) conjugated antibodies: goat- α -rabbit HRP and goat- α -mouse-HRP (1:10,000 dilution in 0.1% TBS-T w 3% BSA (Carl Roth)). Chemiluminescence was detected using the WesternBright Quantum HRP substrate (Advansta) with a FUSION FX7 digital imaging system.

Computational structure prediction and molecular docking

The structure of the receptor monomers/dimers and complexes were first predicted using Boltz-2 [88] and sub-atomic deviations in the fold were verified using known crystal structure data if available. For each condition, diffusion samples were set to 25 via the “--diffusion samples 25” flag during inference with multiple sequence alignments generated through the --use msa server flag. All models were generated with Boltz version 2.2.0. For molecular docking approach, coordinate constrained relaxation protocol was performed to optimize H-bonds based on the Rosetta Energy Function 2015 (ref2015_cst) within RosettaScripts [89]. To obtain a physically more validated structure while simultaneously scoring the interface, local Monte-Carlo-based protein-protein docking was performed in 2000 replicas. Here, first, Rosetta’s “Docking” mover was used with low resolution (backbone plus centroid) flags before a high resolution (full atom) docking (using ref2015 scoring function). Other parameters as distance perturbation and angle perturbation were left on default. The Interface RMSD (iRMSD) was plotted against the score of the interface residues (i_SC) of the top 500 structures (determined by i_SC) and mean binding energies were calculated. Lower iRMSD correlates to docking structures closer to the starting Boltz-2-multimer conformation while a lower interface score correlates with favourable protein-protein binding energies. For structural representations using PyMol [90], the top 10 hits showing the lowest i_SC were manually investigated. Scripts and example prediction .yaml files: data availability [52].

Ethics, animal welfare statement and permissions

Our research complies with all relevant ethical regulations. Beyond holding, no animal experiments were performed; minimal invasive DNA-sampling methods were applied whenever possible; the relevant animal protection committees of Jagiellonian University - Kraków (Poland) or the Leibniz IGB (Germany) in collaboration with the city of Berlin (LaGeSo), as well as the Czech Academy of Sciences, Institute of Animal Physiology and Genetics, who approved the study protocol (permissions 0926/PO/2008/AOPK and S/0673/PO/2008/AOPK) where relevant. Juvenile offspring from one *B. bombina* and *B. variegata* genetic family (JMS) were collected under permits issued by The Regional, and General Director of Environmental Protection, Poland (OP-I.6401.193.2013.MMr and DZP-WG.6401.101.2022Ł); and studies approved by Local Ethical Committee on Animal Testing, Kraków, Poland (94/V/2013 Nr 86/2013 and 05/05/2022 Nr 137/2022). Juvenile *B. orientalis*, juvenile *B. variegata* and two adult males of *B. microdeladigitora* were sourced from commercial breeders. Juvenile toads were humanely euthanized using an overdose of Tricaine PHARMAQ 1000 MG/G (MS222; concentration: 500 mg/L), a method recognized for causing no pain, suffering, or distress. Following euthanasia, animals were promptly decapitated to ensure death. We received a permit for minimal invasive swab sampling of wild-caught *B. bombina* (C45_DD-8493/343/1; 19/01/2024) from the State Directorate of Saxony. The lower nature conservation authority of the district of Görlitz (Saxonia,

Germany, Löbau) registered and permitted the in part protected other *Bombina* species for captive breeding by one of us (MH). All tadpoles and toads were kept according to approved practice, involving appropriate tank size, sufficient waterflow, natural photo-period, *ad libitum* food, and temperatures within the species' thermal tolerance range. All of this guaranteed that no pain, suffering, distress or harm was inflicted on the animals.

Supplementary Text

Text S1:

Estrogen related receptors, estrogen receptors and their interactions

The SD-candidate *err1Y* encodes estrogen-related receptors (in mammals: ERRs), a subfamily of the nuclear receptor superfamily that includes ligand-activated transcription factors like estrogen receptors (ERs). ERRs and ERs likely originated via tandem duplication from a common ancestor. Most jawed vertebrates have multiple ERRs from ancient duplications [44,45]. Unlike ERs, ERRs are orphan receptors known to regulate energy metabolism and mitochondrial biogenesis but do not bind estrogen [46,91-93]. However, ER α , ERR α , and ERR β can bind and activate transcription through the estrogen response element (ERE), ERR elements (ERREs) [91,94], and the SF-1 response element (SFRE). ERR1 activates genes via these elements without ligands [95]. ERR1 is expressed in a broader range of tissues than ERR β , including mouse testis [96]. ERRs are also present in *Drosophila* gonadal tissues, where knockdown causes underdeveloped testes and reduced fertility [97]. ERR1 can activate ER α target genes via imperfect ERE half-sites, especially in ER-negative cells [46,91-93]. Although ERR1 has a ligand-binding pocket, no natural ligand is known, and it constitutively binds coactivators at the back of its ligand binding domain (LBD) without the need to be switched on by another molecule [98-100]. ERR1 may also modulate estrogen signaling by competing with ER for EREs and shared co-regulators [47,48].

Text S2: An overview of the recent history of genome annotation including *err1Y* of *B. variegata*

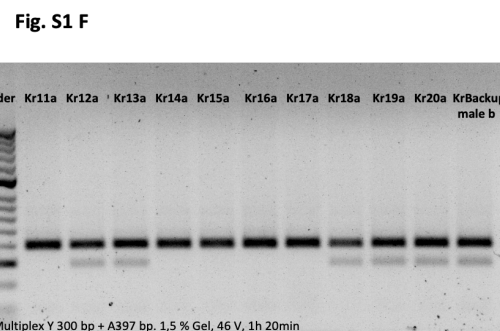
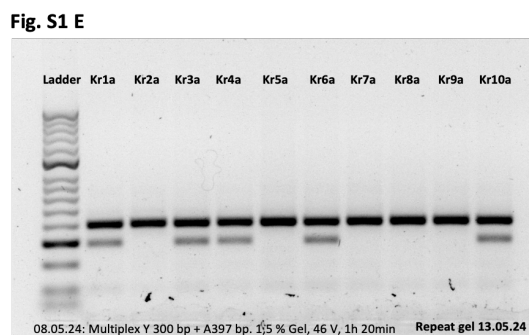
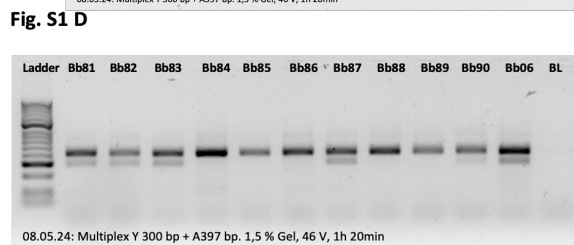
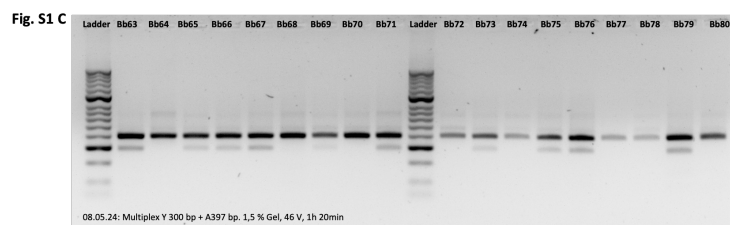
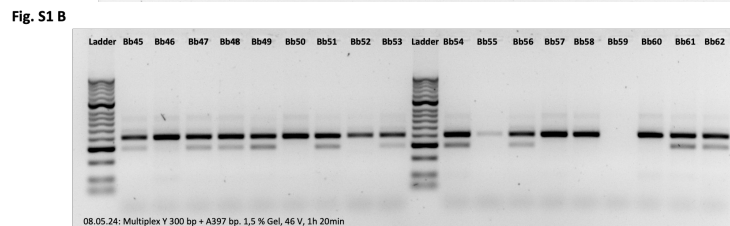
The candidate male sex-determining gene *err1Y* is located on *B. bombina* LG2 (*err1-like*: 845,166,599–845,362,339). In the 'paternal' *B. variegata* haplotype assembly (GCA_964205305.1, Aug-01-2024), no orthologous gene was found; the most similar match was the autosomal *err1* on LG7. Translating *B. bombina err1Y* (XM_053700116.1) and mapping it to the *B. variegata* genome revealed the gene in the wrongly assigned 'maternal' haplotype on an unplaced contig (CAXNTK010000102.1:851,852–966,674), enabling extraction of the corresponding *B. variegata err1Y* coding and protein sequences (Data S3). To address possible inconsistencies in haplotype assignment labels, correspondence with the original authors was initiated. The paternal and maternal designations have since been verified and corrected in the NCBI database. As of the October 17, 2024 update, the '*err1-like*' (= *err1Y*) sequence (CAXNTK010000102.1:851,852–966,674) has been removed and is not currently accessible. *Err1Y* was also detected in a draft genome of a male *B. variegata* (GCA_905336975.1).

Fig. S1 A

Bb3 Bb13 Bb14 Bb15 Bb16 Bb17 Bb18 Bb19 Bb20 Ladder Bb21 Bb22 Bb23 Bb24 Bb25 Bb26 Bb27 Bb28 Bb29 Bb30

Bb31 Bb32 Bb33 Bb34 Bb35 Bb36 Bb37 Ladder Bb38 Bb39 Bb40 Bb41 Bb42 Bb43 Bb44

08.05.24: Multiplex Y 300 bp + A397 bp, 1.5 % Gel, 46 V, 1h 20 min



E-F: Molecular sexing of wild-caught minimal-invasively sampled *B. bombina* from Germany (Saxonia, Kreba).

