Report

Current Biology

# Late Pleistocene stickleback environmental genomes reveal the chronology of freshwater adaptation

## Highlights

- Sequencing Late Pleistocene three-spined stickleback environmental genomes from sedaDNA
- These genomes have a majority of marine-adapted ancestry at adaptive loci
- Freshwater ancestry of large effect reflects findings from experimental releases
- Sedimentary DNA can be used in retrospective "evolve and resequence" experiments

## Authors

Jan Laine, Sarah S.T. Mak, Nuno F.G. Martins, ..., Mikkel Winther Pedersen, Anders Romundset, Andrew D. Foote

## Correspondence

andrew.foote@ibv.uio.no

## In brief

Laine et al. sequence stickleback genomes from Late Pleistocene sediments of a post-glacial lake, which spans an ecological transition from marine to freshwater habitat. Most genes show marine ancestry; however, freshwater ancestry is found in loci of large effect size as early as the brackish phase, highlighting the chronology and tempo of adaptation.

Laine et al., 2024, Current Biology *34*, 1142–1147
March 11, 2024 
https://doi.org/10.1016/j.cub.2024.01.056

CellPress

CellPress
OPEN ACCESS

Current Biology

Report

# Late Pleistocene stickleback environmental genomes reveal the chronology of freshwater adaptation

Jan Laine,[1,7] Sarah S.T. Mak,[2,7] Nuno F.G. Martins,[2,7] Xihan Chen,[3] M. Thomas P. Gilbert,[1,2] Felicity C. Jones,[4] Mikkel Winther Pedersen,[3] Anders Romundset,[5] and Andrew D. Foote[1,6,8,9,*]

[1]Department of Natural History, NTNU University Museum, Norwegian University of Science and Technology (NTNU), Erling Skakkes gate 47A, 7012 Trondheim, Norway
[2]Center for Evolutionary Hologenomics, GLOBE Institute, Faculty of Health and Medical Sciences, 1353 Copenhagen, Denmark
[3]Lundbeck Foundation GeoGenetics Centre, GLOBE Institute, University of Copenhagen, 1350 Copenhagen, Denmark
[4]Friedrich Miescher Laboratory of the Max Planck Society, Max-Planck-Ring 9, 72076 Tübingen, Germany
[5]Geological Survey of Norway (NGU), 7491 Trondheim, Norway
[6]Centre for Ecological and Evolutionary Synthesis, Department of Biosciences, University of Oslo, 0316 Oslo, Norway
[7]These authors contributed equally
[8]X (formerly Twitter): @AndrewFoote2
[9]Lead contact
*Correspondence: andrew.foote@ibv.uio.no
https://doi.org/10.1016/j.cub.2024.01.056

## SUMMARY

Directly observing the chronology and tempo of adaptation in response to ecological change is rarely possible in natural ecosystems. Sedimentary ancient DNA (sedaDNA) has been shown to be a tractable source of genome-scale data of long-dead organisms[1–3] and to thereby potentially provide an understanding of the evolutionary histories of past populations.[4,5] To date, time series of ecosystem biodiversity have been reconstructed from sedaDNA, typically using DNA metabarcoding or shotgun sequence data generated from less than 1 g of sediment.[6,7] Here, we maximize sequence coverage by extracting DNA from ~50× more sediment per sample than the majority of previous studies[1–3] to achieve genotype resolution. From a time series of Late Pleistocene sediments spanning from a marine to freshwater ecosystem, we compare adaptive genotypes reconstructed from the environmental genomes of three-spined stickleback at key time points of this transition. We find a staggered temporal dynamic in which freshwater alleles at known loci of large effect in marine-freshwater divergence of three-spined stickleback (e.g., *EDA*)[8] were already established during the brackish phase of the formation of the isolation basin. However, marine alleles were still detected across the majority of marine-freshwater divergence-associated loci, even after the complete isolation of the lake from marine ingression. Our retrospective approach to studying adaptation from environmental genomes of three-spined sticklebacks at the end of the last glacial period complements contemporary experimental approaches[9–11] and highlights the untapped potential for retrospective "evolve and resequence" natural experiments using sedaDNA.

## RESULTS AND DISCUSSION

Marine and freshwater ecotypes of three-spined stickleback *Gasterosteus aculeatus* are a model system for studying the genomics of parallel adaptation in wild populations.[12,13] Repeated morphological and physiological changes in association with marine and freshwater ecosystems have been found to have evolved in parallel in three-spined stickleback across the Northern hemisphere.[14] The genomic architecture associated with adaptive divergence in the three-spined stickleback is well characterized, with many alleles underpinning adaptive divergence tightly clustered on chromosomes and globally shared as standing genetic variation (SGV).[8,12,13] Freshwater adaptive alleles are proposed to be maintained as low-frequency SGV in marine populations[15,16] (Figure 1A). These characteristics are thought to have been key to the colonization of newly emerged freshwater habitats at the end of the last glacial period.[11,14,16] This model system for parallel adaptation is thus ideally suited for testing the potential for sedimentary ancient DNA (sedaDNA) to track adaptive genetic changes associated with ecological changes at the end of the Late Pleistocene. Specifically, we track adaptive changes during the formation of a post-glacial isolation lake, resulting from isostatic rebound: the rise of land mass following the retreat of the glacial ice sheets that previously depressed them[17] (Figure 1B).

