**Current Biology, Volume *31***

## Supplemental Information

# A supernumerary “B-sex” chromosome drives male sex determination in the Pachón cavefish, *Astyanax mexicanus*

**Boudjema Imarazene, Kang Du, Séverine Beille, Elodie Jouanno, Romain Feron, Qiaowei Pan, Jorge Torres-Paz, Céline Lopez-Roques, Adrien Castinel, Lisa Gil, Claire Kuchly, Cécile Donnadieu, Hugues Parrinello, Laurent Journot, Cédric Cabau, Margot Zahm, Christophe Klopp, Tomáš Pavlica, Ahmed Al-Rikabi, Thomas Liehr, Sergey A. Simanovsky, Joerg Bohlen, Alexandr Sember, Julie Perez, Frédéric Veyrunes, Thomas D. Mueller, John H. Postlethwait, Manfred Schartl, Amaury Herpin, Sylvie Rétaux, and Yann Guiguen**

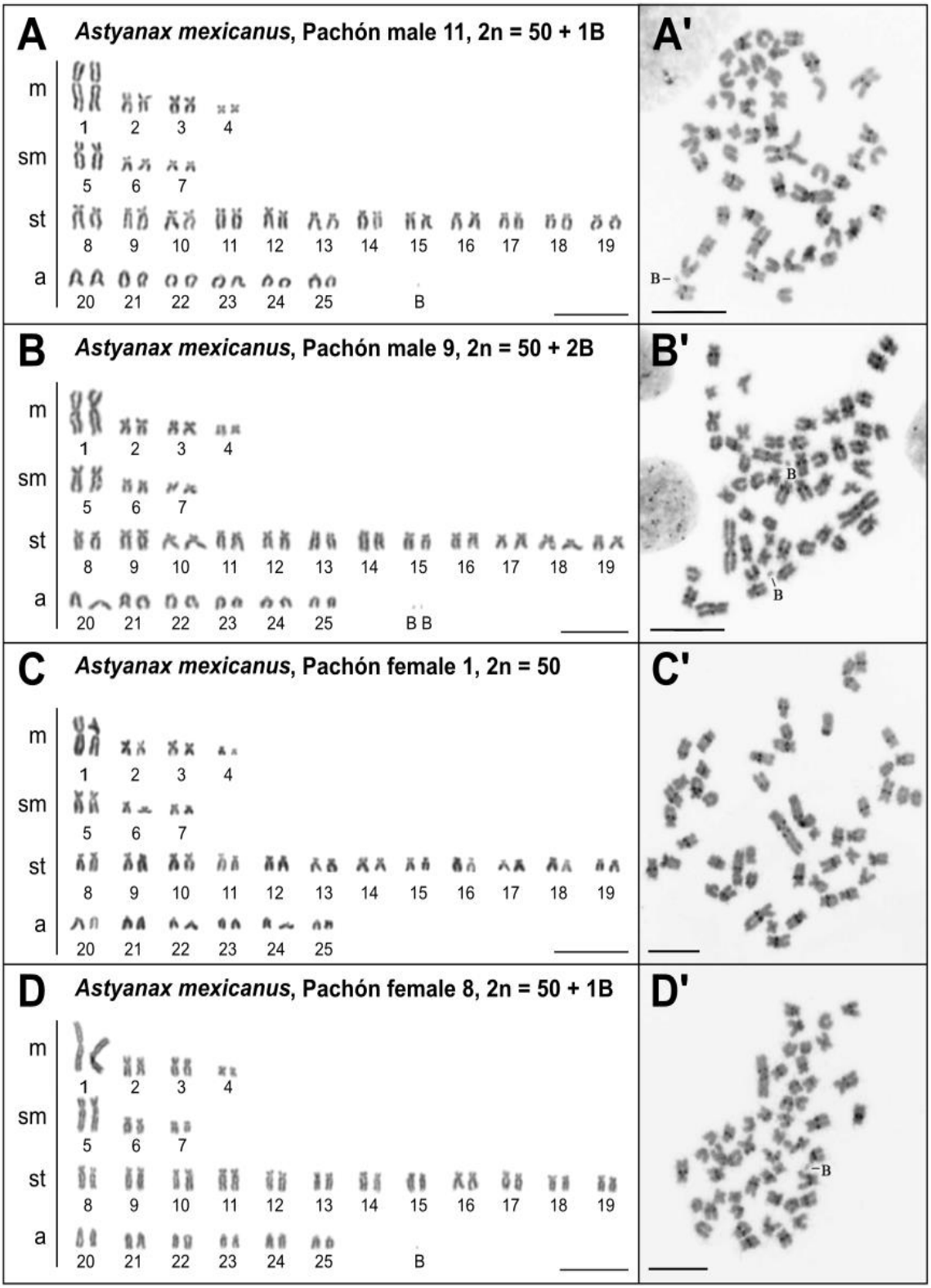


**Figure S1. Karyotypes and corresponding C-banded mitotic metaphases of Pachón cave *Astyanax mexicanus,* with different male and female B chromosome (B) constitution. Related to Figure 1.** Representative male and female Pachón cave karyotypes arranged from Giemsa-stained mitotic chromosomes (panels **A**-**D**) and their corresponding C-banding patterns (panels **A'**-**D'**). B numbers were found to be variable among individuals (from 0 to 3 Bs) with all males having a single (panels **A**-**A'**) or multiple Bs (panels **B**-**B'**) in most of their metaphases, most females having no B (panels **C**-**C'**), and only a few females having rare B positive metaphases (panels **D**-**D'**) Notice also the lack of C-bands on Bs suggesting that these Pachón cavefish male-predominant Bs are largely euchromatic. Scale bar = 10 µm. Male and female numbers referred to Data S1A.

