nature genetics

Matters arising

https://doi.org/10.1038/s41588-024-01784-w

Phylogenomic and comparative genomic analyses support a single evolutionary origin of flatfish asymmetry

Received: 26 July 2022

Accepted: 25 March 2024

Published online: 27 May 2024

Check for updates

Emanuell Duarte-Ribeiro[1,2] ✉, Ulises Rosas-Puchuri[1], Matt Friedman[3], Gavin C. Woodruff[1], Lily C. Hughes[4], Kent E. Carpenter[5], William T. White[6], John J. Pogonoski[6], Mark Westneat[7], Juan Martin Diaz de Astarloa[8], Jeffrey T. Williams[9], Mudjekeewis D. Santos[10], Omar Domínguez-Domínguez[11], Guillermo Ortí[12], Dahiana Arcila[1,13] & Ricardo Betancur-R[1,13] ✉

ARISING FROM Lü et al. *Nature Genetics* https://doi.org/10.1038/s41588-021-00836-9 (2020)

The eye migration that characterizes flatfish cranial development provides a unique opportunity to study the molecular mechanisms underlying anatomical asymmetry. In a paper in Nature Genetics, Lü et al.[1] (LEA) reported a comparative genomic assessment of flatfish asymmetry and claimed that the two major lineages (Pleuronectoidei and Psettodoidei) are polyphyletic, each evolving asymmetric bodies convergently from different symmetric ancestors (flatfish polyphyly (FP); Fig. 1b). Here we revisit this finding by analyzing three independent genome-scale datasets, including LEA's, showing that support for FP results from a failure to accommodate lineage-specific variation in base composition. We also re-analyzed LEA's genomic dataset to identify positively selected genes (PSGs) but using instead an inferred tree that groups flatfishes as monophyletic (FM), implying a single evolutionary origin of the asymmetric body plan (Fig. 1a). These results suggest a key evolutionary role of thyroid hormones (THs) and bone morphogenetic proteins (BMPs), bridging evidence from the fossil record[2] and single-species molecular assays[3].

Evolutionary nonhomogeneous processes leading to base compositional nonstationarity (BCNS) are a major source of phylogenetic error[4,5]. Unaccounted BCNS among carangarian lineages, a large clade that comprises flatfishes and their closest symmetrical relatives, has been shown to mislead phylogenetic inference, challenging the FM hypothesis[4]. We assessed the effects of BCNS in flatfishes by analyzing the following three genome-scale datasets: the LEA dataset (18 species and 1,693 exon markers), a noncoding ultraconserved elements (UCE) dataset[6] (45 species and 596 markers) and a newly generated exonic dataset[7] (up to 389 species and 990 markers; Supplementary Note 1). We used the following two models of nucleotide substitution: a standard homogeneous model (HM) and a nonhomogeneous model (NHM[8]) that permits rate variation among lineages (Supplementary Information). Contrary to LEA, our results support the single-origin FM hypothesis (Fig. 1c and Extended Data Fig. 1). While both HM and NHM analyses using noncoding UCE data consistently resolve the FM tree (see also ref. 6), many HM analyses based on protein-coding sequences favor the FP topology. These results align with recent findings[5], suggesting that the effects of BCNS are more severe when using protein-coding markers than with introns or UCEs. Together, our analyses indicate that appropriate modeling of lineage-specific variations in protein-coding data is sufficient to favor the FM topology, reconciling the results obtained with coding and noncoding datasets.

LEA discussed additional morphological and molecular evidence to support their assertion that the flatfish cranial asymmetry has two

[1]School of Biological Sciences, University of Oklahoma, Norman, OK, USA. [2]Zoological Institute, Department of Environmental Sciences, University of Basel, Basel, Switzerland. [3]Museum of Paleontology, Department of Earth and Environmental Sciences, University of Michigan, Ann Arbor, MI, USA. [4]North Carolina State University, North Carolina Museum of Natural Sciences, Raleigh, NC, USA. [5]Department of Biological Sciences, Old Dominion University, Norfolk, VA, USA. [6]CSIRO Australian National Fish Collection, National Research Collections Australia, Hobart, Tasmania, Australia. [7]Department of Organismal Biology and Anatomy, University of Chicago, Chicago, IL, USA. [8]Instituto de Investigaciones Marinas y Costeras, Universidad Nacional de Mar del Plata, Mar del Plata, Argentina. [9]Department of Vertebrate Zoology, National Museum of Natural History, Smithsonian Institution, Washington, DC, USA. [10]National Fisheries Research and Development Institute, Department of Agriculture, Quezon City, Philippines. [11]Facultad de Biología, Universidad Michoacana de San Nicolás de Hidalgo, Morelia, México. [12]Department of Biological Sciences, The George Washington University, Washington, DC, USA. [13]Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, USA. ✉e-mail: emanuell.duarteribeiro@unibas.ch; rbetancur@ucsd.edu