Fig. S2 A

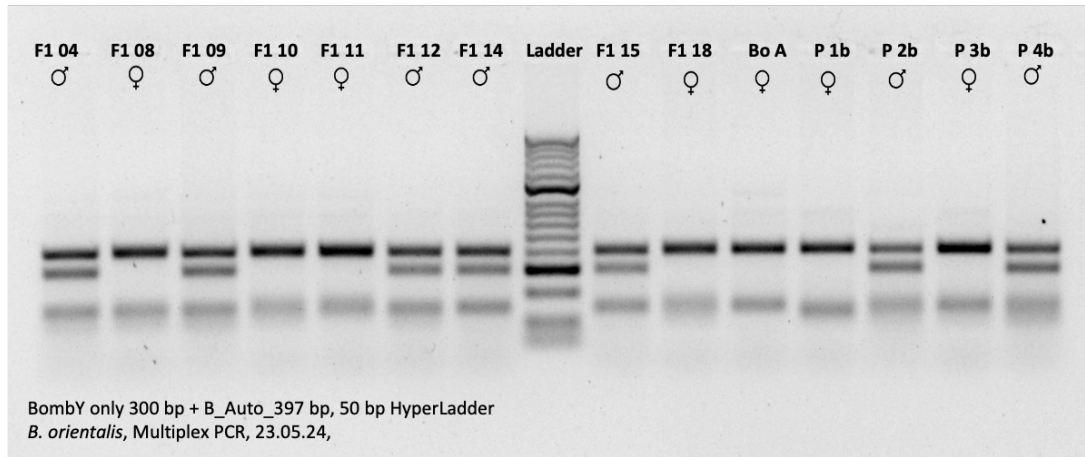


Fig. S2 B

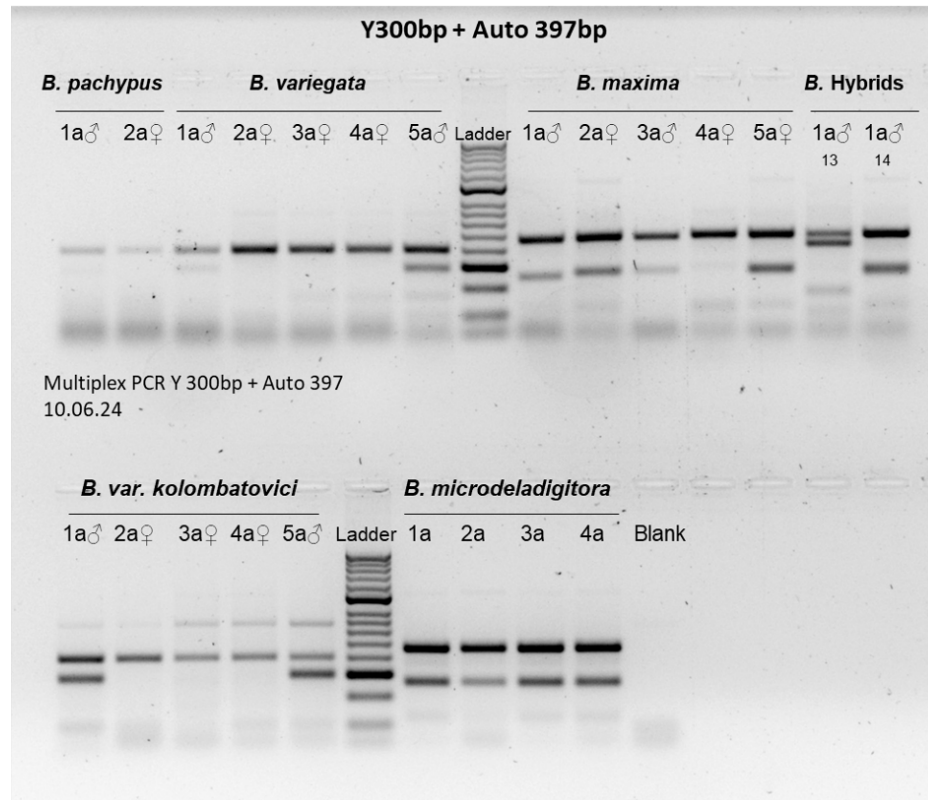


Fig. S2.

Molecular sexing of various *Bombina* species and interspecies hybrids with known phenotypic sex.

A: Molecular sexing of the *B. orientalis*.

B: Molecular sexing of wild-caught minimal-invasively sampled of nominal taxa *B. v. pachypus*, *B. variegata variegata*, *B. variegata kolombatovici*, *B. maxima*, *B. microdeladigitara* and several interspecies hybrids with known phenotypic sex.

Fig. S3A

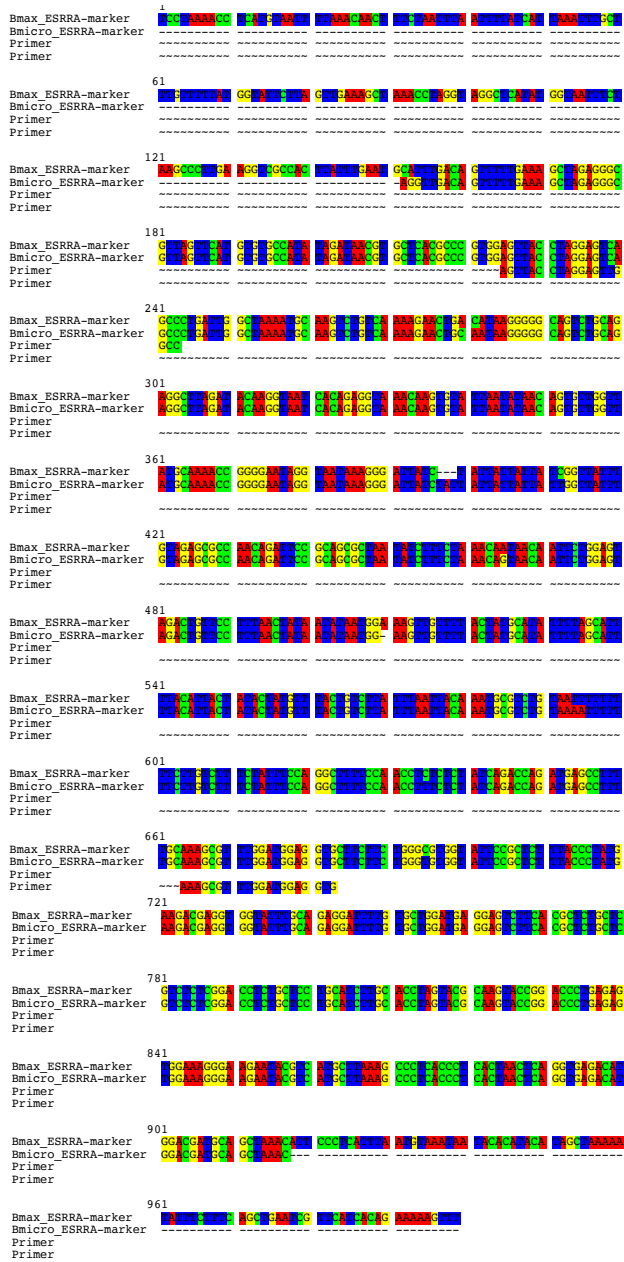


Fig. S3B

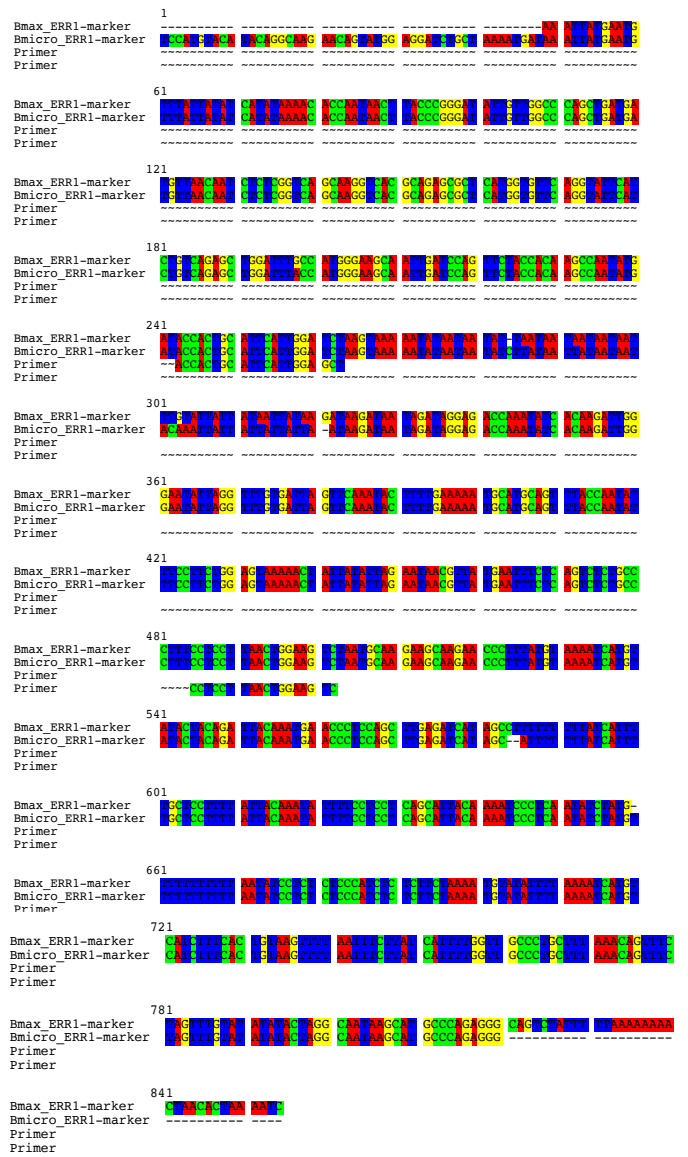


Fig. S3.

A: The fragment of the autosomal *err1* (*essra*) gene, reconstructed from short-read whole genome sequencing, in *Bombina microdeladigitata* and *B. maxima*, which do not exhibit sex-specificity of *err1-like*; with the fully matching primer sites of marker BombAuto_397bp_F/R.

B: The fragment of the *err1-like* gene reconstructed from short-read whole genome sequencing in *Bombina microdeladigitata* and *B. maxima* in which it does not exhibit sex-specificity; with the fully matching primer sites of Bomb_Yonly_300bp_F/R.



Fig. S4.

The *err1*-like gene exhibits a consistent open reading frame disruption caused by a 2 bp insertion in both *B. maxima* and *B. microdeladigitara*. In these 28 chromosome species, the gene is present in both sexes, suggesting that it has either lost or never possessed a sex-determining function.

Fig. S5 A



4500 - male Austria
 4501 - male Waksmund
 4502 - female Waksmund - Trio female parent
 4503 - male Biała Woda - Trio male parent
 4504 - male Waksmund
 4505 - male Waksmund

WK - Waksmund
 WKBW - Waksmund + Biała Woda
 AU - Austria (coll. Lukas Cholewa);
 Animals were anesthetized using MS222

Fig. S5 B

Bombina variegata samples 14Sept1024

4501- Male, Waksmund
 4502- Female, Waksmund - Trio genome female parent
 4503- Male, Biała Woda - Trio genome male parent
 4504- Male, Waksmund
 4505- Male, Waksmund
 4506- Female, F1 progeny Bv x Bv 4502 and 4503
 4507- Female, F1 progeny Bv x Bv 4502 and 4503
 4508- Male, F1 progeny Bv x Bv 4502 and 4503
 4510- Male, F1 progeny Bv x Bv 4502 and 4503
 4511- Male, Kamienica
 4512- Male01, F1 progeny Bv x Bv 4502 and 4503
 4513- Female02, F1 progeny Bv x Bv 4502 and 4503
 4514- Female03, F1 progeny Bv x Bv 4502 and 4503
 4515- Female04, F1 progeny Bv x Bv 4502 and 4503
 ...
 4540- Male05, F1 progeny Bv x Bv 4502 and 4503
 4541- Female06, F1 progeny Bv x Bv 4502 and 4503
 4542- Male07, F1 progeny Bv x Bv 4502 and 4503
 4543- Female08, F1 progeny Bv x Bv 4502 and 4503
 4544- Female09, F1 progeny Bv x Bv 4502 and 4503
 4545- Female10, F1 progeny Bv x Bv 4502 and 4503
 4546- Male11, F1 progeny Bv x Bv 4502 and 4503
 4547- Female12, F1 progeny Bv x Bv 4502 and 4503
 4548- Female13, F1 progeny Bv x Bv 4502 and 4503
 4549- Female14, F1 progeny Bv x Bv 4502 and 4503