A recent study[19] presented analyses of the genome of a Late Pleistocene stickleback generated from bones found in the transition sediment layer deposited during the formation of a post-glacial lake (Jossavannet freshwater lake in Finnmark, Norway, 70°270 N, 23°470 E). This study found freshwater alleles on a background of predominantly marine ancestry in the genome of this early colonist stickleback.[19] However, the ancestry of a single individual provides a snapshot in time and may not reflect the population's average allele frequencies (Figure 1A). Environmental genomes from sedaDNA promise the potential to yield time series of local

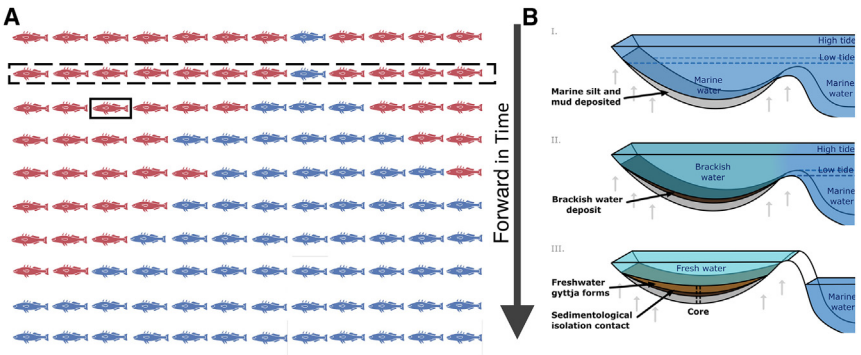


**Figure 1. Changes in freshwater and marine ancestry in three-spined sticklebacks during the formation of a post-glacial isolation lake**

(A) Schematic diagram of allele frequency changes from marine (red silhouettes) to freshwater (blue silhouettes) ancestry through successive generations. Sampling genomes from the present generation of sticklebacks can provide limited insights into past ancestry if much of the genetic variation coalesces in a recent common ancestor. Sampling of DNA from ancient bones (solid black rectangle) from an early generation, alive during the transition of the lake from a marine habitat, provides a "snapshot" of ancestry not found in the present generation. Sampling of DNA from consecutive sediment layers (dashed rectangle) potentially provides a genotype representative of the ancestral population at each time point and can track the chronology and tempo of genomic changes through time.

(B) Formation of isolation basin and freshwater lake. (I) New marine bays were exposed following the retreat of the ice sheets after the Last Glacial Maximum. (II) Isostatic rebound, indicated by gray arrows, causes the rise above sea level of the land mass. This reduces ingression of seawater at low tides. (III) The basin becomes fully isolated from the marine environment and becomes a freshwater habitat. Each phase leaves a distinct sedimentary layer.[18] The process is estimated to have occurred over a few decades (<100 years) based on the narrow ingression layer and well-defined isolation layer in the study lake.

population genetic variation and thereby elucidate evolutionary processes (Figure 1A). To test this, we collected sediment cores from Jossavannet for sedaDNA extraction. The lake, which is 38.4 m above sea level, was formed from a post-glacial isolation basin, with an isolation age of 12.9 ka BP (before present).[19] The core included marine, brackish, and freshwater phases of the isolation process (Figure 2A).

To maximize the complexity of our sedaDNA libraries, we performed extractions on bulk sediment (~25 g) from 10 consecutive sediment layers, each of approximately 1–2 cm thickness. The sequence contained 2 marine sediment layers, 7 layers from the brackish phase when seawater could still ingress the lake at high tides, and 1 layer after the isolation of the lake from the marine environment (Figure 2B). We generated single- and double-stranded libraries of DNA extracts from Late Pleistocene bones and sediment layers from Jossavannet, spanning the transition from marine to freshwater. Metagenomic analyses of species presence were consistent with the geological inference of paleoecology of each layer. From the bones, we generated additional collapsed paired-end read sequence data, which mapped at $\geq 1\times$ depth of coverage for 8.7 Mb (≈2.6%) of the repeat-masked genome. We mapped the collapsed paired-end sequence reads generated from the sedaDNA to the three-spined stickleback reference genome,[12] resulting in up to 5.7 million bp of the genome being covered by at least one sequencing read (Table S1).

Post-mortem DNA damage is characteristic of degradation found in ancient DNA samples, specifically deamination results in changes from cytosine to uracil primarily at the 5′ read termini, read by the sequencing platform as thyamine.[20–22] This characteristic damage pattern of an excess of C → T changes at the 5′ termini and the corresponding G → A changes at the 3′ termini of the double-stranded libraries used in the metagenomic analyses confirmed that the sedaDNA content was primarily of ancient origin (Figure 2C). Approximately 30% of the read-ends of the single-stranded sedaDNA libraries had an excess of C → T changes relative to the reference genome when all reads were included (Figures 2D and 2E), characteristic of post-mortem deamination.[20–22] There was a lower proportion of damaged reads in collapsed shotgun sequenced reads (~0.1) than target capture reads (~0.3) (Figure S1; Dryad: https://doi.org/10.5061/dryad.z8w9ghxkj). Post-mortem DNA processes such as hydrolysis result in fragmentation of DNA strands[20–22]; consistent with this, the mean read length was 77 bp. Such damage and fragmentation patterns are considered authentication of the ancient origin of the mapped DNA content in a sample.[20–22]

We next assessed the tempo and chronology of adaptive changes in regions of the genome associated with global divergence of marine and freshwater stickleback ecotypes.[12] In a principal-component analysis (PCA) of variation at the marine-freshwater divergent region of the genome among present-day stickleback genomes, freshwater samples from Jossavannet segregate along PC1 (PC1, $p < 0.001$) from the adjacent Altafjord marine population, explaining more than 30% of the genetic variation among samples (Figure S2). We found no clear correspondence between the chronological pattern of the sediment layers and the position of the environmental genomes along PC1. We attribute this to the variation in coverage of the marine-freshwater genomic regions among environmental genomes and particularly differences in coverage at loci of large effect on three-spined stickleback adaptation. We therefore focused on the two environmental genomes from which we obtained the greatest sequence coverage of the marine-freshwater divergent regions of the genome. These two genomes were from the sediment layer (769–770 cm) deposited after isolation from the marine ingression and a layer (774–775 cm) from the middle of the brackish phase of the isolation process when marine water could ingress at high tide. This brackish layer was the first in which we failed to detect *Fucus* and *Silvetia* in the associated metagenomic data (Figure 2C), suggesting that it represents a transition from an intertidal habitat due to less frequent marine ingression, for example, during the extremes of the tidal cycle.