https://doi.org/10.1038/s41588-024-01784-w

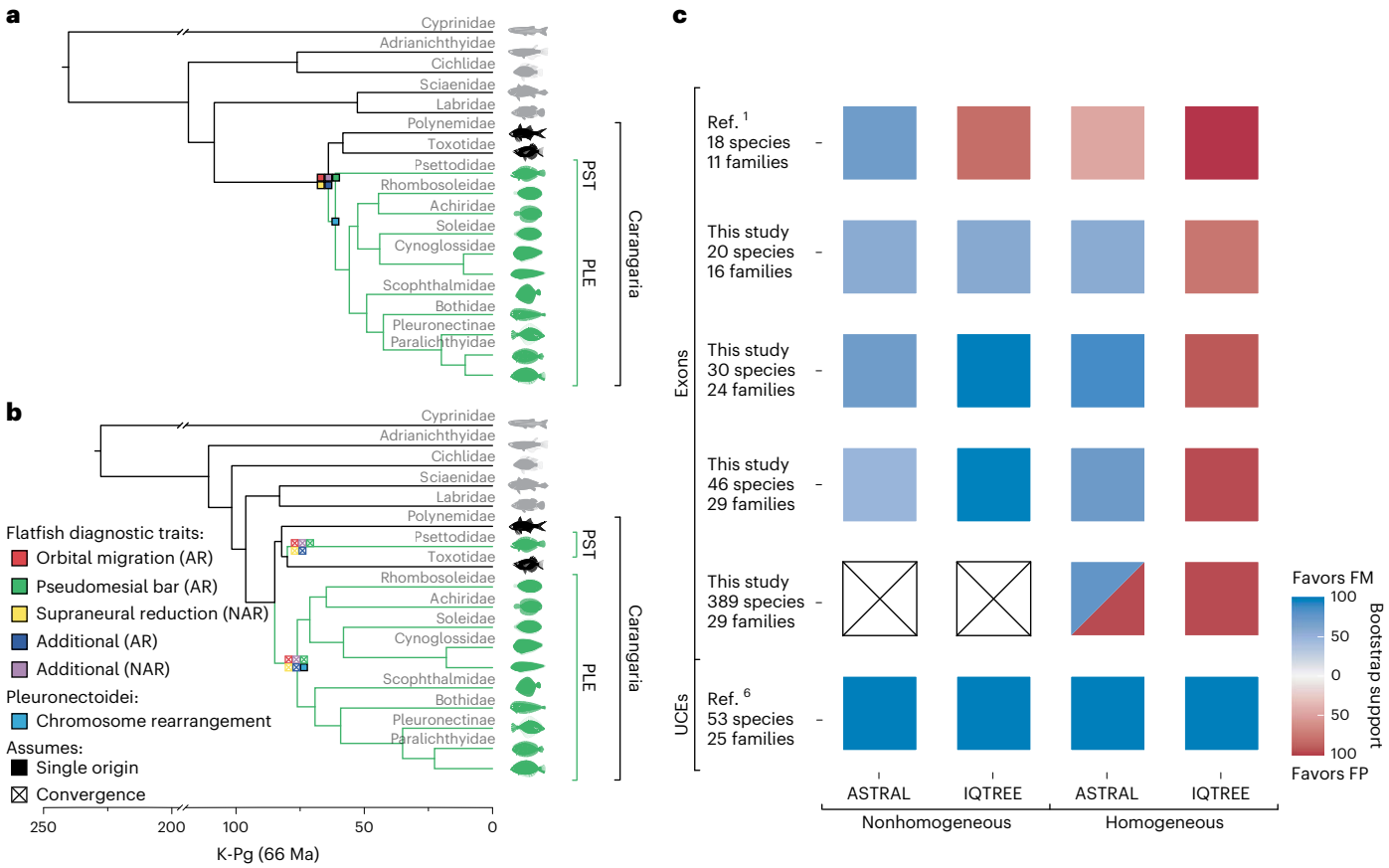


Fig. 1 | Phylogenetic relationships, divergence times and support values for the two competing hypotheses. a, FM tree estimated using LEA's dataset with ASTRAL under an NHM (GHOST) of nucleotide substitution (see Supplementary Note 2 for details on time calibration). **b**, FP tree illustrates the phylogenetic hypothesis and divergence times proposed by LEA. **a**,**b**, Branch and fish silhouette colors denote major flatfish (PLE and PST, green) and nonflatfish carangarian (black) clades. Color-coded squares at branches denote diagnostic characters defining flatfishes. Some diagnostic characters represent morphological features already identifiable in the earliest flatfish fossils, including AR (orbital migration) and NAR (supraneural reduction)[2] traits. Further diagnostic characters shared between extant PLE and PST also include ARs (recessus orbitalis, pseudomesial bar and asymmetrical pigmentation) and NARs (for example, absence of supraneurals, dorsal-fin insertion above skull and depression on the inner face of saccular otolith)[6]. Full squares represent traits evolving under the single-origin assumption, and crossed squares represent traits evolving under the assumption of convergent evolution. **c**, Meta-table shows bootstrap support for the FM and FP hypotheses using different genomic data types, datasets, phylogenetic reconstruction methods and nucleotide substitution models (HM and NHM). Each row represents a different dataset or subset, and each column represents a different phylogenetic method and substitution model. Blue squares denote support for FM, and shades show bootstrap values (0–100%); red squares indicate support for FP; white squares denote lacking analyses (statistically intractable; Supplementary Information). Coalescent analyses were run in ASTRAL using ML gene trees estimated with IQ-TREE (both HM and NHM) based on either the complete gene-tree topology or a less resolved topology after collapsing clades with low bootstrap support (<20%). Note that ASTRAL analyses based on our full exonic dataset produced contrasting results (analyses using collapsed gene trees support FM). Concatenation ML analyses were conducted in IQ-TREE (both HM and NHM). PST, Psettodoidei; PLE, Pleuronectoidei; AR, asymmetry-related; NAR, nonasymmetry-related, ML, maximum-likelihood.

independent origins. We argue that this is subject to interpretation and is inconsistent with other evidence from paleontology and comparative anatomy (Fig. 1b). The earliest definitive flatfish fossils are from the early Eocene of Bolca, Italy (ca. 48.5 Ma) and include the crown pleuronectoid †*Eobothus minimus* plus taxa interpreted as stem pleuronectiforms—†*Amphistium paradoxum*, †*Heteronectes chaneti* and, less certainly, †*Anorevus lorenzoni*[2,9]. In particular, †*Amphistium* and †*Heteronectes* show strong but incomplete orbital migration and are thus considered the link between flatfishes and their symmetrical relatives. Slightly older fossils (ca. 56 Ma) showing incomplete cranial asymmetry are reported in the literature but remain undescribed[9]. Morphologically, multiple anatomical synapomorphies support the FM hypothesis (for example, pseudomesial bar, orbital migration, recessus orbitalis and asymmetrical pigmentation)[6], including traits that are unrelated to asymmetry (for example, axial and fin skeletons and otoliths; Fig. 1a)[6] and therefore are less likely to have evolved via convergent evolution (see also ref. 10).