N=24

*Phenotypic sex corrected by Jacek Szymura,
 Sept. 26, 2024

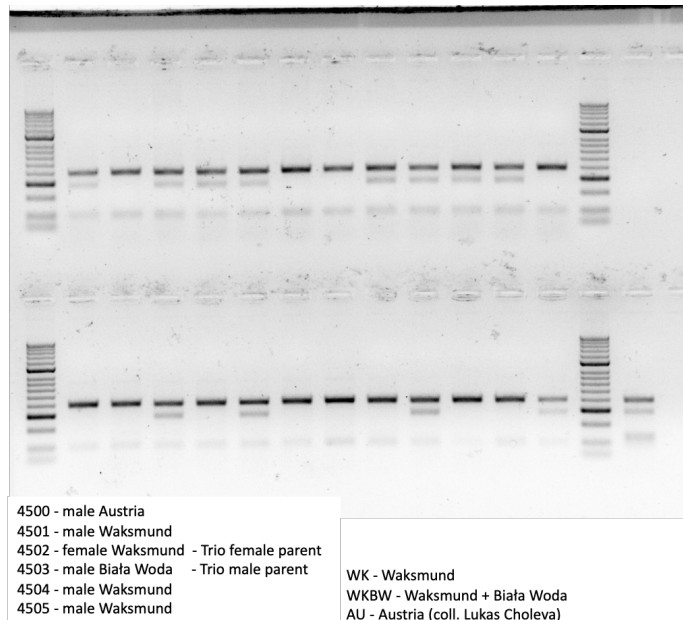
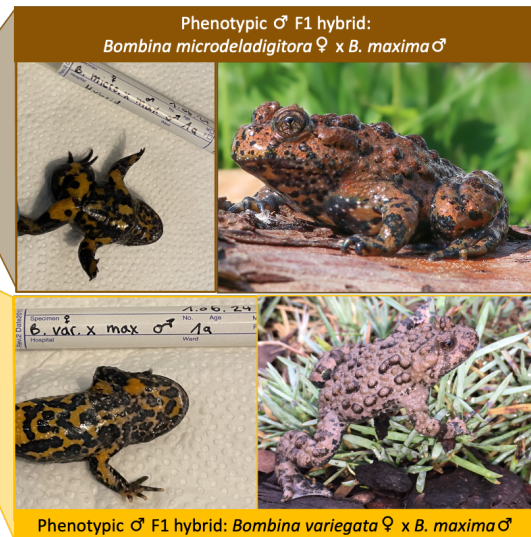
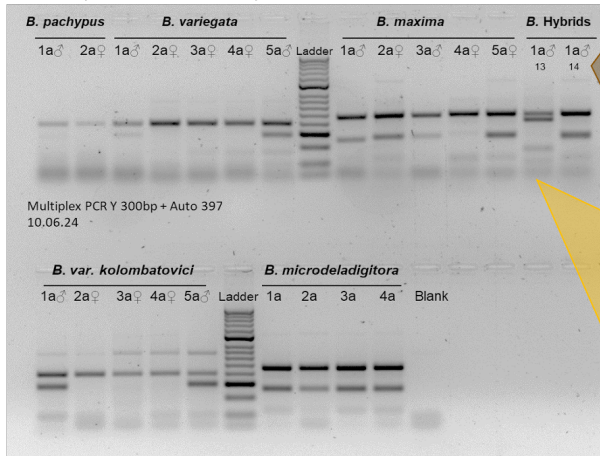


Fig. S5 C**Fig. S5 D**

Bomb_Yonly_300bp_F/R is male-diagnostic in *B. pachypus*, *B. variegata* and *B. v. kolombatovichi*, but not in *B. microdeladigitata* and *B. maxima*; Sanger sequencing and primer alignment with whole genome data of both species confirmed PCR product identities



© Markus Harzdorf

Fig. S5.

A: Inheritance of the male-specific marker in a genetic family of *B. variegata*:

ventral side of phenotypes in parents and progeny (carefully anesthetised); molecular markers shown in Fig. S5B.

B: Molecular sexing of *Bombina variegata*, parents and their F₁-progeny. PCR-products using primers

BombAuto_397bp_F/R and Bomb_Yonly_300bp_F/R

C: Single F₁ specimens of interspecies crosses involving the Chinese species *Bombina microdeladigitata* and *B. maxima* that do not exhibit sex-specificity of *err1-like*; PCR-products using primers BombAuto_397bp_F/R and Bomb_Yonly_300bp_F/R.

D: F₁ of cross *B. bombina* ♀¹ mother x *B. variegata* father ♂², PCR-products using primers BombAuto_397bp_F/R and Bomb_Yonly_300bp_F/R; F₁ bred in captivity, parents (♀¹ Wodzisław Małopolski, (ERS3926743), Poland; ♂² Obidowa, near Nowy Targ, (ERS3926742), Poland).

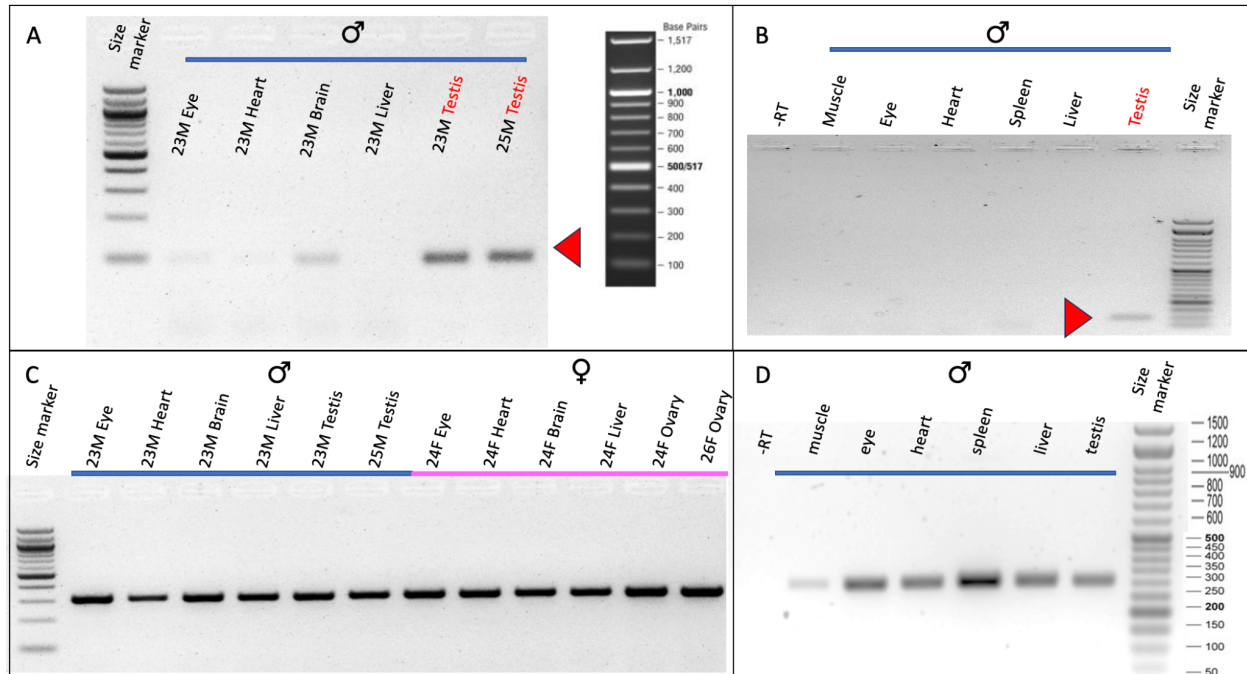


Fig. S6.

Gene expression of Y-specific steroid orphan receptor *err1Y* (A, B: red arrowheads) vs. autosomal *err1 (essra)* (C, D).

A: *Err1Y* transcripts amplified mostly in testes (and male ♂ brain) but were weak (or absent) in eye, heart and liver in two *B. variegata* males (Gosner 46, 4 months post-metamorphosis).

B: *Err1Y* transcripts amplified only in testis of a juvenile male *B. orientalis* (Gosner 46, 6 months post-metamorphosis), but were absent in muscle, eye, heart, spleen and liver.

C: Autosomal *err1 (essra)* was expressed in all examined somatic tissues (eye, heart, brain, liver), and gonads of male (♂) testes and female (♀) ovaries in *B. variegata*.

D: Autosomal *err1 (essra)* was expressed in all examined somatic tissues of a male *B. orientalis*. The same cDNAs were used in A and C as well as B and D, respectively.

-RT: control from RNA without reverse transcription derived from muscle. PCR product identities were confirmed via Sanger sequencing. Size marker shown in A also applies to C, size marker shown in D also applies to B; for at lower right. Primers Bomb_Auto_RNA_F/R for *err1* and Bomb_Y_RNA_F/R for *err1Y* were used.

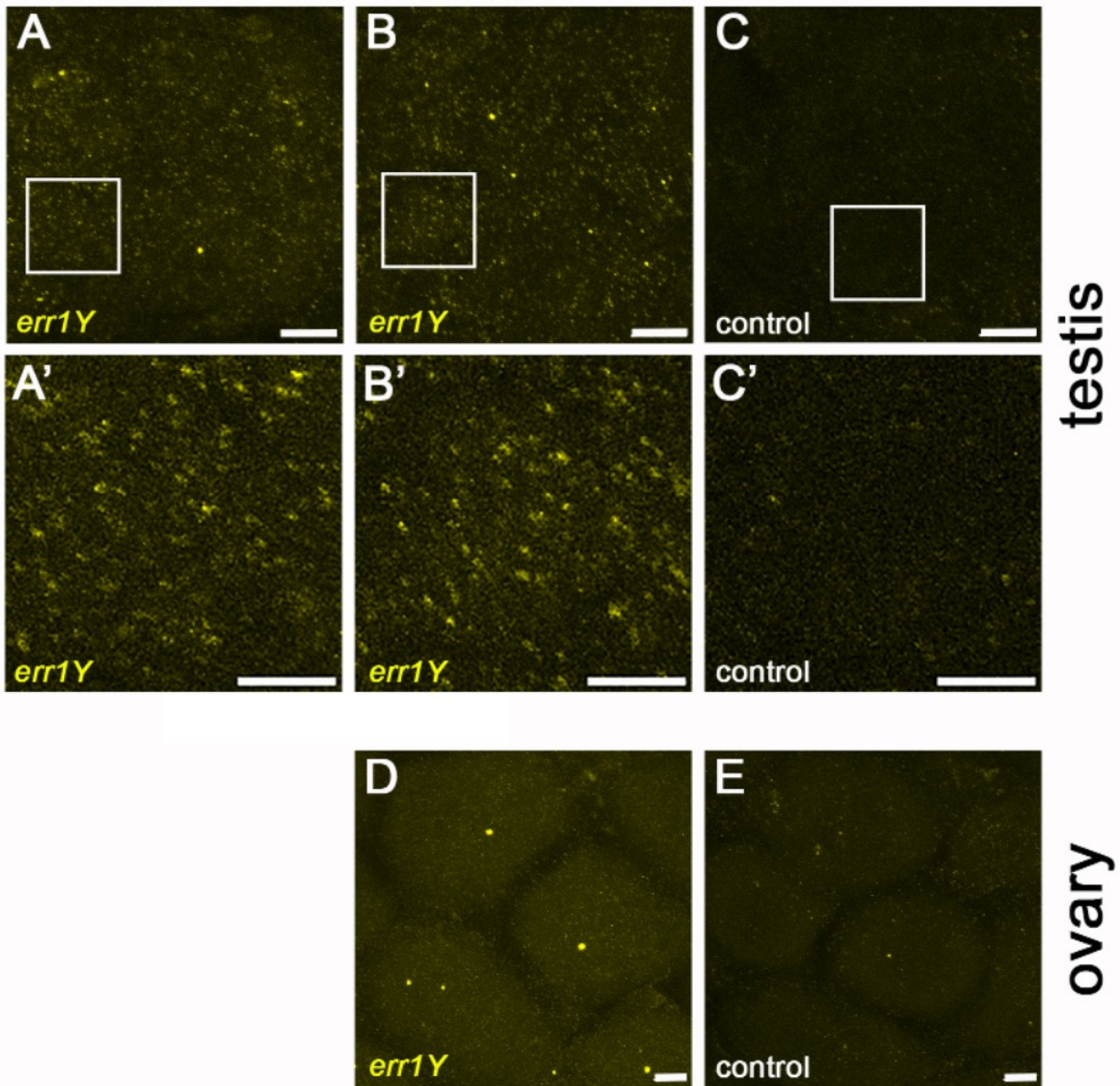


Fig. S7.