As a conservative measure to remove any spuriously mapped reads originating from other species, we used only the sequence data from the target enrichment capture experiments,[23] designed to target transversions in the marine-freshwater divergent regions of the three-spined stickleback genome.[12] To reduce spurious calls due to C → T DNA post-mortem damage patterns caused by deamination and reduce bias due to variation in coverage, we randomly selected an allele at transversions in the marine-freshwater divergent regions of the two environmental genomes and

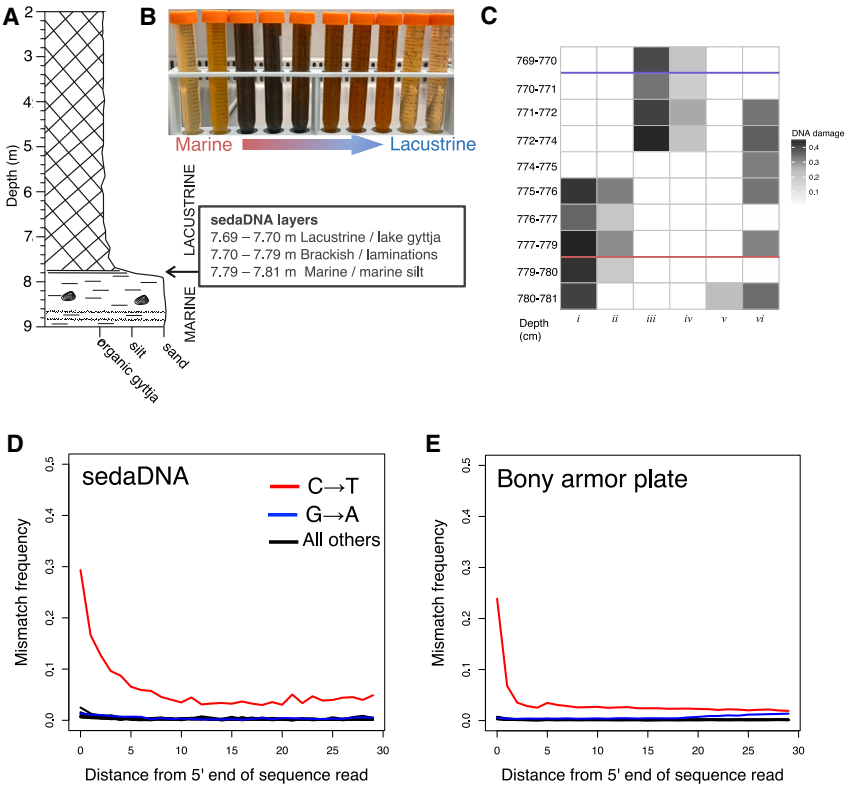


**Figure 2. Stratigraphy of Jossavannet core, showing the ecological and geological context of the sediment layers analyzed in this study, and post-mortem damage patterns authenticating ancient origins of the mapped DNA**

(A) Sediments were sampled from layers corresponding to the isolation phase, dated to 12.9 ka BP[19] and immediately preceding and following isolation. The bones previously analyzed in the study of Kirch et al.[19] were from the same stratigraphic depth, in a layer corresponding to the early isolation phase. The actual physical depth is deeper for the 2021 core analyzed in this study relative to the 2019 core that the bones were sampled from.

(B) Lysate and modified binding buffer from the bulk sediment extraction process, from left to right, represent 2 fully marine layers, 7 transitionary layers following uplift from isostatic rebound in which marine water can still ingress the basin at high tides, and 1 layer following the complete isolation from marine ingression.

(C) Heatmap showing taxa presence (detected by >100 reads; shaded cells) or absence (blank cells), and proportion of reads showing DNA damage (indicated by shading) of key indicator genera: (1) *Fucus*, (2) *Silvetia*, (3) *Microchloropsis*, (4) *Nannochloropsis*, (5) *Thalassiosira*, and (6) *Pseudopedinella*. *Fucus* and *Silvetia* are found in intertidal habitat; *Microchloropsis*, *Nannochloropsis*, *Thalassiosira*, and *Pseudopedinella* are found in marine, brackish, and freshwater habitats. Red and blue lines indicate the start and completion, respectively, of the isolation from marine environment, as determined by the stratigraphy and specifically changes in the distinct sedimentary boundary between the marine and freshwater lacustrine sediment facies (see Kirch et al.[19]).

(D and E) Post-mortem deamination patterns for filtered reads sequenced from (D) a stickleback bony armor plate and (E) sediment from layer 777–779 cm, both dated to the Late Pleistocene (12.9 ka) and mapped against the three-spined stickleback reference genome gasAcu1.0. Plots show a characteristic excess of C → T sequence misincorporation errors at the 5′ read termini relative to the modern stickleback reference.

See also Figure S1.

from present-day Jossavannet and Altafjord genomes. We then conditioned the probability of that allele representing either freshwater or marine ancestry based on the allele frequencies in the contemporary Jossavannet freshwater and adjacent marine populations using the Bayesian approach implemented in Kirch et al.[19] (Figure 3A). The two environmental genomes, which are from sediment layers deposited a few decades apart (<100 years), but from different stages of the isolation process, show similar ancestry profiles, despite differences in the specific sites covered (Figures 3B and 3C). As was previously found for the Late Pleistocene bones,[19] we identify more marine- than freshwater-associated ancestry in the two environmental genomes (Figures 3A and 3B). However, sequences from both environmental genomes have blocks of alleles with a probability of 1 of being freshwater ancestry (Figure 3D). These included several loci of known large effect, e.g., *EDA*, *WNT7B*, and *MUC5B*. For example, raw genotype plots (Figures 3B and 3C) show freshwater alleles at the minimal 16 kb haplotype in the *EDA* locus known to control repeated low armor evolution in sticklebacks[8,12,24] but flanked by marine alleles at neighboring sites.