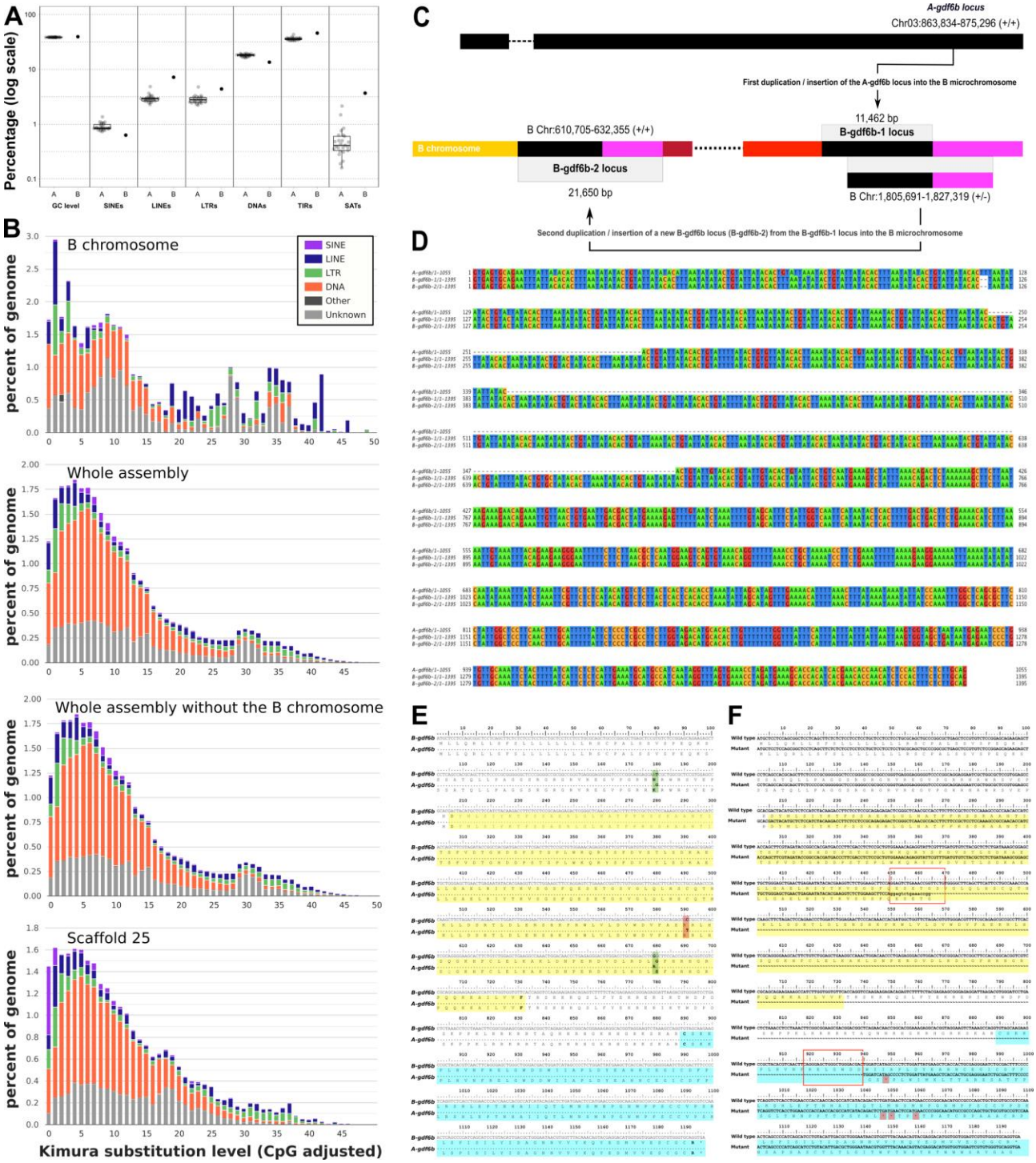


**Figure S2: Genomic repeated elements, *gdf6b* evolution and *gdf6b* sequence alignments. Related to Figure 2. A.** Comparison of the GC content and repeated elements in Pachón cavefish B chromosome (B) versus Pachón cavefish A chromosomes. Percentage (in log scale) of GC content (GC level), short interspersed repeated sequences (SINEs), long interspersed nuclear elements (LINEs), long terminal repeats (LTRs), DNA repeat elements (DNAs), terminal inverted repeat sequences (TIRs), and satellite DNA (SATs) in the 25 A Pachón chromosomes (boxplots, A) versus Pachón B (black dots, B). **B.** Comparison of the repeat landscapes of the Pachón B chromosome (B), the whole assembly with (whole assembly) and without B and of Scaffold_25. Color code for repeat elements is provided in the top right inset of the figure. B repeat landscape is different from the whole assembly with or without the B chromosome and to the repeat landscape of scaffold_25. The B tends to have a higher content of interspersed repeats, mainly due to the expansion of long interspersed

nuclear elements (LINEs). Short interspersed repeated sequences (SINEs), long terminal repeats (LTRs), DNA repeat elements (DNAs), and terminal inverted repeat sequences (TIRs). See Data S1D for additional details. **C.** The *B-gdf6b* loci on the Pachón cave *A. mexicanus* B chromosome (B) stemmed from two successive A and B duplications. Schematic representation of the duplication / insertion history of the two *B-gdf6b* loci on the Pachón cave B (HiC_scaffold_28) with a two-steps duplication hypothesis scheme suggesting that the *B-gdf6b-1* locus originated from an initial duplication of a 11.4 kb region surrounding the *A-gdf6b* locus and inserted in the B, followed by a second independent duplication of a 21.6 kb region surrounding the *B-gdf6b-1* locus that was inserted also in the B. This second independent duplication is probably very recent as the two *B-gdf6b* loci are 99.6 % identical in the 21.6 kb region shared by the two genes. Locations of the duplicated regions are given with respect to the 5'-3' orientation of the *gdf6b* cDNA. The colors of the