Expression of *err1Y* in developing male and female gonads. Single channel (647 nm) images of merged images shown in Fig. 4. C, C' and E are negative controls with no HCR initiator oligos added. Scale bars are as described in Fig. 4. All images were captured with the same intensity settings to allow quantitative comparison.

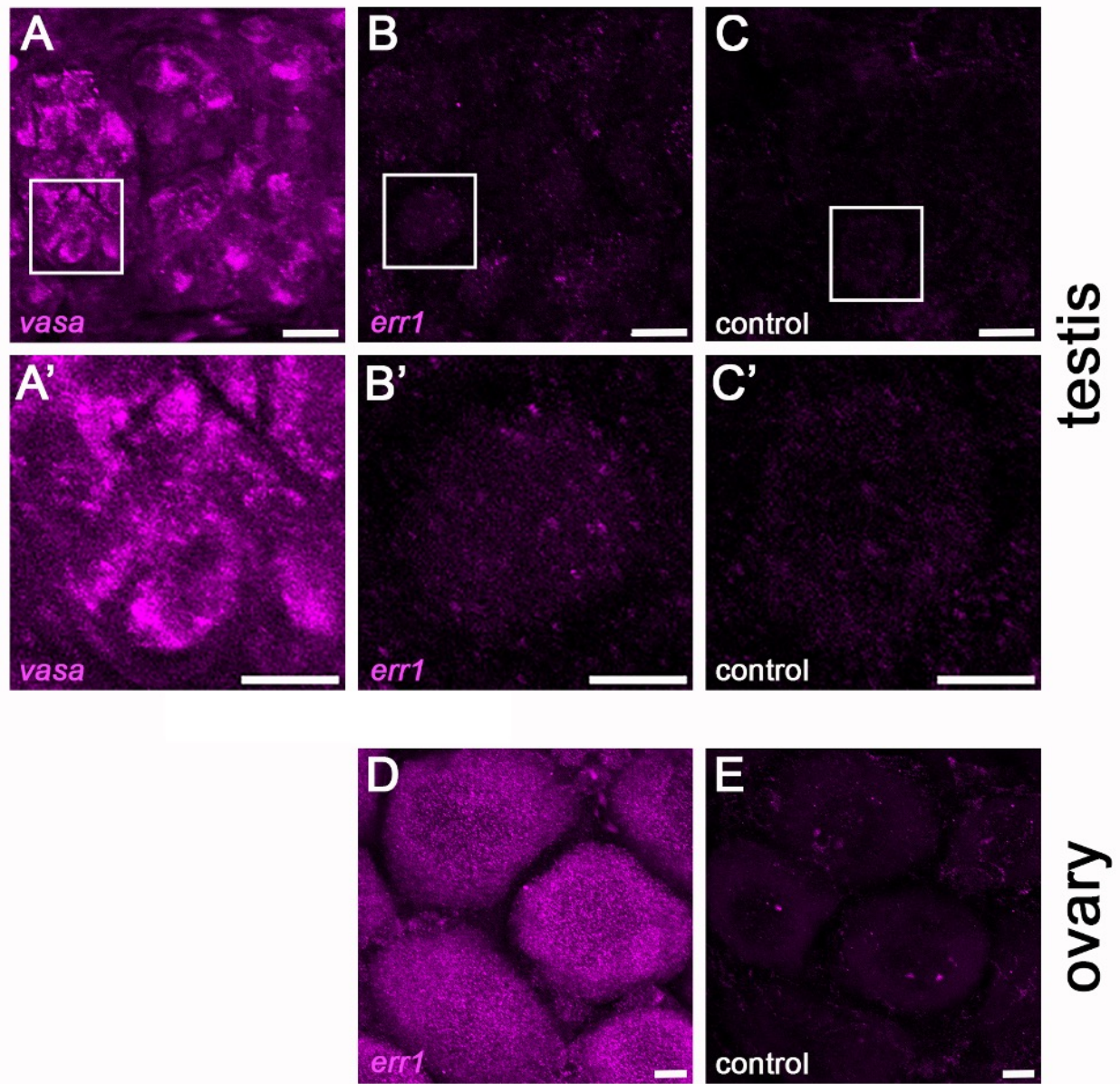


Fig. S8.

Expression of *vasa* and *err1* in developing male and female gonads.

Single channel (546 nm) images of merged images shown in Figure 4 showing expression of *vasa* (A, A') and *err1* (B, B', D). C, C' and E are negative controls with no HCR initiator oligos added. Scale bars are as described in Fig. 4. All images were captured with the same intensity settings to allow quantitative comparison.

BLAST RID: ZH12XEX8014 Query ID: Query_9148019 Database: nr

Tree method: Max Seq Difference: 0.85 Distance: Grishin (protein) Sequence Label: Sequence Title (if a) Hide legend

Fast Minimum Evolution all + TXT Tools Upload

Find: salamanders and caecilians | 7 leaves Multiple organisms | 36 leaves

steroid hormone receptor ERR1 [Epischotus calabaricus]
steroid hormone receptor ERR1 [Anolis sagrei]
steroid hormone receptor ERR1 [Anolis carolinensis]
steroid hormone receptor ERR1 [Sceloporus undulatus]
steroid hormone receptor ERR1 [Pogona vitticeps]
steroid hormone receptor ERR1 [Tiliqua scincoides]
steroid hormone receptor ERR1 [Rhinura floridana]
lizards & snakes | 3 leaves
steroid hormone receptor ERR1 [Zootoca vivipara]
steroid hormone receptor ERR1 [Eublepharis macularius]
steroid hormone receptor ERR1 [Heteronota binoc]
steroid hormone receptor ERR1 [Hemiscordylus capensis]
steroid hormone receptor ERR1 [Elgaria multicarinata webbi]
snakes | 3 leaves
steroid hormone receptor ERR1 isoform X2 [Varanus komodoensis]
steroid hormone receptor ERR1 isoform X1 [Varanus komodoensis]
lizards & snakes | 2 leaves
steroid hormone receptor ERR1 isoform X2 [Xenopus laevis]
steroid hormone receptor ERR1 isoform X1 [Xenopus laevis]
hypothetical protein G2086_08602 [Hyemochirus boettgeri]
steroid hormone receptor ERR1 [Xenopus tropicalis]
steroid hormone receptor ERR1 [Xenopus laevis]
steroid hormone receptor ERR1 isoform X2 [Pelodactylus fuscus]
steroid hormone receptor ERR1 isoform X1 [Pelodactylus fuscus]
steroid hormone receptor ERR1 [Bombina orientalis]
steroid hormone receptor ERR1 [Spina bombifrons]
steroid hormone receptor ERR1 [Engystomops pustulosus]
steroid hormone receptor ERR1 isoform X1 [Hyperolius rissigebachi]
steroid hormone receptor ERR1 isoform X2 [Hyperolius rissigebachi]
PREDICTED: steroid hormone receptor ERR1 [Nanorana parkeri]
steroid hormone receptor ERR1 [Bufo viridis]
steroid hormone receptor ERR1 [Bufo viridis]
steroid hormone receptor ERR1 [Physicophagus adspersus]
unnamed protein product [Staurois parvus]
Steroid hormone receptor ERR1 [Bufo viridis]
Steroid hormone receptor ERR1 [Bufo viridis]
steroid hormone receptor ERR1 isoform X2 [Bufo bufo]
steroid hormone receptor ERR1 isoform X1 [Bufo bufo]
steroid hormone receptor ERR1 [Bufo gargarizans]
steroid hormone receptor ERR1 isoform X2 [Eleutherodactylus coqui]
steroid hormone receptor ERR1 isoform X1 [Eleutherodactylus coqui]
LOW QUALITY PROTEIN: steroid hormone receptor ERR1 [Dendrobates tinctorius]
steroid hormone receptor ERR1 [Dendrobates tinctorius]
steroid hormone receptor ERR1 isoform X2 [Hyla sarda]
steroid hormone receptor ERR1 isoform X1 [Hyla sarda]
steroid hormone receptor ERR1 isoform X2 [Ranitomeya imitator]
steroid hormone receptor ERR1 isoform X1 [Ranitomeya imitator]
steroid hormone receptor ERR1 [Phyllolates terribilis]
steroid hormone receptor ERR1 [Rana temporaria]
steroid hormone receptor ERR1 [Aguaruna catenata]
steroid hormone receptor ERR1 [Pseudophryne corroboreae]
steroid hormone receptor ERXC1-like [Bombina bombina]

err1 = errα (autosomal) Reptiles
err1 = errα (autosomal) Anura
err1-like = err1Y B. bombina

Label color map
Query
From type material
Blast names color map
frogs & toads
caecilians
salamanders
coleocephalus
lungfishes
birds
bony fishes
lizards & snakes
snakes
unknown

Success Notes 2010 (selected) View port at (0.0) of 1057x740

Reset Tree

BLAST RID 71K9MW6014 Query ID: |Query_12256721 Database: refseq_protein

Tree method: Fast Minimum Evolution Max Seq Difference: 0.85 Distance: Grishin (protein) Sequence Label: Sequence Title (if a)

Hide legend

Label color map

- Query
- From type material

Blast names color map

- Frogs & toads
- Cacilians
- Salamanders
- Unknown

err1 = errα, Gymnophiona

err3 = errγ: Anura, Urodela, Gymnophiona

err2 = errβ: Urodela

err1-like = err1γ (autosomal) Anura

err1-like = err1γ B. bombina

steroid hormone receptor ERRI isoform X1 [Rhineura viviparum]

steroid hormone receptor ERRI [Geotrypetes seraphim]

steroid hormone receptor ERRI [Microcaecilia unicolor]

Frogs & toads | 9 leaves

Frogs & toads | 3 leaves

estrogen-related receptor gamma isoform X6 [Spina bombiformis]

Frogs & toads | 12 leaves

Frogs & toads | 3 leaves

estrogen-related receptor gamma isoform X3 [Hyperolius riggenbachii]

estrogen-related receptor gamma isoform X6 [Pseudophryne corroboree]

Frogs & toads | 11 leaves

estrogen-related receptor gamma isoform X5 [Bufo bufo]

estrogen-related receptor gamma isoform X4 [Bufo bufo]

estrogen-related receptor gamma isoform X3 [Bufo gargarizans]

Frogs & toads | 4 leaves

estrogen-related receptor gamma isoform X5 [Pelobates fuscus]

estrogen-related receptor gamma isoform X4 [Pelobates fuscus]

estrogen-related receptor gamma isoform X3 [Pelobates fuscus]

estrogen-related receptor gamma isoform X2 [Pelobates fuscus]

estrogen-related receptor gamma isoform X1 [Bombina orientalis]

estrogen-related receptor gamma isoform X2 [Bombina orientalis]

estrogen-related receptor gamma isoform X1 [Bombina orientalis]

estrogen-related receptor gamma isoform X2 [Microcaecilia unicolor]

estrogen-related receptor gamma isoform X1 [Rhineura viviparum]

estrogen-related receptor gamma isoform X4 [Geotrypetes seraphim]

estrogen-related receptor gamma isoform X2 [Geotrypetes seraphim]

estrogen-related receptor gamma isoform X1 [Microcaecilia unicolor]

estrogen-related receptor gamma isoform X1 [Pseudoeurycea waltii]

estrogen-related receptor gamma isoform X1 [Amphystoma mexicanum]

estrogen-related receptor gamma-like isoform X1 [Amphystoma mexicanum]

estrogen-related receptor gamma-like isoform X2 [Amphystoma mexicanum]

steroid hormone receptor ERRI [Bufo gargarizans]

steroid hormone receptor ERRI isoform X2 [Bufo bufo]

steroid hormone receptor ERRI isoform X1 [Bufo bufo]

steroid hormone receptor ERRI isoform X2 [Eleutherodactylus coqui]

steroid hormone receptor ERRI isoform X1 [Eleutherodactylus coqui]