These findings overlap with the genomic regions identified as those most rapidly evolving in contemporary release experiments. Roberts Kingman et al.[11] found that the most rapidly evolving regions following the introduction of marine sticklebacks to freshwater lakes were low recombination regions proximate to recombination hot spots, which overlapped with large numbers of quantitative trait loci (QTLs) and peaks of genetic differentiation between ecotypes. They included QTLs associated with defensive armor and trophic phenotype (e.g., *EDA*[8]) and kidney function (e.g., *WNT7B*[8]) on chrIV, for which we found freshwater genotypes as early as the brackish stage of the lake formation, differentiating the ancient population from the present-day fjord population (Figure 3). These traits increase fitness in response to the changes in pH, salinity, prey resources, and predators.[8,10,14] Low recombination rates can increase the physical linkage between selected alleles within haploblocks,[8,11–13] whereas neighboring recombination hotspots are proposed to increase the potential for re-assembly of haploblocks from smaller segments.[11] Consistent with the hypothesized mechanism of freshwater ancestry being re-assembled from distinct haploblocks rather than inherited from a single common ancestor, we see that although the *EDA* and *WNT7B* loci on chrIV carry freshwater alleles, the *SULT4A* locus on chrIV carries predominantly marine ancestry in the environmental genomes. This marine ancestry at *SULT4A* (a gene associated with neurotransmitter metabolism) has been replaced by freshwater ancestry in the contemporary lake population (Figures 3B and 3C).

Our findings of freshwater genotypes at key loci as early as the brackish phase of the lake's formation provide a useful

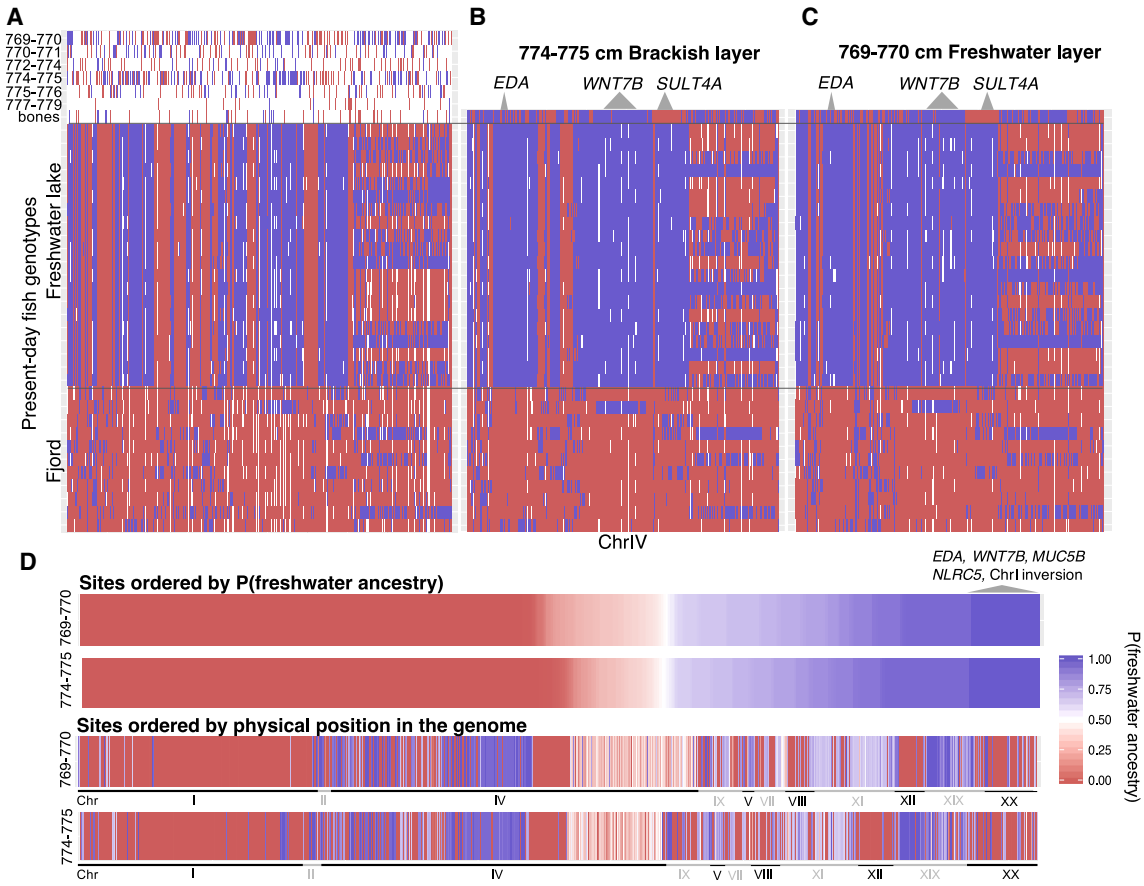


**Figure 3. Patterns of stickleback marine and freshwater ancestry in sediment layers**

(A–C) Visual genotypes at transversions in the marine-freshwater divergent regions of chromosome IV for contemporary Jossavannet (freshwater) and Altafjord (marine) sticklebacks, and (A) all ancient environmental genomes from sediment layers for which >100 sites were covered by capture enriched sequence data (top 6 rows), and the captured ancient stickleback bone library[19] (7th row), (B) 554 transversions in brackish layer (774–775 cm depth, top row), and (C) 465 transversions in freshwater layer (769–770 cm depth, top row). Rows are individual fish and columns are sites, which are shaded red for the most common allele in the marine population, blue for the alternative allele, and white for missing data.

(D) Probability of three-spined stickleback freshwater ancestry at 1,083 transversions from the freshwater sediment layer found at 769–770 cm depth, above the isolation boundary dated to 12.9 cal ka BP[19] and at 1,084 transversions from a brackish sediment layer found at 774–775 cm depth, from a period when sea water could still ingress the basin at high tides. Sites are first ordered based on probabilities from 0 to 1 and second by physical position in the genome. The probability of freshwater ancestry is dependent upon the allele sampled from the sediment sample, and the allele frequencies sampled from the contemporary freshwater Jossavannet lake and marine Altafjord populations. For probabilities to be either 0 or 1, alternative alleles must be fixed in each contemporary population. Sites are restricted to transversions in the marine-freshwater globally divergent regions of the genome[12] that were the targeted for enrichment by baits.