schematic B fragments depict their different A chromosome origin. **D**: Multiple sequence alignment of Pachón cavefish intron 1 of *A-gdf6b* with *B-gdf6b-1* and *B-gdf6b-2*. Sequences were extracted from the whole genome Pachón cavefish assembly with the following coordinates: For *A-gdf6b* = HiC_scaffold_3:864791:865845:-, For *B-gdf6b-1* = HiC_scaffold_28:617028:618422:+ and for *B-gdf6b-2* = HiC_scaffold_28:1832027:1833421:-. The percentage of identity between the *A*- and *B-gdf6b* is respectively 1.7 % for intron 1 (18 differences in 1054 bp after gaps and indels collapsed to 1 bp) compared to 0.58 % for their proximal promoter (7 differences in 1200 bp after gaps and indels collapsed to 1 bp; proximal promoter defined from the ATG to the 3'end of the adjacent gene, ~ 1400 bp). This could suggest different evolutionary constraints between these two different *gdf6b* regions. **E.** Coding sequence and protein alignments of *A-gdf6b* and *B-gdf6b*. Sequences were aligned with CLUSTALW[S1]. As the *gdf6b-B1* and *gdf6b-B2* are 100% identical from their ATG to STOP codons the alignment only shows differences between *B-gdf6b* (*B-gdf6b-1* and *B-gdf6b-2*) and *A-gdf6b* coding sequences (CDS). The three nucleotide variations between the *B-gdf6b* and *A-gdf6b* CDS are boxed (green for nonsynonymous sites and red for synonymous sites) at positions 180 bp, 591 bp and 679 bp positions of the CDS. Regions highlighted in yellow and blue indicated respectively the Pfam "TGF-β propeptide" and "TGF-β-like" domains[S2]. ".": Identical nucleotides; "*": stop codon. **F.** Alignment of nucleotide (*B-gdf6b*) and translated protein sequences (B-Gdf6b) in wild type and one representative example of a mutant male showing a 470 bp deletion in its *B-gdf6b* gene in F0 fish in the exon 2 encoding for the Pfam "TGF-β propeptide" (yellow) and the "TGF-β-like" (blue) domains[S2]. The resulting fish displayed a frame-shifted and truncated Gdf6b-B protein with premature stop codon (red). The positions of the guide RNAs selected to inactivate *gdf6b* in Pachón cavefish are boxed in red. Similar 470 bp deletions have been also within the *A-gdf6b* gene in other fish.