LEA identified a shared chromosome re-arrangement between Psettodoidei and two carangarian species that are also absent in Pleuronectoidei. They interpreted this as supporting the polyphyletic origin of flatfishes. However, without comparisons involving noncarangarian outgroups, we contend that their interpretation is susceptible to polarization issues. An equally plausible explanation is that this re-arrangement evolved within Pleuronectoidei after their split with Psettodoidei. Finally, LEA reported analyses based on relative evolutionary rates and lineage-specific substitutions to support FP and explain the lesser degree of asymmetry found in *Psettodes*. We assert that pleuronectoid species exhibit rates overlapping with *Psettodes*, carangarians and outgroup species. Moreover, the inclusion of key species into the alignments reveals that many substitutions claimed to be pleuronectoid-specific are shared with at least one carangarian species. Consequently, these arguments should not be construed as evidence against FM (see details in Supplementary Note 4 and Extended Data Fig. 3).

https://doi.org/10.1038/s41588-024-01784-w

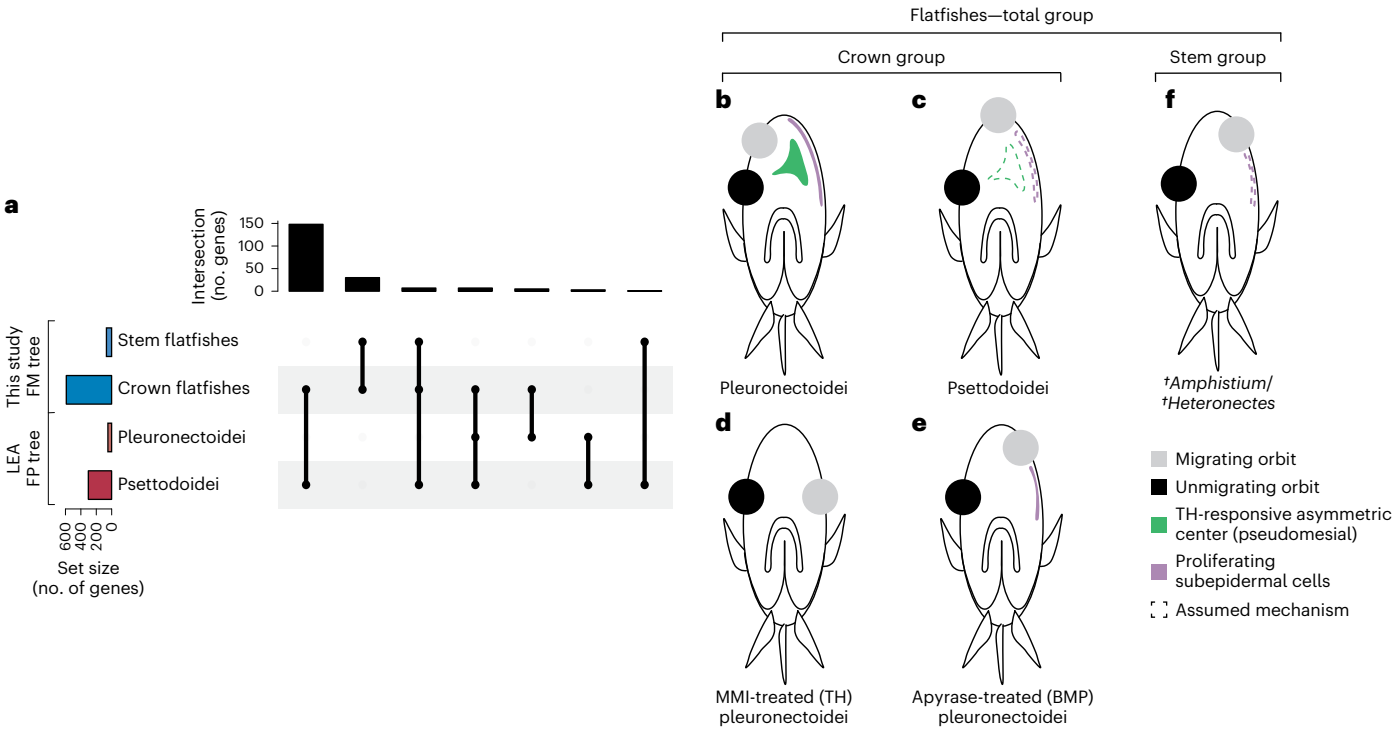


Fig. 2 | Hypothesized molecular mechanisms underlying flatfish asymmetric development. a, Upset plot comparing positively selected genes (PSGs) identified with aBSREL for the stem and crown (foreground) flatfishes using the FM tree (horizontal blue bars) and the PSGs reported by LEA for Pleuronectoidei and Psettodoidei using the FP tree (horizontal red bars). Black dots connected by black lines represent all existent dataset intersection combinations; vertical bars indicate the number of PSGs within a particular intersection. The substantial overlap (intersection size) of PSGs identified in the two major flatfish lineages and across different analyses suggests that Pleuronectoidei and Psettodoidei share the same molecular mechanisms of asymmetric development, ultimately corroborating the single-origin hypothesis. **b**–**f**, The suggested developmental model summarizes the hypothesized evolutionary mechanisms underlying flatfish asymmetric development[2,3]. **b**, Simplified representation of a Pleuronectoidei postmetamorphic juvenile highlighting the two major developmental events responsible for the eye migration—subocular dermal cell proliferation (purple) and the TH-responsive asymmetric center (green) that later develops into the pseudomesial bar[3] (flatfish autapomorphy that frames the migrated orbit[2]). **c**, In postmetamorphic Psettodoidei, the position of the subocular dermal cell proliferation (dashed purple line) and the TH-responsive asymmetric center (dashed green line) are inferred based on the presence of the pseudomesial bar and the complete eye migration in the adult, where migrating orbit eclipses the body mid-line[2]. **d**, Eye migration is entirely inhibited in the MMI-treated larvae due to the disruption of the thyroid axis[3]. **e**,**f**, Apyrase-treated larvae (**e**) have incomplete orbit migration caused by inhibition of the TH-responsive asymmetric center ossification[3], closely resembling the primitive condition found in the earliest flatfish fossils that lack the pseudomesial bar (**f**)[2]. Note that TFF-1 (a transcription factor that regulates the TH signaling pathway) and *bmp4* (involved in the heterotypic ossification of the TH-responsive asymmetric center and the origin of the pseudomesial bar) were identified as PSGs in our analyses using the FM topology, but not in LEA's study using the FP topology. Fish illustrations are adapted from refs. 2,3.