See also Figures S2 and S3.

comparison to experimental release studies that have introduced marine three-spined stickleback to freshwater ecosystems.[11] Several factors predict greater constraints on selection on alleles associated with freshwater adaptation in the marine sticklebacks that colonized the nascent Jossavannet lake[19] compared with those experimentally released into North Pacific lakes.[11] The ecological conditions in Jossavannet changed gradually, passing through several decades of inter-tidal and brackish phases, before becoming an isolated freshwater body. Additionally, the lake's formation followed the retreat of the glacial ice sheets and the emergence of the first suitable habitat for threes-spined sticklebacks in millennia. As such, there would have been limited opportunity at this time for freshwater migrants to re-seed the local marine population with freshwater alleles via the transporter mechanism.[16] Finally, the North Pacific, where the experimental releases were conducted,[11] is thought to be the source of origin of three-spined stickleback marine-freshwater ecotypes.[13,14] Accordingly, North Pacific marine populations can be richer in freshwater adaptive alleles as SGV than their Atlantic counterparts.[13] Despite the gradual ecological change to a freshwater ecosystem and the potential reduced availability of the SGV in the local marine population (relative to present-day North Pacific marine sticklebacks introduced to freshwater in the experimental releases[11]), our findings suggest that the efficacy of selection is sufficient to raise freshwater alleles to detectable frequencies at several known loci of large effect at the earliest stages of the lake's formation. The concordance of the findings from our sedaDNA approach and contemporary experimental releases[11] validates the ability of the environmental genomes derived from sedimentary layers to

provide accurate and valuable insights into evolutionary processes from the ecological changes at the end of the Late Pleistocene.

Important considerations for future studies when designing similar time-series sedaDNA projects include consideration of the timescale sediment cores that will need to be covered in order to track both rapid and prolonged genetic change associated with adaptation. We did not "catch adaptation in the act" so to speak; the signatures of freshwater adaptation were consistent between environmental genomes spanning brackish to freshwater. Additionally, future studies should consider how coverage of loci of different effect sizes biases comparisons between sediment layers. We found coverage of focal genomic regions associated with marine-freshwater adaptation to be highly stochastic among environmental genomes. Target capture enrichment[23] for known adaptive loci was critical for obtaining genotypes at marine-freshwater divergent loci (Table S1), which represent a tiny fraction of the metagenomic sedaDNA libraries. Surprisingly, target enrichment capture of several of the sedaDNA libraries outperformed capture of the ancient bone libraries (Figure 3A). Finally, achieving the complexity of genomic libraries needed for this study requires a large sediment mass input, which, in turn, increases the presence of inhibitors. A key challenge to this approach is thus to maximize library complexity while minimizing inhibition of enzymatic reactions such as library build and PCR steps. Experiments with training cores resulted in library build and PCR failure, despite successful sedaDNA extraction. Although there remain technical hurdles to overcome, we believe that our findings highlight the untapped potential for sedaDNA datasets to track evolution through time. In doing so, we take evolve-and-resequence experiments[25] out of the lab and into the field, retrospectively applying the approach to natural populations as they adapted to the climatic changes at the end of the Late Pleistocene. In turn, we may gain new insights into the potential responses to ongoing climatic change during the Anthropocene.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Sampling sediment cores
  - Present-day stickleback sample collection
- METHOD DETAILS
  - Sub-sampling sediment layers
  - Modern sample collection, library build and sequencing
  - Ancient DNA labwork
  - Single-stranded library build
  - Sediment aDNA extraction
  - Target enrichment capture
  - Mapping, filtering and masking of sedaDNA data
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Assessing postmortem DNA damage and contamination
  - Metagenomic analyses
  - Principal Component Analysis
  - Comparison of ancient and modern genotypes
  - Metagenomic Damage analysis

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2024.01.056.

## ACKNOWLEDGMENTS

We thank three anonymous reviewers whose feedback greatly improved this manuscript. This work was funded by a European Research Council (ERC) consolidator grant (ERC-COG-101045346 "EXPLOAD") and a Norwegian Research Council (NFR) ERC application support grant (324353) awarded to A.D.F.

## AUTHOR CONTRIBUTIONS

A.D.F. conceived and designed the study. A.R. collected and analyzed the geology of the sediment cores. A.D.F. conducted fieldwork to collect contemporary sticklebacks and subsequent genetic lab work. A.D.F., J.L., S.S.T.M., and N.F.G.M. performed the manual ancient DNA extractions, library builds, amplification, and target enrichment capture experiments. X.C. oversaw automated DNA extraction and library builds. M.W.P. analyzed metagenomic data. A.D.F. and J.L. analyzed stickleback paleogenomics. F.C.J. and M.T.P.G. provided expertise and resources related to stickleback genomics and ancient DNA lab work, respectively. A.D.F. wrote the manuscript with input from all coauthors.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: July 31, 2023
Revised: December 4, 2023
Accepted: January 22, 2024
Published: February 12, 2024

## REFERENCES

1. Gelabert, P., Sawyer, S., Bergström, A., Margaryan, A., Collin, T.C., Meshveliani, T., Belfer-Cohen, A., Lordkipanidze, D., Jakeli, N., Matskevich, Z., et al. (2021). Genome-scale sequencing and analysis of human, wolf, and bison DNA from 25,000-year-old sediment. Curr. Biol. *31*, 3564–3574.e9.
2. Vernot, B., Zavala, E.I., Gómez-Olivencia, A., Jacobs, Z., Slon, V., Mafessoni, F., Romagné, F., Pearson, A., Petr, M., Sala, N., et al. (2021). Unearthing Neanderthal population history using nuclear and mitochondrial DNA from cave sediments. Science *372*, eabf1667.
3. Pedersen, M.W., De Sanctis, B., Saremi, N.F., Sikora, M., Puckett, E.E., Gu, Z., Moon, K.L., Kapp, J.D., Vinner, L., Vardanyan, Z., et al. (2021). Environmental genomics of Late Pleistocene black bears and giant short-faced bears. Curr. Biol. *31*, 2728–2736.e8.
4. Dehasque, M., Ávila-Arcos, M.C., Díez-Del-Molino, D., Fumagalli, M., Guschanski, K., Lorenzen, E.D., Malaspinas, A.S., Marques-Bonet, T., Martin, M.D., Murray, G.G.R., et al. (2020). Inference of natural selection from ancient DNA. Evol. Lett. *4*, 94–108.
5. Ellegaard, M., Clokie, M.R.J., Czypionka, T., Frisch, D., Godhe, A., Kremp, A., Letarov, A., McGenity, T.J., Ribeiro, S., and John Anderson, N. (2020). Dead or alive: sediment DNA archives as tools for tracking aquatic evolution and adaptation. Commun. Biol. *3*, 169.