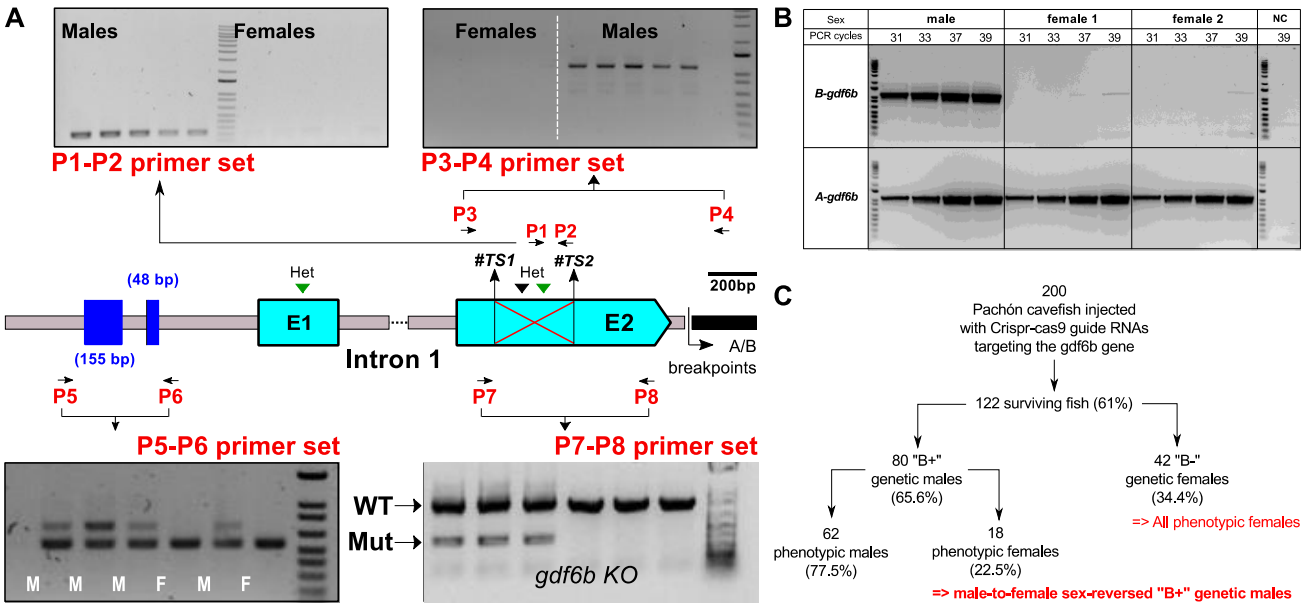


**Figure S3: PCR genotyping of the Pachón cavefish male-predominant B chromosome and the *gdf6b* knockout mutant individuals. Related to Figure 4 and STAR Methods. A.** Absence / presence of the Pachón cavefish B chromosome was detected with different primer pairs based on differences between the *A-gdf6b* and *B-gdf6b loci*. Three sets of primers (P) were designed with two primer sets designed to amplify specifically the *B-gdf6b* copy based either on a single base variation between the A/B *gdf6b* CDS at position 679 bp (P1-P2 primer set), or based on primers located on both sides of the A/B breakpoints downstream of the *B-gdf6b* gene (P3-P4 primer set). The third set of primers (P5-P6 primer set) was designed on gaps/indels variations between *A-gdf6b* and *B-gdf6b* in the proximal promoter of *gdf6b* genes. Primer sets P1-P2 and P3-P4 produce a single PCR fragment only in B+ individuals, and primer set P5-P6 amplifies two bands in B+ individuals and only a single band in B- individuals. Gene knockout (KO) was performed by genome editing using the CRISPR-cas9 method with two guide RNAs target sites (#TS1 and #TS2) designed in order to target *gdf6b* exon 2 (E2). KO individuals were genotyped using primers (P7-P8 primer set) flanking the 2 target sites that induced a 470 bp deletion in *gdf6b* mutant individuals (Mut) compared to the wild type (WT) *gdf6b* sequence. **B.** Increasing PCR cycle numbers allows the detection of a faint PCR fragment specific to the B chromosome in Pachón cavefish females. PCRs were carried out in one male and two female Pachón cavefish with the protocol described for the P3-P4 primer set and with increasing PCR cycle numbers. The B chromosome was detected using primers specifically designed to amplify the B-gdf6b loci (primer set P3-P4) and a control was incorporated with primers specifically designed to amplify the *A-gdf6b* locus (primer set P9-P10). NC: PCR negative control. **C.** Numbers of *gdf6b* knockout generated by CRISPR-Cas9 method including the number of fish injected, and the number and percentage of sex-reversed males obtained (in bold red type). Out of the 200 micro-injected eggs (at 1 cell stage), 122 adult fishes were obtained including 80 genetic males and 42 genetic females. Among the 80 genetic males, we found 62 phenotypic males (77.5%), and 18 phenotypic females (22.5%) displaying a 470 bp deletion on the exon 2 of their *A-gdf6b* and/or *B-gdf6b* gene.