To investigate the presence of lineage-specific adaptations under the FM topology (Fig. 1a), we tested for positive selection among the 1,693 orthologs identified by LEA using the aBSREL model of codon evolution[11]. We used the following two different foreground schemes (Supplementary Note 3 and Extended Data Fig. 2): (1) stem flatfishes, which aimed to detect genes responsible for the initial break of symmetry in the single branch leading to all extant flatfish species, and (2) crown flatfishes, which aimed to detect genes responsible for further adaptations experienced later in the flatfish radiation. We identified 67 PSGs in the stem flatfish lineage, and 588 PSGs shared between Psettodoidei and at least one pleuronectoid lineage (Fig. 2a). Remarkably, 15 (31%) and 162 (53%) of the lineage-specific PSGs that LEA reported using their FP topology for Pleuronectoidei and Psettodoidei, respectively, are under positive selection in the crown flatfish lineage when analyzed using the FM tree (Fig. 2a). Although a scenario involving parallelism cannot be disregarded, such substantial overlap in PSGs provides additional corroboration for the single-origin hypothesis.

Gene Ontology characterization of PSGs revealed genes involved in the determination of left/right symmetry during embryonic development (*bmp4, mkks and flr*) and components of several essential signaling pathways responsible for tissue development and cell proliferation in stem flatfishes—BMP (*bmp4*), WNT (*cxxc4*, *dcdc2b*, *carf*, *mkks* and *tinagl1*), RA (*asxl1*), FGF (*mil*), NOTCH (*dlk1*, *dlk2*) and HOX (*shox*). Several of these are regulated by THs, which comprise a key signaling pathway that orchestrates vertebrate metamorphosis[12]. In flatfishes, the disruption of the thyroid axis with methimazole (MMI) inhibits asymmetric skull development by (1) reducing dermal cell proliferation in the subocular region and (2) ablating a TH-responsive asymmetric center, localized just ventral to the migrating eye and determined by deiodinase 2 (*dio2*) expression (Fig. 2b–f)[3]. While the mechanisms by which THs regulate *dio2* asymmetric expression in teleosts remain elusive, in humans it is well known that *dio2* expression is stimulated by the thyroid transcription factor-1 (TTF-1)[13]. TTF-1 is required for the preoptic region development in zebrafish[14] and is differentially expressed in the metamorphosing head of the Atlantic halibut[15]. Notably, TFF-1 was also identified as a PSG in stem flatfishes when using the FM topology, a result not reported by LEA using their FP tree.

Another important developmental aspect involving the TH-responsive asymmetric center is that its ossification correlates with the origin of the pseudomesial bar. This bone structure has a central role in driving the complete eye migration, which is unique to crown flatfishes[3]. Apyrase treatment hinders eye migration in the metamorphosing Senegalese sole by inhibiting the heterotypic dermal ossification regulated by BMP[3]. Remarkably, apyrase-treated larvae

develop other indexes of metamorphic progressions, such as subocular dermal cell proliferation, closely resembling the primitive condition found in the earliest flatfish fossils that lack the pseudomesial bar (Fig. 2f)[2,3]. The accumulation of amino acid substitutions at the *bmp4* prodomain affects its regulation and has been identified as the major driver of craniofacial adaptations in African cichlids and Galapagos finches[16]. Contrary to the LEA findings, *bmp4* is positively selected in stem flatfishes, suggesting that its regulation has a major role in the origin of the flatfish asymmetry.

Our findings reveal that improperly modeled genome-scale datasets are prone to phylogenetic biases, artificially supporting the dual-origin asymmetry hypothesis. Using the FM topology as the evolutionary framework, our results bridge evidence from the fossil record[2] and single-species molecular assays[3], suggesting that the molecular basis of the complex phenotypic modifications in flatfishes may involve an intricate interaction between THs and many important developmental pathways. These results ultimately underscore the importance of accounting for topological uncertainty and model violations in phylogenetically informed comparative genomic analyses.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41588-024-01784-w.