6. Alsos, I.G., Rijal, D.P., Ehrich, D., Karger, D.N., Yoccoz, N.G., Heintzman, P.D., Brown, A.G., Lammers, Y., Pellissier, L., Alm, T., et al. (2022). Postglacial species arrival and diversity buildup of northern ecosystems took millennia. Sci. Adv. *8*, eabo7434.
7. Garcés-Pastor, S., Coissac, E., Lavergne, S., Schwörer, C., Theurillat, J.P., Heintzman, P.D., Wangensteen, O.S., Tinner, W., Rey, F., Heer, M., et al. (2022). High resolution ancient sedimentary DNA shows that alpine plant diversity is associated with human land use and climate change. Nat. Commun. *13*, 6559.
8. Peichel, C.L., and Marques, D.A. (2017). The genetic and molecular architecture of phenotypic diversity in sticklebacks. Philos. Trans. R. Soc. Lond. B Biol. Sci. *372*, 20150486.
9. Bell, M.A., Aguirre, W.E., and Buck, N.J. (2004). Twelve years of contemporary armor evolution in a threespine stickleback population. Evolution *58*, 814–824.
10. Barrett, R.D., Rogers, S.M., and Schluter, D. (2008). Natural selection on a major armor gene in threespine stickleback. Science *322*, 255–257.
11. Roberts Kingman, G.A., Vyas, D.N., Jones, F.C., Brady, S.D., Chen, H.I., Reid, K., Milhaven, M., Bertino, T.S., Aguirre, W.E., Heins, D.C., et al. (2021). Predicting future from past: The genomic basis of recurrent and rapid stickleback evolution. Sci. Adv. *7*. eabg52852021.
12. Jones, F.C., Grabherr, M.G., Chan, Y.F., Russell, P., Mauceli, E., Johnson, J., Swofford, R., Pirun, M., Zody, M.C., White, S., et al. (2012). The genomic basis of adaptive evolution in threespine sticklebacks. Nature *484*, 55–61.
13. Reid, K., Bell, M.A., and Veeramah, K.R. (2021). Threespine stickleback: A model system for evolutionary genomics. Annu. Rev. Genomics Hum. Genet. *22*, 357–383.
14. Bell, M.A., and Foster, S.A. (1994). The Evolutionary Biology of the Threespine Stickleback (Oxford University Press).
15. Barrett, R.D.H., and Schluter, D. (2008). Adaptation from standing genetic variation. Trends Ecol. Evol. *23*, 38–44.
16. Schluter, D., and Conte, G.L. (2009). Genetics and ecological speciation. Proc. Natl. Acad. Sci. USA *106*, 9955–9962.
17. Romundset, A., Bondevik, S., and Bennike, O. (2011). Postglacial uplift and relative sea level changes in Finnmark, northern Norway. Quat. Sci. Rev. *30*, 2398–2421.
18. Romundset, A., Lakeman, T.R., and Høgaas, F. (2018). Quantifying variable rates of postglacial relative sea level fall from a cluster of 24 isolation basins in southern Norway. Quat. Sci. Rev. *197*, 175–192.
19. Kirch, M., Romundset, A., Gilbert, M.T.P., Jones, F.C., and Foote, A.D. (2021). Ancient and modern stickleback genomes reveal the demographic constraints on adaptation. Curr. Biol. *31*, 2027–2036.e8.
20. Dabney, J., Meyer, M., and Pääbo, S. (2013). Ancient DNA Damage. Cold Spring Harb. Perspect. Biol. *5*, a012567.
21. Briggs, A.W., Stenzel, U., Johnson, P.L.F., Green, R.E., Kelso, J., Prüfer, K., Meyer, M., Krause, J., Ronan, M.T., Lachmann, M., and Pääbo, S. (2007). Patterns of damage in genomic DNA sequences from a Neandertal. Proc. Natl. Acad. Sci. USA *104*, 14616–14621.
22. Orlando, L., Allaby, R., Skoglund, P., Der Sarkissian, C., Stockhammer, P.W., Ávila-Arcos, M.C., Fu, Q., Krause, J., Willerslev, E., Stone, A.C., and Warinner, C. (2021). Ancient DNA analysis. Nat. Rev. Methods Primers *1*, 14.
23. Enk, J.M., Devault, A.M., Kuch, M., Murgha, Y.E., Rouillard, J.M., and Poinar, H.N. (2014). Ancient Whole Genome Enrichment Using Baits Built from Modern DNA. Mol. Biol. Evol. *31*, 1292–1294.
24. Colosimo, P.F., Hosemann, K.E., Balabhadra, S., Villarreal, G., Jr., Dickson, M., Grimwood, J., Schmutz, J., Myers, R.M., Schluter, D., and Kingsley, D.M. (2005). Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. Science *307*, 1928–1933.
25. Kofler, R., and Schlötterer, C. (2014). A guide for the design of evolve and resequencing studies. Mol. Biol. Evol. *31*, 474–483.
26. Schubert, M., Lindgreen, S., and Orlando, L. (2016). AdapterRemoval v2: rapid adapter trimming, identification, and read merging. BMC Res. Notes *9*, 88.
27. Korneliussen, T.S., Albrechtsen, A., and Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. BMC Bioinformatics *15*, 356.
28. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760.
29. Skoglund, P., Northoff, B.H., Shunkov, M.V., Derevianko, A.P., Pääbo, S., Krause, J., and Jakobsson, M. (2014). Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal. Proc. Natl. Acad. Sci. USA *111*, 2229–2234.
30. Smit, A., Hubley, R., and Green, P. (2015). RepeatMasker Open-4.0. http://www.repeatmasker.org.
31. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.
32. Morgulis, A., Gertz, E.M., Schäffer, A.A., and Agarwala, R. (2006). WindowMasker: window-based masker for sequenced genomes. Bioinformatics *22*, 134–141.
33. Michelsen, C., Pedersen, M.W., Fernandez-Guerra, A., Zhao, L., Petersen, T.C., and Korneliussen, T.S. (2022). metaDMG-A Fast and Accurate Ancient DNA Damage Toolkit for Metagenomic Data. Preprint at bioRxiv. https://doi.org/10.1101/2022.12.06.519264.
34. Quinlan, A.R. (2014). BEDTools: the Swiss-army tool for genome feature analysis. Curr. Protoc. Bioinformatics *47*, 11.12.1–11.1234.
35. R Core Team (2023). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. https://www.R-project.org/.
36. Jowsey, P.C. (1966). An Improved Peat Sampler. New Phytol. *65*, 245–248.
37. Carøe, C., Gopalakrishnan, S., Vinner, L., Mak, S.S.T., Sinding, M.H.S., Samaniego, J.A., Wales, N., Sicheritz-Pontén, T., and Gilbert, M.T.P. (2018). Single-tube library preparation for degraded DNA. Methods Ecol. Evol. *9*, 410–419.
38. Kapp, J.D., Green, R.E., and Shapiro, B. (2021). A fast and efficient single-stranded genomic library preparation method optimized for ancient DNA. J. Hered. *112*, 241–249.
39. Rohland, N., Glocke, I., Aximu-Petri, A., and Meyer, M. (2018). Extraction of highly degraded DNA from ancient bones, teeth and sediments for high-throughput sequencing. Nat. Protoc. *13*, 2447–2461.
40. Meyer, M., and Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb. Protoc. *2010*. pdb.prot5448.
41. Schubert, M., Ginolhac, A., Lindgreen, S., Thompson, J.F., Al-Rasheid, K.A., Willerslev, E., Krogh, A., and Orlando, L. (2012). Improving ancient DNA read mapping against modern reference genomes. BMC Genomics *13*, 178.
42. Tracy, C.A., and Widom, H. (1994). Level-Spacing Distributions and the Airy Kernel. Commun. Math. Phys. *159*, 151–174.
43. Wang, L., Zhang, W., and Li, Q. (2020). AssocTests: An R Package for Genetic Association Studies. J. Stat. Soft. *94*, 1–26.
44. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.
45. Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., and Kuhn, M. (2019). Welcome to the tidyverse. J. Open Source Softw. *4*, 1686.
46. Wickham, H., and Chang, W. (2016). Package 'ggplot2'. The Comprehensive R Archive Network, Ver 2.