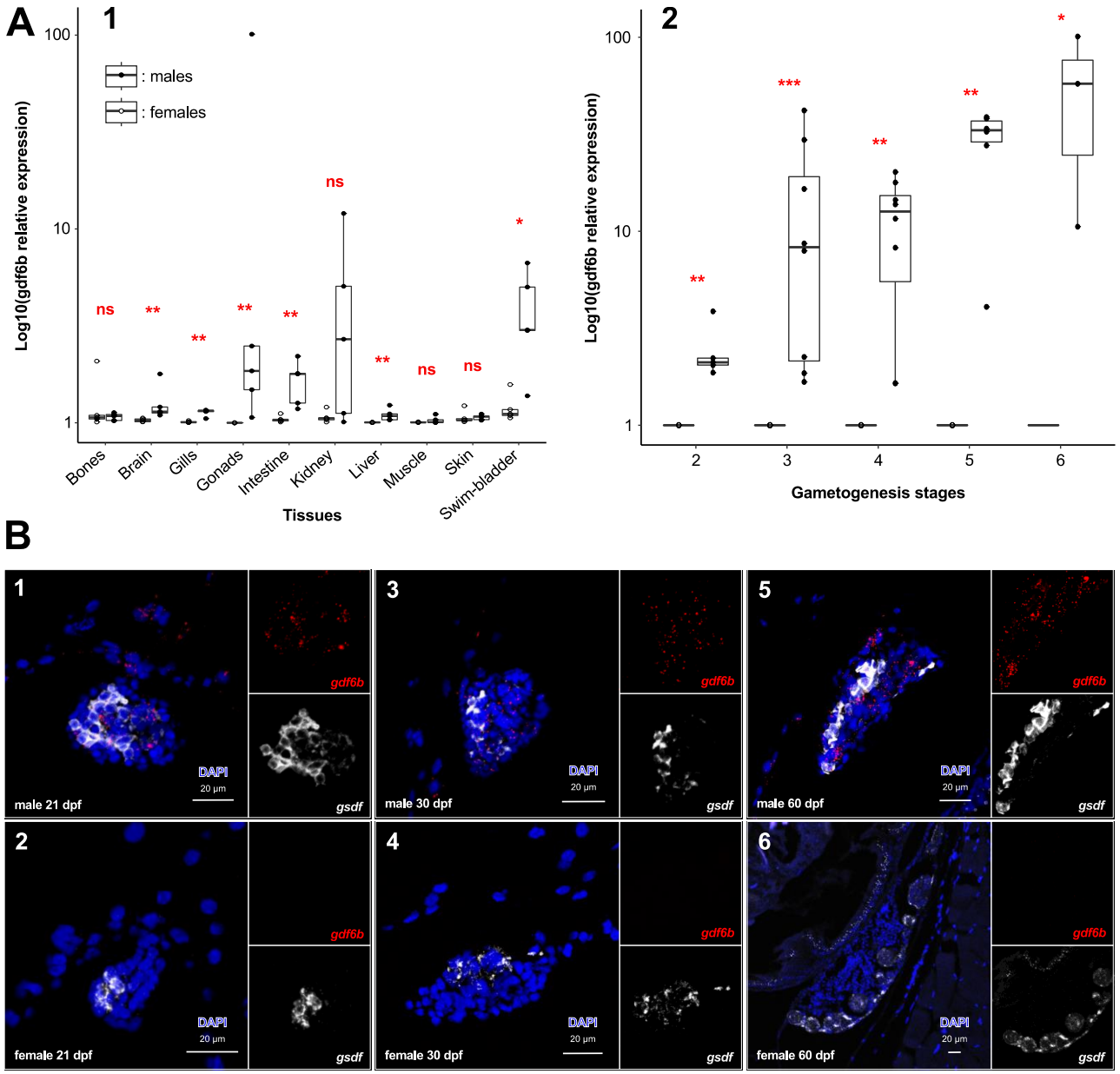


**Figure S4. Expression patterns of *gdf6b* in adult Pachón cave *Astyanax mexicanus*. Related to Figure 4. A.** Expression profiles of *gdf6b* (*A-gdf6b* and *B-gdf6b*) in different adult tissues (**A1**) in males (light grey) and females (dark grey) and during male (light grey) and female (dark grey) gametogenesis (**A2**). Gametogenesis stages 2 to 6 were defined as previously described[S3]. Results are presented as boxplots with individual expression values displayed as dots, the expression median as a line, and the box displaying the first and third quartiles of expression. Statistical significance between males and females were tested with the Wilcoxon Rank Sum Test (Wilcoxon-Mann-Whitney Test) and significant differences are *** = $P < 0.01$; ** = $P < 0.01$; * = $P < 0.05$. **B.