References

1. Lü, Z. et al. Large-scale sequencing of flatfish genomes provides insights into the polyphyletic origin of their specialized body plan. *Nat. Genet.* **53**, 742–751 (2021).
2. Friedman, M. The evolutionary origin of flatfish asymmetry. *Nature* **454**, 209–212 (2008).
3. Campinho, M. A. et al. A thyroid hormone regulated asymmetric responsive centre is correlated with eye migration during flatfish metamorphosis. *Sci Rep.* **8**, 12267 (2018).
4. Betancur-R, R. et al. Addressing gene tree discordance and non-stationarity to resolve a multi-locus phylogeny of the flatfishes (Teleostei: *Pleuronectiformes*). *Syst. Biol.* **62**, 763–785 (2013).
5. Jarvis, E. D. et al. A phylogeny of modern birds. *Science* **346**, 1126–1138 (2014).
6. Harrington, R. C. et al. Phylogenomic analysis of carangimorph fishes reveals flatfish asymmetry arose in a blink of the evolutionary eye. *BMC Evol. Biol.* **16**, 224 (2016).
7. Duarte-Ribeiro, E. et al. Phylogenomic and comparative genomic analyses support a single evolutionary origin of flatfish asymmetry. *figshare* https://doi.org/10.6084/m9.figshare.25808329.v1 (2024).
8. Crotty, S. M. et al. GHOST: recovering historical signal from heterotachously evolved sequence alignments. *Syst. Biol.* **69**, 249–264 (2020).
9. Bannikov, A. F. & Zorzin, R. A new genus and species of percomorph fish ("stem pleuronectiform") from the Eocene of Bolca in northern Italy. *Miscellanea Paleontologica* **17**, 5–14 (2020).
10. Chanet, B., Mondéjar-Fernández, J. & Lecointre, G. Flatfishes interrelationships revisited based on anatomical characters. *Cybium* **44**, 9–18 (2020).
11. Smith, M. D. et al. Less is more: an adaptive branch-site random effects model for efficient detection of episodic diversifying selection. *Mol. Biol. Evol.* **32**, 1342–1353 (2015).
12. Mourouzis, I., Lavecchia, A. M. & Xinaris, C. Thyroid hormone signalling: from the dawn of life to the bedside. *J. Mol. Evol.* **88**, 88–103 (2020).
13. Gereben, B., Salvatore, D., Harney, J. W., Tu, H. M. & Larsen, P. R. The human, but not rat, *dio2* gene is stimulated by thyroid transcription factor-1 (TTF-1). *Mol. Endocrinol.* **15**, 112–124 (2001).
14. Manoli, M. & Driever, W. *Nkx2.1* and *Nkx2.4* genes function partially redundant during development of the zebrafish hypothalamus, preoptic region, and pallidum. *Front. Neuroanat.* **8**, 145 (2014).
15. Alves, R. N. et al. The transcriptome of metamorphosing flatfish. *BMC Genomics* **17**, 413 (2016).
16. Parsons, K. J. & Albertson, R. C. Roles for Bmp4 and CaM1 in shaping the jaw: evo-devo and beyond. *Annu. Rev. Genet.* **43**, 369–388 (2009).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All newly sequenced data are available from NCBI under the BioProject accession number PRJNA862198. Detailed taxonomic sampling information and phylogenetic datasets are available in the Figshare digital repository[7] (see also Supplementary Information). Source data are provided with this paper.

Code availability

We conducted analyses using single commands from publicly available software. All settings are fully reported in the Supplementary Information.

Acknowledgements

The computational work for this project was conducted at the University of Oklahoma Supercomputing Center for Education & Research (OSCER). Funding for the project was secured through grants from the National Science Foundation (NSF; DEB-1932759 and DEB-2225130 to R.B., DEB-2015404 and DEB-2144325 to D.A. and DEB-1541554 to G.O.). Additional financial support was provided by the Office of the Vice President for Research and Partnerships, as well as the Office of the Provost at the University of Oklahoma.

Author contributions

E.D.R. and R.B.-R. conceived and supervised the project. E.D.R., R.B.-R., K.E.C., W.T.W., J.J.P., M.W., J.M.D.A., J.T.W., M.D.S., O.D.D., G.O. and D.A. collected samples. E.D.R., R.B.-R., U.R.P., M.F., L.C.H., G.C.W., G.O. and D.A. designed and performed bioinformatic analyses. E.D.R., R.B.-R. and M.F. wrote the manuscript. All authors have read, revised and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