LOW QUALITY PROTEIN: steroid hormone receptor ERRI [Dendrobates]

steroid hormone receptor ERRI [Dendrobates ebraecatus]

steroid hormone receptor ERRI isoform X2 [Hyla sarda]

steroid hormone receptor ERRI isoform X1 [Hyla sarda]

steroid hormone receptor ERRI isoform X2 [Rhinophrynus imitator]

steroid hormone receptor ERRI isoform X1 [Rhinophrynus imitator]

steroid hormone receptor ERRI [Phyllomedusa terribilis]

steroid hormone receptor ERRI [Rana siliacea]

steroid hormone receptor ERRI [Aquarana castibonensis]

steroid hormone receptor ERRI [Pseudophryne corroboree]

steroid hormone receptor ERRI isoform X1 [Hyperolius riggenbachii]

steroid hormone receptor ERRI isoform X2 [Hyperolius riggenbachii]

PRE-ICTED: steroid hormone receptor ERRI [Nannophrys parkeri]

steroid hormone receptor ERRI [Physiphallus adspersus]

steroid hormone receptor ERRI [Nannophrys parkeri]

steroid hormone receptor ERRI isoform X2 [Pelobates fuscus]

steroid hormone receptor ERRI isoform X1 [Pelobates fuscus]

steroid hormone receptor ERRI [Bombina orientalis]

steroid hormone receptor ERRI [Spina bombiformis]

steroid hormone receptor ERRI [Xenopus laevis]

steroid hormone receptor ERRI [Xenopus tropicalis]

steroid hormone receptor ERRI isoform X2 [Xenopus laevis]

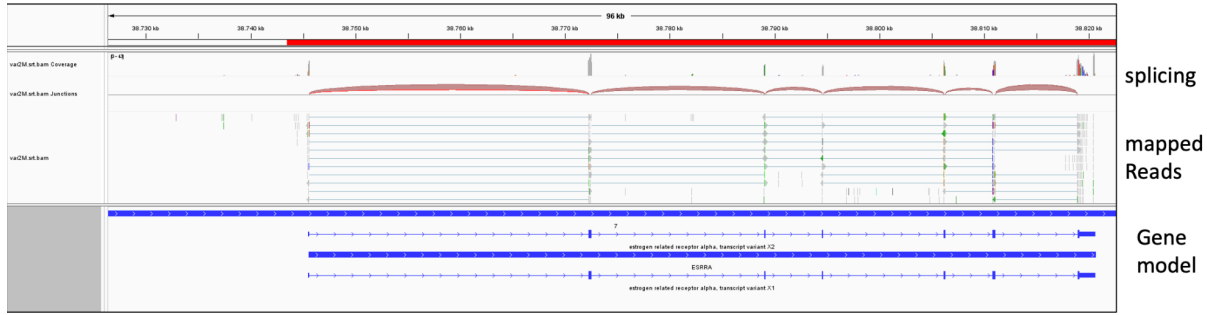
steroid hormone receptor ERRI isoform X1 [Xenopus laevis]

steroid hormone receptor ERRI-like [Bombina orientalis]

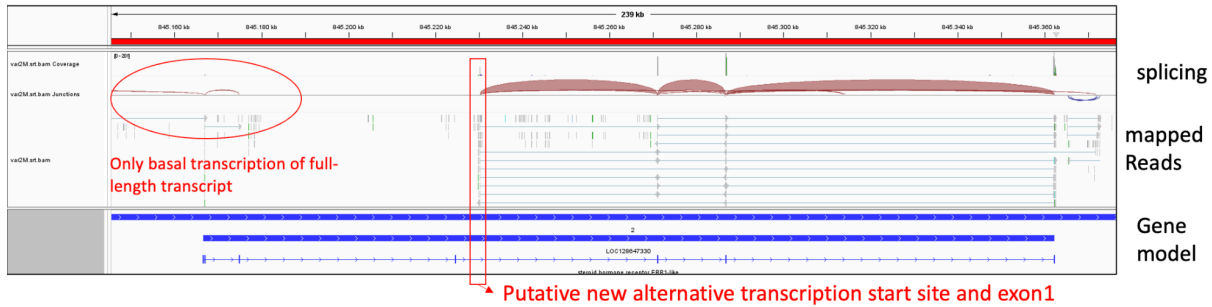
20

Err1 autosomal, RNAseq coverage of gene model 100%:

Mapped RNAseq reads from *B. variegata* transcriptome Bvar2M



Err1-Y, RNAseq coverage of gene model <50% (only 3 terminal exons (ligand binding domain), ERE binding domain missing!):



>Err1Y-fragment de novo assembled transcript-60_1

GAGTGTCAAAGTGACCAATCTCTGAAAGGGAAATGCGAAGCTGCGGACGGGTGGCAGCATTGACCGGCCCTCTATAGTCCATGAATGCAGTGGTATCATATGGCTGGTGGTAAATGGATC
AGTTGCTTCCCATGTCAAATCCAGTCTGACAGATGCATACCTGAACACCATGAACACTCTGCGTGACCTTGTGACCAAGAGATTGTTAACATAATCAGCTGGGCCAACAACTCCGGGCTTCTCTCCCTCTCTCA
GTCAGACCAAGAGCCTTTTGGGAAGTGCCTGGATGGAGGCTCTCTGCTTTGCGTGATATTTCGCTCCTTACCCTACGAAGACAAGTTGATTGTCAGAGGATATTGCTGCAAGGATCTTCATGCTCTG
CTGCTCTCTGACCTTTGCTCCGACATCTTGCATCTTATACGCAAGTACGAGACCTGTATTGGAAAGGGAGGAGTACAACATTCTGAAAGGCCTCGCCTGACTAATCAGATTCTGTGGACATAGAGGATCCAG
AAGCGGTTCTTCAGCTCCGTGACAATCTATGGGAAACCTCAATGATTATGAGTCAACAGCATCCAGAAAGAACCGCGCATGAAAGGCTGCTGTAACACTCCCACTCTGCGCCAGATGGCTGGACATGTT
ATGACGATTTTATAGTCAATAAATATCCAAAAAGCACAGTCTGTTGTATAACATGATTAAATCTATCTAGCATTATTAGTGTTCATTAAAGAAATACCATGAAAACCACTTGCAGGCTGCAATAGAAA
AACAAAGTACTAGAAATAAGAAAAATAGCATTGTGATTCTTAAGGCATCAAGCCGAATATGGCCGAGAACACAGCATTACGCTTTTATGTCAGATTATTAGTGTTCATTAAAGAAATACCATGAAAACCACTTGCAGGCTGCAATAGAAA
AAATCGTATTCGCATTGTGTTTGCATGTGCTGAAATAGCCAAATGCTGCTGACTGCGGCAATATTGACATTGTTCTATGAATCTGAAATTAATTTTGTCTGTGCAAAACAAAGTAAATTTGTACTTTATAATA
TGATATTAAATATGCACATTAATAAT

>|cl|ORF1 from *err1Y* fragment

MNAIVSYWLVKLDQLPMSNPVLTDAYLNTMNTLRDLADQEIWNISWANNIPGFSLSQSDQKLLRSASWMEVFLLCVIFRSLPYEDKVVFAEDIVLHKDSSCARLSDLCSILHLIRKYQTLYLEREYNILKGLALNTSD
SVDIEDPEAVLQLRDLNWLWETLNDYESNQHPPEPRRHERLLTLPRLQRMAGHVMQHFISQ

Putative conserved domains

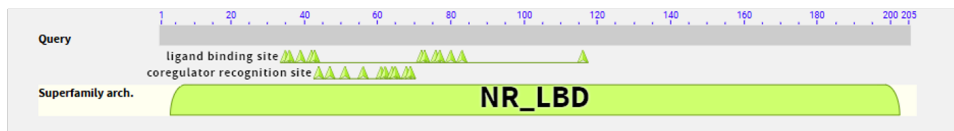


Fig. S10.

Mapping reads from the testis transcriptome of juvenile *B. variegata* to the original gene models of the reference genome of *B. bombina* for the autosomal *err1* and the male-specific *err1Y*. While the RNAseq coverage completely fitted the gene model for the autosomal gene (*err1*), reads for *err1Y* suggested only basal transcription at the 5'-end and the greatest expression covering the three terminal, originally predicted exons with the ligand binding domain were expressed. *Err1Y* expression was 5–6 × higher than *err1* or *era*, and *de novo* assembly recovered only this truncated isoform. The transcript contains an intact open reading frame with an ATG start in originally predicted exon 5, 94 bp downstream of the new transcription start site, encoding a protein that retains the ligand-binding (LBD) but lacks the DNA-binding domain (DBD) for estrogen response elements (ERE).

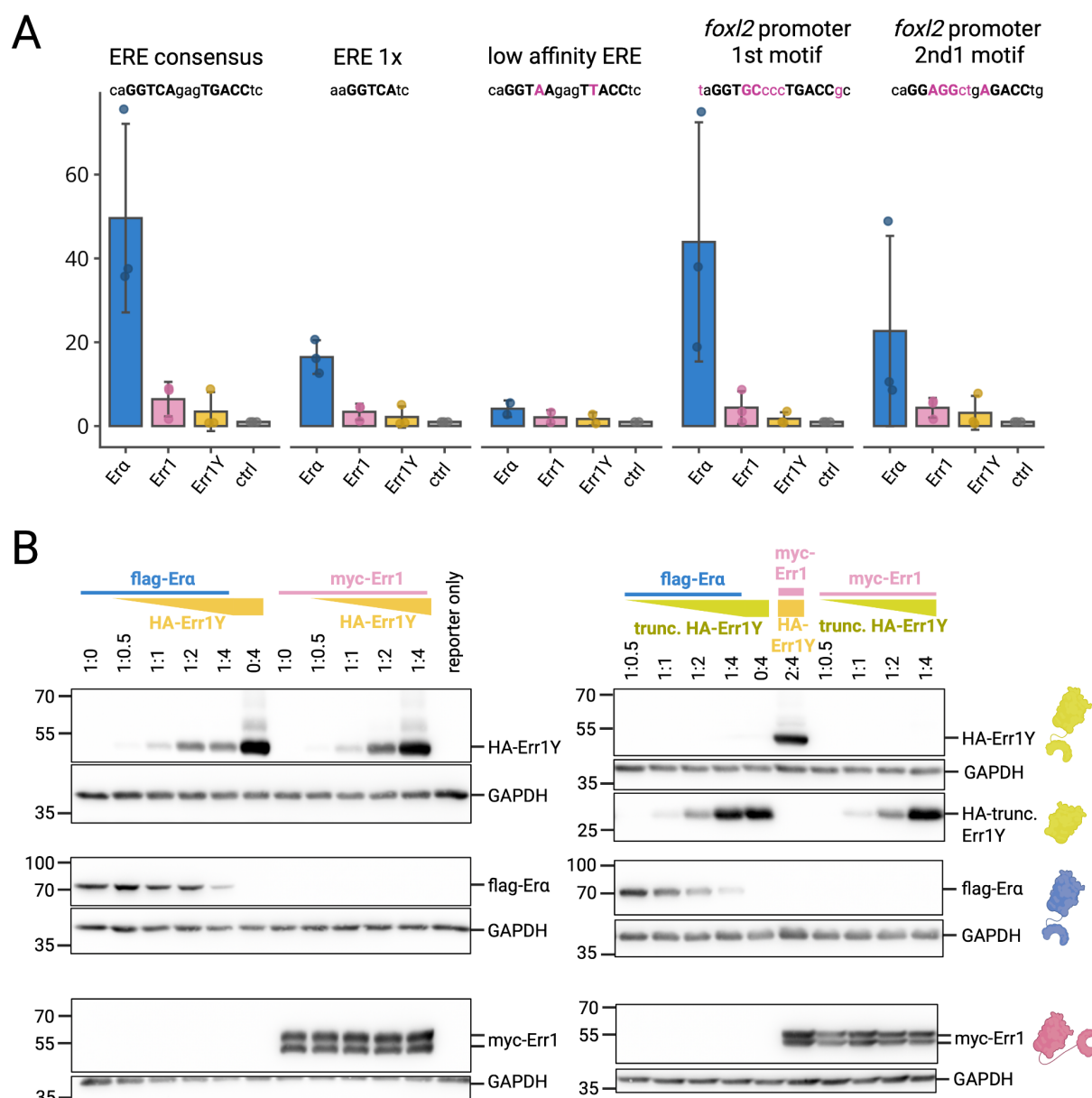


Fig. S11.