** Gonadal expression of *gdf6b* (in red) and the Sertoli and granulosa supporting cell marker *gsdf* (in white) in male (B1, B3 and B5) and female (B2, B4 and B6) differentiating gonads. At all stages i.e., 21, 30- and 60-days post-fertilization (dpf) *gdf6b* is specifically expressed in male gonads with no strict colocalization with *gsdf* in males. Nuclei were stained with DAPI (in blue). Scale bar = 20 µm.

## SUPPLEMENTAL REFERENCES


S1. Higgins, D.G., and Sharp, P.M. (1988). CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. Gene *73*, 237–244.

S2. El-Gebali, S., Mistry, J., Bateman, A., Eddy, S.R., Luciani, A., Potter, S.C., Qureshi, M., Richardson, L.J., Salazar, G.A., Smart, A., et al. (2019). The Pfam protein families database in 2019. Nucleic Acids Res *47*, D427–D432.

S3. Imarazene, B., Beille, S., Jouanno, E., Branthonne, A., Thermes, V., Thomas, M., Herpin, A., Rétaux, S., and Guiguen, Y. (2021). Primordial Germ Cell Migration and Histological and Molecular Characterization of Gonadal Differentiation in Pachón Cavefish Astyanax mexicanus. Sex Dev, 1–18.