## STAR★METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Biological samples | | |
| Stickleback biosample data | This study | NCBI: PRJNA693136 |
| Chemicals, peptides, and recombinant proteins | | |
| AmpliTaq Gold | Thermo Fisher Scientific | Cat# N8080241 |
| 3M Sodium Acetate | Sigma-Aldrich | Cat# S7899 |
| Proteinase K 100MG | Sigma-Aldrich | Cat# 3115844001 |
| H2O, Molecular Biology Grade | Thermo Fisher Scientific | Cat# 10490025 |
| Tween 20 100ML | Thermo Fisher Scientific | Cat# 10113103 |
| Ethanol, Absolute, Mol Biology Grade | Thermo Fisher Scientific | Cat# 10644795 |
| 5M Sodium Chloride 100ML | Thermo Fisher Scientific | Cat# 10609823 |
| T4 DNA Ligase (2,000,000 U/mL) | New England Biolabs Inc. | Cat#M0202M |
| T4 Polynucleotide Kinase | New England Biolabs Inc. | Cat#M0201S |
| T4 RNA Ligase Buffer | New England Biolabs Inc. | Cat# B0216L |
| ET SSB | New England Biolabs Inc. | Cat# M2401S |
| ATP (100mM) | Thermo Fisher Scientific | Cat# R0441 |
| BSA | New England Biolabs Inc. | Cat# B9000 |
| PfuTurbo Cx Hotstart DNA Polymerase | Agilent | Cat#600414 |
| HighPrep PCR Clean-up System (50ml) | MagBio Genomics | Cat#AC-60050 |
| Buffer PE | QIAGEN | Cat# 19065 |
| Buffer PB | QIAGEN | Cat# 19066 |
| Buffer EB | QIAGEN | Cat# 19086 |
| Dithiothreitol | Thermo Fisher Scientific | Cat# R0861 |
| dNTP Set 100mM 100mL | Thermo Fisher Scientific | Cat# 10336653 |
| EDTA 0.5M pH 8.0 Fisher Bioreagents 500ML | Thermo Fisher Scientific | Cat# 10182903 |
| Tris HCl, 1M, pH 8.0, 100ML | Thermo Fisher Scientific | Cat# 10336763 |
| PEG8000 | Sigma-Aldrich | Cat# P2139 |
| Critical commercial assays | | |
| MinElute PCR Purification kit | QIAGEN | Cat# 28006 |
| Tapestation screenTape D1000 HS | Agilent | Cat# 5067-5584 |
| Qubit dsDNA HS Assay Kit | Thermo Fisher Scientific | Cat# Q32854 |
| Dneasy Blood & Tissue kit | QIAGEN | Cat# 69504 |
| NEBNext Ultra II DNA Library Prep Kit for Illumina | New England Biolabs Inc. | Cat# E7645S |
| MagAttract Power Soil Pro kit | QIAGEN | Cat# 47199 |
| Bioanalyzer High Sensitivity DNA Analysis kit | Agilent | Cat# 5067-4626 |
| Fragment Analyzer NGS Fragment Kit (1-6000bp) | Agilent | Cat# DNF-473-0500 |
| Deposited data | | |
| Raw sequence data (fastq format) | This study and Kirch et al.[19] | NCBI: PRJNA693136 |
| metaDMG-A output<br>Capture bait design details and sequences<br>Competitive mapping to RefSeq<br>Mitogenome Database output<br>Post-mortem damage plots | This study | DRYAD: https://doi.org/10.5061/dryad.z8w9ghxkj |
| Code for generating genotype probability plots | This study and Kirch et al.[19] | GitHub: 10.5281/zenodo.10519326 |
| Software and algorithms | | |
| AdapterRemoval v2 | Schubert et al.[26] | https://github.com/MikkelSchubert/adapterremoval |