A: Luciferase reporter assays in HEK cells transfected with flag-Era, myc-Err1, HA-Err1Y, or empty pcDNA3.1(-) vector (ctrl) alongside reporters driven by an ERE consensus sequence, a single ERE half-site (ERE 1x), a mutated non-functional ERE (Bad ERE), or two *foxl2* promoter-proximal ERE-containing sequences (*foxl2* closest promoter motif and second closest motif). ERE core motifs are highlighted in capitals and bold, deviations from consensus in magenta. **B:** Western blot validation of protein expression in the luciferase reporter gene assays (Fig. 5B) using antibodies against the short peptide tags. HEK cells were transfected with titrated ratios of flag-ERα and HA-Err1Y (left panels) or flag-ERα, myc-Err1, and truncated HA-Err1Y (right panels) at the indicated molar ratios (1 = 50ng DNA). GAPDH serves as a loading control. Schematic representation and ratios are shown above the blot corresponding to Fig. 5B.

Taxon (or interspecies hybrid)	Origin	Phenotypic sex established				Remarks
		Approach	Females	Males	Un- known	
<i>B. bombina</i>	Wodzisław Małopolski, Poland	Secondary sex. characters anatomy	46	33	13	Family with parents and 90 offspring
<i>B. bombina</i>	Kreba, Saxony, Germany	Secondary sex. characters	4	9	8	Buccal swab sampling during migration
<i>B. bombina</i>	Thuringia, Germany	Secondary sex. characters	1	1	0	Phenotypic sex determined by collector
<i>B. variegata variegata</i>	unknown, bred in captivity	Reproduction	3	2	0	1 albino ♂
<i>B. variegata variegata</i>	Waksmund and Biała Woda, Poland	Secondary sex. characters	10	12	0	Family with 22 offspring
<i>B. variegata variegata</i>	unknown, bred in captivity, Balkans	Anatomy	3	9	0	Gosner 46, at metamorphosis
<i>B. variegata variegata</i>	unknown, bred in captivity, Balkans	Anatomy	6	7	1	Gosner 46, 4 months after metamorphosis
<i>B. variegata</i> (nominal taxon 'kolombatovici')	Zadar, Croatia	Secondary sex. characters	3	2	0	-
<i>B. v. pachypus</i>	Calabria, Italy	Secondary sex. characters	1	1	0	-
<i>B. v. scabra</i>	North of Greece, bred in captivity	Secondary sex. characters	2	2	0	-
<i>B. orientalis</i>	unknown, bred in captivity	Anatomy	11	9	4	Adult, 1 ♂ Gosner 46, 6 months after metamorphosis
<i>B. orientalis</i>	unknown, bred in captivity	Anatomy	8	4	1	Gosner 46, at metamorphosis
<i>B. orientalis</i>	unknown, bred in captivity	Anatomy	3	3	0	Gosner 46, 1-2 month after metamorphosis
<i>B. maxima</i>	unknown, bred in captivity	Reproduction	2	3	0	-
<i>B. microdeladigitora</i>	unknown, bred in captivity	Anatomy or external morphology	0	2	4	-
Hybrids						
<i>B. bombina</i> ♀ ¹ x <i>B. variegata</i> ♂ ²	F ₁ bred in captivity, parents: ♀ ¹ Wodzisław Małopolski, (ERS3926743), ♂ ² Obidowa, near Nowy Targ, (ERS3926742), Poland	Reproduction, anatomy	11	17	-	-
<i>B. microdeladigitora</i> ♀ x <i>B. maxima</i> ♂	F ₁ bred in captivity	External morphology	0	1	-	-
<i>B. variegata</i> ♀ x <i>B. maxima</i> ♂	F ₁ bred in captivity	External morphology	0	1	-	-

Table S1.

Overview of animals from the genus *Bombina* of which DNA-samples and phenotypic sex information were available for the present study. The multiplex PCR-test (autosomal primers Bomb_Auto_397bp_F/R and male-specific BombYonly_300bp_F/R) was successfully applied in all species, except *B. maxima* and *B. microdeladigitora*, where the Y-marker did not yield sex-specific PCR-products.

Autosomal gene on LG 7

Gene: *err1 (esrra)*
Name: *estrogen related receptor alpha*
RNA title: mRNA-estrogen related receptor alpha
Protein title: steroid hormone receptor ERR1
Merged features: XM_053720047.1; XP_053576022.1
Location (chr7): 38,745,492...38,820,666 [Length]
Span on NC_069505.1: 75,175 nt
Aligned length: 2,919 nt (including UTR)
CDS length: 1,233 nt
Protein length: 410 aa

Y-specific gene on LG 2

Gene: *err1Y (err1-like)*
Name: *steroid hormone receptor err1-like*
RNA title: mRNA-steroid hormone receptor err1-like
Protein title: steroid hormone receptor err1-like
Merged features: XM_053700116.1 and XP_053556091.1
Location (chr2): 845,166,599...845,362,339 [Length]
Span on NC_069500.1: 195,741 nt
Aligned length: 1,212 nt (excluding UTR as not annotated)
CDS length: 1,212 nt
Protein length: 403 aa

Table S2.

Comparison of gene parameters of autosomal gene *err1* (= *esrra*) on linkage group 7 and Y-specific *err1Y* (= *err1-like*) on linkage group 7 (homologous to *Xenopus tropicalis* LG1) in *Bombina bombina*.

Protein matches <i>X. tropicalis</i> => <i>B. bombina</i>	<i>X. tropicalis</i> accession	<i>X. tropicalis</i> chromosome	<i>B. bombina</i> accession	<i>B. bombina</i> chromosome
2729	NC_030684.2	XT chr8	NC_069499.1	BB chr1
2008	NC_030686.2	XT chr10	NC_069499.1	BB chr1
1368	NC_030685.2	XT chr9	NC_069499.1	BB chr1
889	NC_030680.2	XT chr4	NC_069499.1	BB chr1
5441	NC_030677.2	XT chr1	NC_069500.1	BB chr2 XY sex chromosomes, this paper
4776	NC_030678.2	XT chr2	NC_069501.1	BB chr3
227	NC_030679.2	XT chr3	NC_069502.1	BB chr4
252	NC_030685.2	XT chr9	NC_069502.1	BB chr4
3753	NC_030681.2	XT chr5	NC_069502.1	BB chr4; Nürnberg <i>et al.</i> (2021): "B. variegata sex chromosome"; therein as "LG5"
3516	NC_030682.2	XT chr6	NC_069503.1	BB chr5
272	NC_030685.2	XT chr9	NC_069503.1	BB chr5
207	NC_030684.2	XT chr8	NC_069504.1	BB chr6
5009	NC_030679.2	XT chr3	NC_069504.1	BB chr6
2324	NC_030680.2	XT chr4	NC_069505.1	BB chr7
828	NC_030684.2	XT chr8	NC_069505.1	BB chr7
297	NC_030685.2	XT chr9	NC_069505.1	BB chr7
205	NC_030679.2	XT chr3	NC_069505.1	BB chr7
1954	NC_030683.2	XT chr7	NC_069506.1	BB chr8
1897	NC_030683.2	XT chr7	NC_069507.1	BB chr9
1127	NC_030680.2	XT chr4	NC_069508.1	BB chr10
1445	NC_030685.2	XT chr9	NC_069509.1	BB chr11
1079	NC_030684.2	XT chr8	NC_069510.1	BB chr12

Table S3.

Synteny analysis of the *B. bombina* reference genome (GCF_027579735.1) linkage groups compared to those of *X. tropicalis* (GCF_000004195.4).

Gene	Oligo-ID_GENE-ABBR_No	Sequence
<i>err1Y</i>	CW100_B1-BV_ERR1Y_1A	GAGGAGGGCAGCAAACGGAATGTCCAGCCGTCTGGCGCAGGAGT
<i>err1Y</i>	CW101_B1-BV_ERR1Y_1B	TTATTGACTAATAAAATGCTGCATATAGAAGAGTCTTCCTTTACG
<i>err1Y</i>	CW102_B1-BV_ERR1Y_2A	GAGGAGGGCAGCAAACGGAACGTGGTTCTTCTGGATGCTGGTTTG
<i>err1Y</i>	CW103_B1-BV_ERR1Y_2B	AGTGTTAGCAGCATCCTTTCATGCCTAGAAGAGTCTTCCTTTACG
<i>err1Y</i>	CW104_B1-BV_ERR1Y_3A	GAGGAGGGCAGCAAACGGAATGTCACGGAGCTGAAGAACGGCTTC
<i>err1Y</i>	CW105_B1-BV_ERR1Y_3B	CATAATCATTGAGGGTTTCCCATAGTAGAAGAGTCTTCCTTTACG
<i>err1Y</i>	CW106_B1-BV_ERR1Y_4A	GAGGAGGGCAGCAAACGGAAGTCAGGGCGAGGCATTTCAGAATG
<i>err1Y</i>	CW107_B1-BV_ERR1Y_4B	ATCCTCTATGTCCACAGAATCTGAATAGAAGAGTCTTCCTTTACG
<i>err1Y</i>	CW108_B1-BV_ERR1Y_5A	GAGGAGGGCAGCAAACGGAATGGTACTTGCGTATAAGATGCAAGA
<i>err1Y</i>	CW109_B1-BV_ERR1Y_5B	TACTCCTCCCTTTCCAAATACAGGGTAGAAGAGTCTTCCTTTACG
<i>err1Y</i>	CW110_B1-BV_ERR1Y_6A	GAGGAGGGCAGCAAACGGAAGCATGAAGAATCCTTGTGTAGCAC
<i>err1Y</i>	CW111_B1-BV_ERR1Y_6B	CGGAGCAAAGGTCAGAGAGACGAGCTAGAAGAGTCTTCCTTTACG
<i>err1Y</i>	CW112_B1-BV_ERR1Y_7A	GAGGAGGGCAGCAAACGGAAGTAGGGTAAGGAGCGAAATATCACG
<i>err1Y</i>	CW113_B1-BV_ERR1Y_7B	ATCCTCTGCAAATACAACCTTGCTTAGAAGAGTCTTCCTTTACG
<i>err1Y</i>	CW114_B1-BV_ERR1Y_8A	GAGGAGGGCAGCAAACGGAATGCAAAAGGCTCTTCTGGTCTGACT
<i>err1Y</i>	CW115_B1-BV_ERR1Y_8B	AGCAGGAAGACCTCCATCCAGGCGCTAGAAGAGTCTTCCTTTACG
<i>err1Y</i>	CW116_B1-BV_ERR1Y_9A	GAGGAGGGCAGCAAACGGAATGGCCCAGCTGATTATGTTAACAAT
<i>err1Y</i>	CW117_B1-BV_ERR1Y_9B	AGAGGGAGGAGAAGCCCAGGAATGTTTAGAAGAGTCTTCCTTTACG
<i>err1Y</i>	CW118_B1-BV_ERR1Y_10A	GAGGAGGGCAGCAAACGGAACATGGTGTTTCAGGTATGCATCTGTC
<i>err1Y</i>	CW119_B1-BV_ERR1Y_10B	TTGGTCAGCAAGGTCACGCAGAGTGTAGAAGAGTCTTCCTTTACG