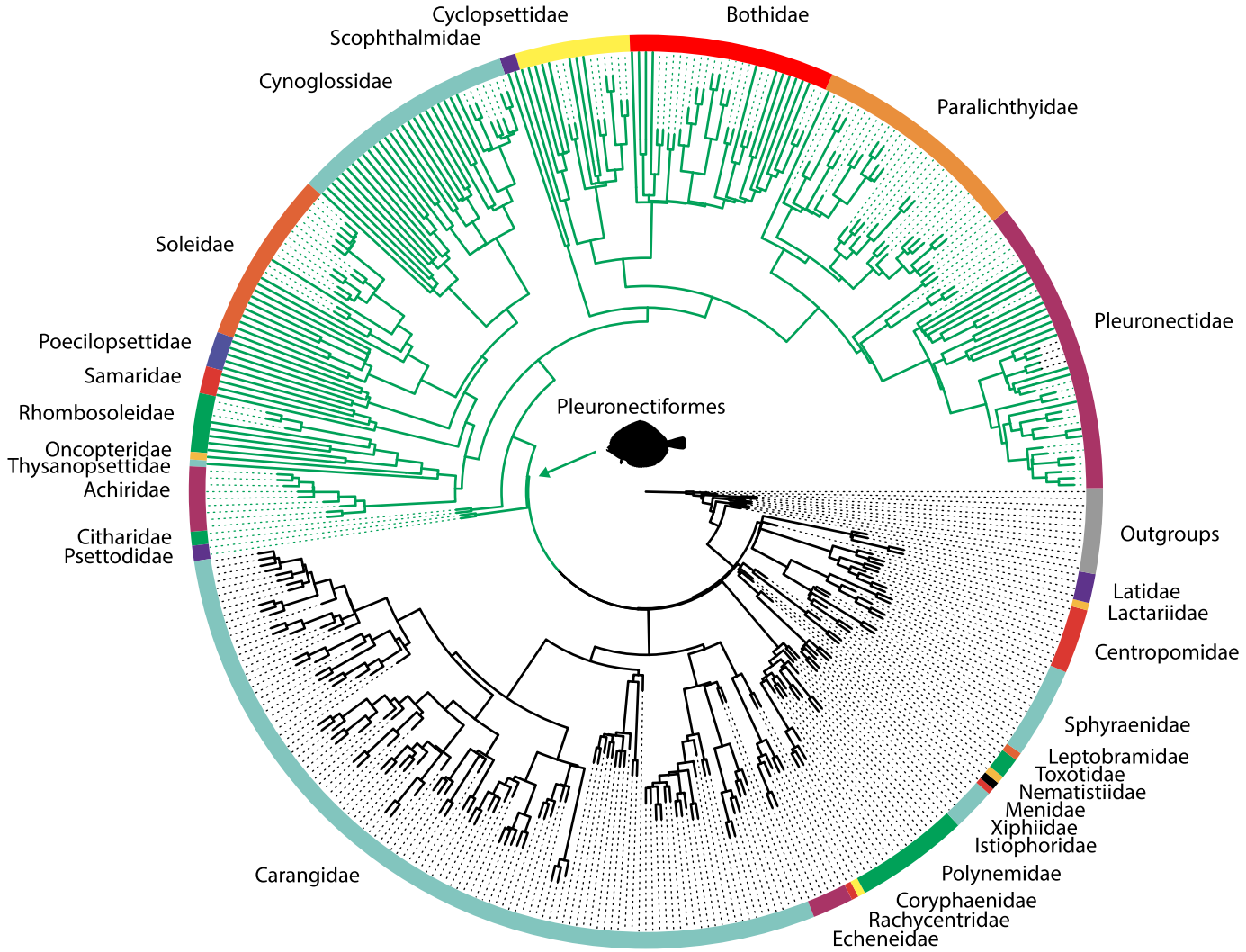
Extended data is available for this paper at https://doi.org/10.1038/s41588-024-01784-w.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41588-024-01784-w.

Correspondence and requests for materials should be addressed to Emanuell Duarte-Ribeiro or Ricardo Betancur-R.

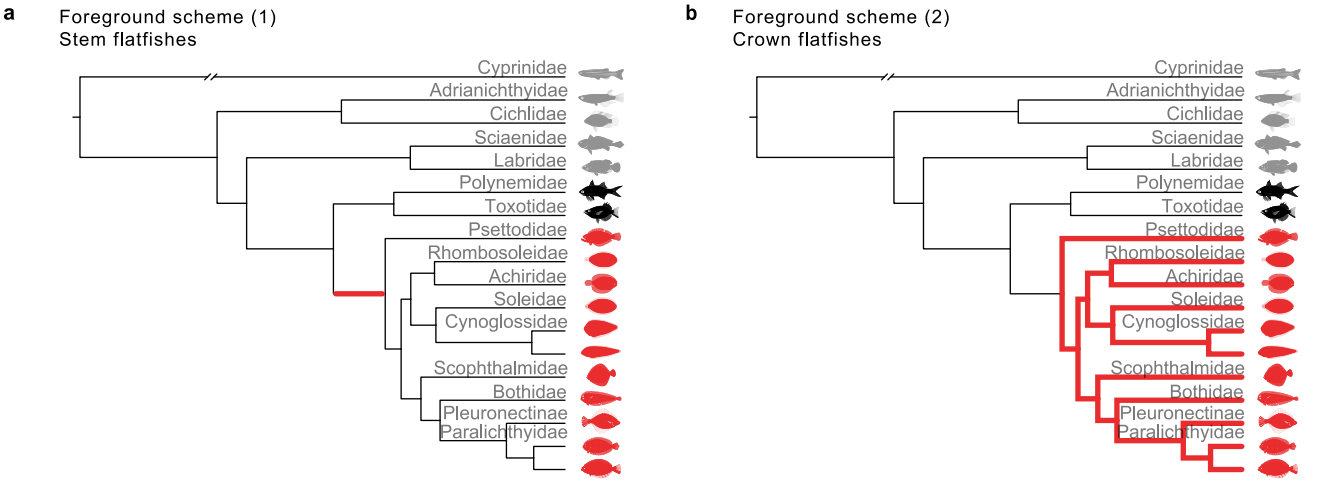
Peer review information *Nature Genetics* thanks Jeramiah Smith and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