<i>err1Y</i>	CW120_B1-BV_ERR1Y_11A	GAGGAGGGCAGCAAACGGAATCCAATTTTACCACCAGCCAATATG
<i>err1Y</i>	CW121_B1-BV_ERR1Y_11B	ACTGGATTGACATGGGAAGCAACTTAGAAGAGTCTTCCTTTACG
<i>err1</i>	CW21_B2_ERR1_20A	CCTCGTAAATCCTCATCAAATCTTGCTCCTCGTCCCTCCACGGC
<i>err1</i>	CW22_B2_ERR1_20B	ACATACTTGCCCCGTCCTGGGGTGTAATCATCCAGTAAACCGCC
<i>err1</i>	CW23_B2_ERR1_19A	CCTCGTAAATCCTCATCAAAACACAGTCGCTTAGGAATGGAGTTC
<i>err1</i>	CW24_B2_ERR1_19B	ACCAGATGCCACATCTCCGCACACTAAATCATCCAGTAAACCGCC
<i>err1</i>	CW25_B2_ERR1_21A	CCTCGTAAATCCTCATCAAAGACCGCTACTGTCACTAGAGCCGCT
<i>err1</i>	CW26_B2_ERR1_21B	ACCGGTTCCCTGCAAGTCAGGTCCAAATCATCCAGTAAACCGCC
<i>err1</i>	CW27_B2_ERR1_14A	CCTCGTAAATCCTCATCAAATGTAGAACATCTCCGTCTGGTCTTC
<i>err1</i>	CW28_B2_ERR1_14B	TGCTGGGCCTGCCCCCTGGGTTATAAATCATCCAGTAAACCGCC
<i>err1</i>	CW29_B2_ERR1_12A	CCTCGTAAATCCTCATCAAAACAATTATCCGGTTCTGCCACCAG
<i>err1</i>	CW30_B2_ERR1_12B	CCGGCAAAGCTGGGTCTGGCATGGCAAATCATCCAGTAAACCGCC
<i>err1</i>	CW31_B2_ERR1_6A	CCTCGTAAATCCTCATCAAATCAGGGTCCGGTACTTGCGTACTAG
<i>err1</i>	CW32_B2_ERR1_6B	GCATGACGTATTCTTCCCTTTCCACAAATCATCCAGTAAACCGCC
<i>err1</i>	CW33_B2_ERR1_13A	CCTCGTAAATCCTCATCAAATTGTTTTTAACTACTGCTCCAGAT
<i>err1</i>	CW34_B2_ERR1_13B	GTGCGACACCACTGCATTGGAATCATCCAGTAAACCGCC
<i>err1</i>	CW35_B2_ERR1_4A	CCTCGTAAATCCTCATCAAACTGTAGAGTGTCACGGAGCCGAAGA
<i>err1</i>	CW36_B2_ERR1_4B	GTTTGACTCATACTCACTGAGGGCTAAATCATCCAGTAAACCGCC
<i>err1</i>	CW37_B2_ERR1_5A	CCTCGTAAATCCTCATCAAATCTGAGTTAGTGAGGGTGAGGGCTT
<i>err1</i>	CW38_B2_ERR1_5B	GCCTCAGGGTCCTCTATGTGCACAGAAATCATCCAGTAAACCGCC
<i>err1</i>	CW39_B2_ERR1_7A	CCTCGTAAATCCTCATCAAAACGAGCAGAGCGTGAAGACTCCTCG
<i>err1</i>	CW40_B2_ERR1_7B	CAAGATGCAGGAGCAGAGGTCCGAGAAATCATCCAGTAAACCGCC
<i>vasa</i>	CW41_B2_Bo_VASA_1A	CCTCGTAAATCCTCATCAAAAGGTACCTCTTGTTGAGCGTCACTT
<i>vasa</i>	CW42_B2-Bo_VASA_2A	CCTCGTAAATCCTCATCAAAATATCTTCATCATCTTTCTATCAA

<i>vasa</i>	CW43_B2-Bo_VASA_3A	CCTCGTAAATCCTCATCAAAAACGACCAGTGCGTCCAATTCTGTG
<i>vasa</i>	CW44_B2-Bo_VASA_4A	CCTCGTAAATCCTCATCAAAAAGATTAATTACATGTTGAACATTT
<i>vasa</i>	CW45_B2-Bo_VASA_5A	CCTCGTAAATCCTCATCAAAGTAGCTACAAGTATAGGACACTCTC
<i>vasa</i>	CW46_B2-Bo_VASA_6A	CCTCGTAAATCCTCATCAAAATCTCCTGTTGAGGTCTGTCACCATG
<i>vasa</i>	CW47_B2-Bo_VASA_7A	CCTCGTAAATCCTCATCAAAAAGAATAAGAGCAATATAATCTGCT
<i>vasa</i>	CW48_B2-Bo_VASA_8A	CCTCGTAAATCCTCATCAAACGTTTACCTATGGTATGTAGAA
<i>vasa</i>	CW49_B2-Bo_VASA_9A	CCTCGTAAATCCTCATCAAAGTCCAACCTCAAGAATGGTTTGTTT
<i>vasa</i>	CW50_B2-Bo_VASA_10A	CCTCGTAAATCCTCATCAAAAACCGTGAGAAATAAATAATCTGGT
<i>vasa</i>	CW51_B2-Bo_VASA_1B	ACTGAAAGCAATTTCTCTAACCACAAATCATCCAGTAAACCGCC
<i>vasa</i>	CW52_B2-Bo_VASA_2B	ACTTTTACGAGTGAACGAGCAATACAAATCATCCAGTAAACCGCC
<i>vasa</i>	CW53_B2-Bo_VASA_3B	ATGAAATTGCTTTTCCAATGTTGCCAAATCATCCAGTAAACCGCC
<i>vasa</i>	CW54_B2-Bo_VASA_4B	ATATTCATCAATTTCTTTAGGCAGAAAATCATCCAGTAAACCGCC
<i>vasa</i>	CW55_B2-Bo_VASA_5B	ATATCCAAGCCTCTTGACGCTACTGAAATCATCCAGTAAACCGCC
<i>vasa</i>	CW56_B2-Bo_VASA_6B	AACGAAATTTCCAGAGGGCTGTTTCAAATCATCCAGTAAACCGCC
<i>vasa</i>	CW57_B2-Bo_VASA_7B	GCTGGTTGAAGGAAGCTGTTCTGAAAATCATCCAGTAAACCGCC
<i>vasa</i>	CW58_B2-Bo_VASA_8B	TTTTTTGTCTTTACAAACACCAAAGAAATCATCCAGTAAACCGCC
<i>vasa</i>	CW59_B2-Bo_VASA_9B	CAAGCAGCTTGCTTTTCTTTTCAAAAAATCATCCAGTAAACCGCC
<i>vasa</i>	CW60_B2-Bo_VASA_10B	ATCACTACAAGCTCCGCCAACTTGAAAATCATCCAGTAAACCGCC

Table S4.

Initiator oligos used for Hybrid Chain Reaction (HCR) to detect *err1*, *err1Y* and *vasa* expression in primordial gonads of *Bombina variegata*.

Primer name	Sequence
ERE-consensus2x-fw	caGGTCAgagTGACCtc
ERE-consensus2x-rev	tcgagaGGTCActcTGACCtggtac
FOXL2-promoter-I-fw	ctaGGTGCcccTGACCgcc
FOXL2-promoter-I-rev	tcgaggcGGTCAgggGCACCTaggtac
FOXL2-promoter-II-fw	ccaGGAGGctgAGACCtgc
FOXL2-promoter-II-rev	tcgagcaGGTCTcagCCTCCTgggtac
ERE-consens1x-fw	cgtgataaGGTCAtc
ERE-consens1x-rev	tcgagaTGACCttatcacggtac
ERE-double-BAD-fw	caGGTAAgagTTACCtc
ERE-double-BAD-rev	tcgagaGGTAActcTTACCtggtac

Table S5.

Sequences of primers used for estrogen response elements (ERE) and estrogen-related receptor response element (ERRE) motifs cloning in firefly luciferase reporter constructs.

HA-Err1Y	GCTAGCGCCACCATGTATCCATACGATGTTCCAGATTACGCTGGTGGTGGTAGCGAATTCGGCGGCGGGGATCCCGGCAAGGAGGCCGCCATCAGCTTCCCCACCTGCAGTGCGGCGGCCACCTGAACCAGATCAGCAGGGCCTGCTGGACGGCGAGCCCGTGTTCGTGATCTGCACCGCCGGCCAGCAGTACGCCGTGGACCTGTGCAATCAAGGGCCTGCACAAGCCCTGGACCACCAAGAGCATCGGCCACCAGATCCCCAGCAGCGGCCACGGCCCGACATGCAGGGCACCCTGTGCTGCAACACCAGCTACAAGGAGCAGCTGTGCTGGTGTGCGGCACGAGGGCCCTGGGCTACCACTACGGCGTGAGCAGCTGCGAGGCCTGCAAGGCCTTCTTCATGAGGACCATCAGGGCTACATCGAGTACAGGTGCAGGGCCAGCAACGACTGCAAGATCATCAAGAGGCACAGGAACGTGTGCAAGGGCTGCAGGTTCCGCAAGTGCCTGCTGGTGGGCATGAAGAAGGAGTGCGTGAAGCTGGACAGGGTGAA GTGCGGCCAGAGAAGAACTACAAGAGGAAGCCAGCGACGTGGTGTGGACAACCCCCAGAGCCACATCCAGCAGGGCAGCGCCACCCTGGTGACCAAGCAGGATCCCATGAACGCCGTGGTGAGCTACTGGCTGGTGGTGAGCTGGACCAGCTGCTGCCCATGAGCAACCCCGTGTGACCGACGCCTACCTGAACACCATGAACACCCTGAGGGACCTGGCCGACCAGGAGATCGTGAACATCATCAGCTGGGCCAACAACATCCCCGGCTTCAGCAGCCTGAGCCAGAGCGACCAGAAGAGCCTGCTGAGGAGCGCCTGGATGGAGGTGTTCTGCTGTGCGTGATCTTCAGGAGCCTGCCCTACGAGGACAAGGTGGTGTTCGCCGAGGACATCGTGTGCACAAGGACAGCAGCTGCAGCGCCAGGCTGAGCGACCTGTGCAGCGACATCCTGCACCTGATCAGGAAGTACCAGACCCTGTACCTGAGAGGGAGGAGTACAACATCCTGAAGGGCCTGGCCCTGACCAACAGCGACAGCGTGGACATCGAAGACC CCGAGGCCGTGCTGCAGCTGAGGGACAACCTGTGGGAGACCCTGAACGACTACGAGAGCAACCAGCACCC CGAGGAGCCCAGGAGGCACGAGAGGCTGTGCTGACCCTGCCCTGCTGAGGCAGATGGCCGGCCACGTG ATGCAGCACTTCATCAGCCAGTGA
flag-Era	GCTAGCGCCACCATGCACTACAAAGACGATGACGACAAGGGTGGTGGTAGCGAATTCGGTGGTGGTGGATCGACCATGCCCTGCACAGCAAGGCCCTAGAGTGGCCCTGCTGCATCAGATTCAGACCACCGAGCTGGA GAACAGAAACAGAACCCCCGTGAAGATCCCCCTGGAGAGACCCCTGGGCGAGATGTACGTGGAGAACAAC AGAACCGGCATCTTCAACTACCCCGAGGGCGCCACCTACGACTTTGCCGCTGCCGCTCCCGTGTACTCCT CCGCCTCCCTGAGCTACGCTGCTTCTCCGAGAGCTTCGGCAGCGGCAGCCTGAGCGGCCCTGCACAGCCT GAACAACGTGCCCTTAGCCCCGTGGTGTTCCTGCAGACCGCCCCCTCAGATCAGCCCCCTTCATCCACCAT CACAGCCAACAAGTGCCCTACTACCTGGAGAGCGAGCAAGGCAACTTCGGCGTGAGAGAGAGCGCCCCC CCACCTTCTACAGAAGCAGCAGCGACAACAGAAGACAGAGCGGCAGAGAGAGAGTGAACACCGCCAACGA