https://doi.org/10.1038/s41588-024-01784-w



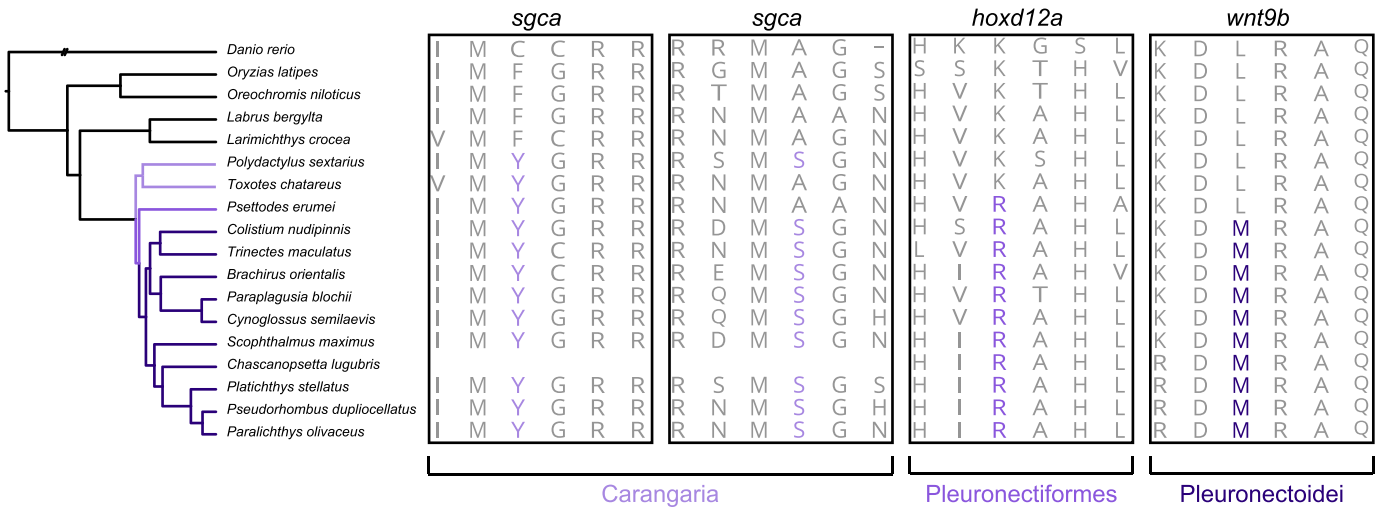
Extended Data Fig. 1 | Species tree for Carangaria estimated with ASTRAL-III. Species tree for Carangaria estimated with ASTRAL-III using gene trees with weakly supported nodes (bootstrap values [BS] <20%) collapsed into polytomies to reduce the effects of gene-tree error. Gene trees were estimated using IQ-TREE (HM; GTR + G) based on newly sequenced data that covers 990 loci and 389 species. Collapsing gene-tree branches with low support helped reduce the negative effects of gene-tree estimation error, successfully resolving flatfish monophyly (FM [BS 76%]; green arrow) using the complete dataset. Tree files with tip labels and support values are available from the Figshare digital repository.

https://doi.org/10.1038/s41588-024-01784-w



Extended Data Fig. 2 | Schematic representation of the two alternative foreground schemes used to identify positively selected genes in flatfishes using the aBSREL model in HyPhy. a, Stem flatfishes, which aimed to detect genes responsible for the initial break of symmetry in the single branch leading to all extant flatfish species. **b**, Crown flatfishes, which aimed to detect genes responsible for further adaptations experienced later in the flatfish radiation.

https://doi.org/10.1038/s41588-024-01784-w



Extended Data Fig. 3 | Amino acid alignments highlighting substitutions (in purple) claimed to be Pleuronectoidei-specific by LEA[1]. By adding sequences from *Psettodes* and other non-flatfish carangarians into the alignments, we show that three out of the four Pleuronectoidei-specific substitutions discussed by the authors are in fact shared with Psettodoidei or other non-flatfish carangarian species. These include two missense substitutions in the musculature development-related gene *sgca*, which the authors suggest is related to the formation of the flat phenotype. Both substitutions are shared with *Polydactylus sextarius*, a symmetrical carangarian species, and, therefore, may not necessarily explain the evolution of the flatfish thin musculature. The *hoxd12a* substitution is shared between Pleuronectoidei and Psettodoidei. This gene is involved in the development of the flatfish dorsal fin, an important locomotion adaptation that improves their maneuverability in the benthic environment, and that may provide additional evidence for the single origin of the asymmetric body plan.

nature portfolio

Corresponding author(s): Ricardo Betancur-R. Emanuell Duarte-Ribeiro

Last updated by author(s): 12/15/2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☒	☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☒	☐ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used during data collection
Data analysis	Trimmomatic (v0.36); Samtools (v1.9); Velvet (v1.2.10); aTRAM (v2.0); TranslatorX; Mafft (v7.421); IQ-TREE (v2.0); ASTRAL-III; PartitionFinder2; treePL; HyPhy; aBSREL

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

We archived the sequence data generated for this manuscript as raw reads in the NCBI Sequence Repository (PRJNA862198). Details of the taxa and sequences removed are available in the Figshare digital repository (https://doi.org/10.6084/m9.figshare.24799245.v1). Phylogenetic reconstruction results for all datasets and

subsets are presented in the Figshare digital repository (https://doi.org/10.6084/m9.figshare.20372565.v2). Results for the stem and crown foreground schemes are presented in Tables S2-3 (see Figshare digital repository - https://doi.org/10.6084/m9.figshare.20505675.v2).