	CAAGGGCCCCCGCCATGGAGAGCACCAAGGAGACAAGATACTGCGCCGTGTGCAGCGACTACGCCTCC GGCTACCACTACGGCGTGTGGAGCTGCGAGGGCTGCAAGGCCCTTCTTCAAGAGAAGCATCCAAGGCCACA ACGACTACATGTGCCCCGCCACCAATCAGTGCACCATCGACAAGAACAGAAAGAGCTGCCAAGCCTG CAGACTGAGAAAAGTGCTACGAGGTGGGCATGATGAAGGGCGGCATCAGAAAAGGAGAGAAGAGCGGCAGAG ATGATCAAGCACAAAGAGACAGAAGGAGGAGCAAGAGACCAAGAACGAGGTGGACACAAGCGAGATCAGAA CCACAAGCATCTGGGTGAACCCACCGTGAATCCATCAAACCTCAGCCCCGTGCTGAGCCTGAGCGCCGA GCAGCTGATCAGCGCCCTGATGGAGGCCGAGCCCCCATCGTGTACAGCGAGTACGACAGCACCAAGCCC CTGAGCGAGGCCCTCCATGATGACCTGCTGACCAACCTGGCCGACAAGGAGCTGGTGCACATGATCAACT GGGCCAAGAGAGTGCCCGGCTTCGTGGACCTGACCCTGCACGACCAAGTGCACCTGCTGGAGTGCGCCTG GCTGGAGATCCTGATGGTGGGCTGATCTGGAGAAGCATGGAGCACCCCGCAAGCTGAGCTTCGCCCCC AATCTGCTCCTGGACAGAAACCAAGGCAGATGCGTGGAGGGCTGGTGGAGATCTTCGACATGCTGGTGA CCACCGCCACAAGATTGAGAATGATGAGACTGAGAGGCGAGGAGTTTATCTGCATGAAGAGCATCATTCT CCTGAACAGCGGCGTGTACACCTTCCTGAGCAGCACCTGGAGACCCTGGAGGACACCGATCAGATCCAC CTGATCCTGGACAAGCTGATCGACACCCTGGTGCACCTTCATGGCCAAGAGCGCCTGAGCCTGCAGCAAC AGCAGAGAAGACTGGCTCAGCTGCTCCTGATCCTGAGCCACGTGAGACACATGAGCAACAAGGGCATGGA ACACCTCTATAGCATGAAGTGAAGAAGCTGGTGGCCCTGTACGACCTGCTCCTGGAGATGCTGGAGCGCC CAGAGAATGCACAGCCCCAAGGACAAGCTGGGCACCCCAAGAGGAGGACAGCCAAAGCCCCATCACCACCG TGCCACCGTGAGCCCCCTGCCTGCCTAGCTTCTACGCCGACAGAGAAGAGGTGAACCTGCAGAGCGCCGT GTGA GCGGCCGC
myc-Err1	GCTAGC GCCACC ATG GAGCAGAAGCTGATCTCCGAGGAGGACCTGGGTGGTGGTAGCGAATTCGGTGGTG GTGGATCC AGCAGCAGAGACAGAAGACCTGACCTGTGTATCAAGGCCGAGCCCGGCACCCCGAGAGCAT TGGCCGGAGAAGCCCTAGCGGCAGCTCCGATTCTCCGGCCACGGGCCGACCTGCAAGGCAACAGATGC TGCAGAGGCGGCAGCAGGAGCAAGACGACACCCCGCAGAGGCAAGTACGTGCTGAACAGCATCCCCA ACGAGACTGTGCCTGGTGTGCGCGACGTGGCCTCCGGCTACCACTACGGCGTGGCCTCCTGCGAGGCCTG CAAGGCCTTCTTCAAGAGAACCATCCAAGGCAACATCGAGTACAGCTGCCCCGCTCCAACGAGTGCGAG ATACCAAGAGACGGAGAAAGCCTGCCAAGCCTGCAGATTACCAAGTGCCTGAGAGTGGGCATGCTGA AGGAGGGCGTGAGACTGGACAGAGTGAGAGGCGGCAGACAGAAGTACAAGAGAAGACCCGACGGCGACGT GCTGCAATATAACCCCGGGGGCCAAGTGCAGCACGGCAGCGCGCCGTGGTGAAGAAGCAAGCCCCCATG AACGCCGTGGTGAGCCACCTGCTGGTGGCCGAGCCCGACAAGCTGTTTCGCCATGCCCCGACCCCGCCCTGC CCGACGGCTACCTGAAGAGCATGAGCACCTGTGCGACCTGGCCGACAGAGAGATCGTGATCATTATCAG CTGGGCCAAGAACATCCCCGGCTTCAGCAACCTGAGCCTGAGCGATCAGATGAGCCTGCTGCAGAGCGTG TGGATGGAGGTGCTGCTCCTGGGCGTGGTGTTCAGAAGCCTGCCCTACGAGGACGAGGTGGTGTTCGCCG AGGACTTCGTGCTGGACGAGGAGAGCAGCAGAAGCGCTAGACTGAGCGACCTGTGCAGCTGCATCCTGCA CCTGGTGAGAAAGTACAGAACCCTGAGAGTGGAGAGAGAGGAGTACGTGATGCTGAAGGCCCTGACCCTG ACCAACAGCGACAGCGTGCACATCGAGGACCCCGAGGCCGTGCTGAGACTGAGAGACACCTGCAAGAGG CCCTGAGCGAGTACGAGAGCAACAGACACCCCGAGGAGCCCTGCAGAGACGGCAAGCTGCTCCTGACCCT GCCCCTGCTGAGACAGACCGCCGCGAGAGTGTGCGACACTTCCACGCCCTGAGAGAGGAGGCGACCGTG CCCATGCACAAGCTGTTCTGGAGATGCTGGAGGCCATGATGGACTGA GCGGCCGC
truncated HA-Err1Y	GCTAGC GCCACC ATG TATCCATACGATGTTCCAGATTACGCTGGTGGTGGTAGCGAATTCGGCGGCGGGG ATCC CATGAACGCCGTGGTGAGCTACTGGCTGGTGGTGAAGCTGGACCAGCTGCTGCCCATGAGCAACCC CGTGCTGACCGACGCCTACCTGAACACCATGAACACCCTGAGGGACCTGGCCGACCAGGAGATCGTGAAC ATCATCAGCTGGGCCAACAACATCCCCGGCTTCAGCAGCCTGAGCCAGAGCGACCAGAAGAGCCTGCTGA GGAGCGCCTGGATGGAGGTGTTCTGCTGTGCGTGATCTTCAGGAGCCTGCCCTACGAGGACAAGGTGGT GTTCCGCCGAGGACATCGTGCTGCACAAGGACAGCAGCTGCAGCGCCAGGCTGAGCGACCTGTGCAGCGAC ATCCTGCACCTGATCAGGAAGTACCAGACCCTGTACCTGGAGAGGGAGGAGTACAACATCCTGAAGGGCC TGGCCCTGACCAACAGCGACAGCGTGACATCGAAGACCCCGAGGCCGTGCTGCAGCTGAGGGACAACCT GTGGGAGACCCTGAACGACTACGAGAGCAACCAGCACCCCGAGGAGCCAGGAGGCACGAGAGGCTGCTG CTGACCCTGCCCCCTGCTGAGGCGAGATGGCCGGCCACGTGATGCAGCACTTCATCAGCCAGTGA GCGGCCGC G

Table S6.

Sequences of designed receptors and primers for design of short construct. Red: restriction cloning sites, yellow: kozak sequence and start codon, green: short peptide tags. Sequences were cloned via restriction digest and ligation into pcDNA3.1(-).

Genome Assembly Statistics — <i>Bombina</i> spp.				
	<i>B. bombina</i> (fire-bellied toad)		<i>B. variegata</i> (yellow-bellied toad)	
	aBomBom1.pri GCF_027579735.1	aBomBom1.alt GCA_027579415.1	aBomVar4 paternal.1 GCA_964205295.1	aBomVar4 maternal.1 GCA_964205305.1
Assembly Overview				
Assembly level	Chromosome	Contig	Chromosome	Chromosome
Assembly type	Haploid (principal)	Haploid (alternate)	Haploid (paternal)	Haploid (maternal)
Submitter	Vertebrate Genomes Project	Vertebrate Genomes Project	Wellcome Sanger Institute	Wellcome Sanger Institute
Date submitted	Jan. 9, 2023	Jan. 6, 2023	Jul 29, 2024	Jul 29, 2024
Sequencing tech.	PacBio Sequel II CLR, Bionano optical maps, 10X Genomics Linked reads, Hi-C	PacBio Sequel II CLR	PacBio HIFI, Hi-C, parental short reads	PacBio HIFI, Hi-C, parental short reads
Long read genome coverage	92.75×	92.75×	26×	26×
Size & Composition				
Genome size	10 Gb	1.9 Gb	9.4 Gb	9.1 Gb
Total ungapped length	9.8 Gb	1.9 Gb	9.4 Gb	9.1 Gb
GC content	39%	38.5%	39%	39%
No. chromosomes	12	—	12	12
No. organelles	1 (mitochondrion)	—	1 (mitochondrion)	—
Scaffold Statistics				
No. scaffolds	2,467	13,751	4,756	3,497
Scaffold N50	1.2 Gb	171.9 kb	1.1 Gb	1.1 Gb
Scaffold L50	4	3,040	4	4
Contig Statistics				
No. contigs	9,446	13,751	11,427	10,308
Contig N50	3.8 Mb	171.9 kb	1.7 Mb	1.7 Mb
Contig L50	737	3,040	1,632	1,652
Annotation (RefSeq — aBomBom1.pri only)				
Total genes	31,613	—	—	—
Protein-coding genes	22,468	—	—	—
BUSCO (tetrapoda_odb10)	C:93.8% [S:89.1%, D:4.7%] F:1.6%, M:4.7%	—	—	—

Table S7. Statistical overview of fire- and yellow-bellied toad assemblies.

Supplementary Data

Data S1. (separate file)

B. bombina err1Y region with analyses of male-specific ddRAD markers and long reads as mapped to the *Bombina bombina* reference genome GCF_027579735.1 (.xlsx).

Data S2. (separate file)

Presence/absence of ddRADsex male-specific markers in the genetic family of *B. bombina*, including mother, father and phenotypically sexed females, males and few specimens with unknown sex (.xlsx).

Data S3. (separate file)

Err1Y-Miniprot gene prediction on original *B. variegata* contig CAXNTK010000102.1:851852 - 966674, amino acid sequence and *B. variegata* genomic contig CAXNTK010000102.1 from first NCBI version; predicted *err1Y* cds exons are marked in red lowercase letters (page 234 and following). (PDF)