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	n/a
Reporting on race, ethnicity, or other socially relevant groupings	n/a
Population characteristics	n/a
Recruitment	n/a
Ethics oversight	n/a

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	n/a
Data exclusions	n/a
Replication	n/a
Randomization	n/a
Blinding	n/a

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	n/a
Research sample	n/a
Sampling strategy	n/a
Data collection	n/a
Timing	n/a
Data exclusions	n/a
Non-participation	n/a
Randomization	n/a

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	In a recent paper in Nature Genetics, Lü et al. (LEA) conducted the first comparative genomic assessment of flatfish asymmetry, and claimed that the two major lineages (Pleuronectoidei and Psettodoidei) are polyphyletic, each evolving asymmetric bodies convergently from different symmetric ancestors. We revisit this finding by analyzing three independent genome-scale datasets, including LEA's, showing that support for flatfish polyphyly results from a failure to accommodate lineage-specific variation in base composition. We also reanalyzed LEA's genomic dataset to identify positively selected genes but using instead an inferred tree that groups flatfishes as monophyletic, implying a single evolutionary origin of the asymmetric body plan.
Research sample	We assessed flatfish phylogenetic relationships by analyzing three genome-scale datasets: the LEA dataset (18 species and 1693 exon markers), a non-coding ultraconserved elements (UCE) dataset (45 species and 596 markers), and a newly generated exonic dataset (up to 389 species and 990 markers).
Sampling strategy	The newly generated exonic dataset that included 990 exonic markers and covered 389 carangarian species. Flatfish diversity in the new dataset is represented by a sample of 207 out of ca. 600 species, including representatives from all 15 valid families.
Data collection	Sequencing library preparation was performed by Arbor Biosciences using the dual round ('touchdown') capture protocol. Exon capture probes were designed based on alignments of 1,104 single-copy exons for seven species strategically selected to cover Carangaria diversity (i.e., Trinectes inscriptus, Carangoides armatus, Cynoglossus maculipinnis, Remora remora, Microstomus kitt, Rachycentron canadum, Samariscus triocellatus). Legacy markers widely used in fish phylogenetics were also added to the initial probe set: TBR1, MYH6, KIAA1239, PLAGL2, PTCHD1, RIPK4, SH3PX3, SIDKEY, SREB2, ZIC1, SVEP1, GPR61, SLC10A3, UBE3A, and UBE3A-like. A total of 412 samples were sequenced at the University of Chicago (https://fgf.uchicago.edu) using one-lane Illumina HiSeq 4000 with paired-end 100 bp reads.
Timing and spatial scale	As an strategy to increase taxonomic sampling, we assembled the most comprehensive phylogenetic dataset for flatfishes to date (207 species).
Data exclusions	Given the pervasiveness of contamination in phylogenomic datasets, in addition to the raw data quality assessments, we implemented a multistep quality control (QC) pipeline to remove cross-contaminated sequences or misidentified species using an array of sequence- and tree-based approaches. After this final QC step, the final exonic dataset included 990 exonic markers for 389 carangarians, including 207 flatfish species. Details of the taxa and sequences removed are available in the Figshare digital repository.
Reproducibility	We used different datasets and subsets as an way to assess the reproducibility of our results.
Randomization	DNA extraction and Library pooling for Illumina sequencing were specifically randomised to avoid crosscontamination issues. We also used different datasets and subsets as an way to assess the effects of taxonomic sampling.
Blinding	No blinding was applied for data analyses.

Did the study involve field work? ☐ Yes ☒ No

Field work, collection and transport

Field conditions	n/a
Location	n/a
Access & import/export	n/a
Disturbance	n/a

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used: n/a

Validation: n/a

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s): n/a

Authentication: n/a

Mycoplasma contamination: n/a

Commonly misidentified lines (See ICLAC register): n/a

Palaeontology and Archaeology

Specimen provenance: n/a

Specimen deposition: n/a

Dating methods: n/a

☒ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Ethics oversight: n/a

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals: n/a

Wild animals: n/a

Reporting on sex: n/a

Field-collected samples: n/a

Ethics oversight: n/a

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	n/a
Study protocol	n/a
Data collection	n/a
Outcomes	n/a

Dual use research of concern

Policy information about dual use research of concern

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

No	Yes	
☒	☐	Public health
☒	☐	National security
☒	☐	Crops and/or livestock
☒	☐	Ecosystems
☒	☐	Any other significant area

Experiments of concern

Does the work involve any of these experiments of concern:

No	Yes	
☒	☐	Demonstrate how to render a vaccine ineffective
☒	☐	Confer resistance to therapeutically useful antibiotics or antiviral agents
☒	☐	Enhance the virulence of a pathogen or render a nonpathogen virulent
☒	☐	Increase transmissibility of a pathogen
☒	☐	Alter the host range of a pathogen
☒	☐	Enable evasion of diagnostic/detection modalities
☒	☐	Enable the weaponization of a biological agent or toxin
☒	☐	Any other potentially harmful combination of experiments and agents

Plants

Seed stocks	n/a
Novel plant genotypes	n/a
Authentication	n/a

ChIP-seq

Data deposition

☐ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links *May remain private before publication.*	n/a
Files in database submission	n/a
Genome browser session (e.g. UCSC)	n/a

Methodology

Replicates	n/a
Sequencing depth	n/a
Antibodies	n/a
Peak calling parameters	n/a
Data quality	n/a
Software	n/a

Flow Cytometry

Plots

Confirm that:

☐ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☐ All plots are contour plots with outliers or pseudocolor plots.

☐ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	n/a
Instrument	n/a
Software	n/a
Cell population abundance	n/a
Gating strategy	n/a

☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type	n/a
Design specifications	n/a
Behavioral performance measures	n/a

Acquisition

Imaging type(s) n/a

Field strength n/a

Sequence & imaging parameters n/a

Area of acquisition n/a

Diffusion MRI ☐ Used ☒ Not used

Preprocessing

Preprocessing software n/a

Normalization n/a

Normalization template n/a

Noise and artifact removal n/a

Volume censoring n/a

Statistical modeling & inference

Model type and settings n/a

Effect(s) tested n/a

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference n/a

(See Eklund et al. 2016)

Correction n/a

Models & analysis

n/a	Involved in the study
☐	☐ Functional and/or effective connectivity
☐	☐ Graph analysis
☐	☐ Multivariate modeling or predictive analysis

Functional and/or effective connectivity n/a

Graph analysis n/a

Multivariate modeling and predictive analysis n/a