(*Continued on next page*)

*Continued*

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| ANGSD | Korneliussen et al.[27] | http://www.popgen.dk/angsd/index.php/ANGSD |
| BWA | Li and Durbin[28] | https://github.com/lh3/bwa |
| PMDtools | Skoglund et al.[29] | https://github.com/pontussk/PMDtools |
| RepeatMasker | Smit et al.[30] | http://www.repeatmasker.org |
| Samtools | Li et al.[31] | http://www.htslib.org |
| WindowMasker | Morgulis et al.[32] | https://github.com/goeckslab/WindowMasker |
| metaDMG-A | Michelsen et al.[33] | https://metadmg-dev.github.io/metaDMG-core/ |
| Bedtools | Quinlan[34] | https://bedtools.readthedocs.io |
| R-package | R Core Team[35] | https://www.R-project.org |

## RESOURCE AVAILABILITY

### Lead contact

Further information and requests for resources, material and reagents should be addressed and will be fulfilled by the lead contact, Andrew Foote (andrew.foote@ibv.uio.no).

### Materials availability

Raw sequence data and BioSample details are available at the National Centre for Biotechnology Information (NCBI) under BioProject accession number: PRJNA693136.

### Data and code availability

- Sequencing data have been deposited in the NCBI database and are publicly available as of the date of publication. Accession numbers are listed in the key resources table. Additional output files have been deposited in a dryad repository, accession number is listed in the key resources table.
- All original code has been deposited on GitHub and is publicly available as of the date of publication. DOIs are listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

## EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

### Sampling sediment cores

Sediment core samples from Jossavannet freshwater lake (70°27'N, 23°47'E) in the Finnmark region of Norway were collected in late spring 2019 and 2021. The freshwater lake is relatively shallow and was cored with a ''Russian-type'' peat corer.[36] The one meter-long, half-cylinder-shaped samples of lake deposits were collected and transported to the laboratory at the Geological Survey of Norway (NGU: Norges Geologiske Undersøkelse). During basin isolation from the sea, fundamental environmental changes lead to a complete replacement of floral and faunal assemblages, resulting in clear marine-lacustrine transitions apparent in sediment core biostratigraphy.[18] Threespine stickleback bones were found in a core collected in 2019 in sediment formed during the isolation phase of the lake when the habitat would have been brackish.[19] The 2019 core was subsampled for material suitable for radiocarbon dating; 1-cm-thick slices of the core were wet-sieved and residual terrestrial plant remains were identified, picked and dried overnight before being submitted to the Poznan Radiocarbon Laboratory, Poland. A series of four samples were dated across the transition. These comprised of a sample of *Salix* leaves (Poz-115352, 7 mg) picked at the boundary yielded 11,080 ± 50 radiocarbon years, calibrated to 13070-12,800 cal year BP. This is supported by a marine sample of algae found 7 cm deeper (large sample Pox-115492, weighing 61 mg), which yielded 11780 ± 50, calibrated to 13,330- 13,070 cal year BP. It is further supported by a third sample of *Salix* leaves (Poz-115350, 5 mg) found shortly above the transition, yielding 11,420 ± 50, calibrated to 12,920-12,700 cal year BP. All calibration using OxCal software with IntCal.

### Present-day stickleback sample collection

Adult threespine stickleback specimens were collected using minnow traps in 2019 and 2021 from 9800 $m^2$ Jossavannet freshwater lake (70°27'N, 23°47'E) and from the outer branch area (next to the lake sites) of Altafjord (70°27'N, 23°46'E). Samples were collected under permit (201300202-62) from Finnmark Fylkeskommune. Upon sampling, sticklebacks were euthanized and stored in 95% ethanol.

## METHOD DETAILS

### Sub-sampling sediment layers

The three cores collected in 2021 were sampled destructively for sedimentary DNA analyses. These were first split longitudinally in to two halves at the laboratory. One half was wet-sieved with a 125μm mesh for additional stickleback hard parts, such as bones, bony plates and spines. However, no additional stickleback remains were found. Sediment layers spanning the marine-lacustrine transition were sampled from the remaining half. Samples consisted of approximately 1cm depth consecutive layers, with sediment material taken from the inner core to avoid the 'smearing' effect associated with the outer core. To further avoid cross-contamination between sediment layers, all equipment was sterilized using bleach and 70% ethanol, and latex gloves and paper sleeves were changed between each sampling. Sediment layers were transferred to sterile UV-treated 50ml falcon centrifuge tubes for DNA extraction.

### Modern sample collection, library build and sequencing

DNA was extracted from the modern samples collected in 2021 using a Qiagen Blood and Tissue kit following the manufacturer's protocol, and built into dual-indexed libraries using a NEBNext Ultra II library preparation kit in the laboratory facilities of NTNU, Trondheim, Norway. Paired-end 250bp sequencing on a lane of an Illumina NovaSeq 6000 S4 flow cell was performed by a commercial service offered by Novogene.

### Ancient DNA labwork

Ancient DNA lab work was conducted in the dedicated ancient DNA facilities of the Globe Institute, University of Copenhagen. Details for the extraction of DNA from ancient stickleback bones and spine from the Jossavannet sample are given in Kirch et al.[19] Briefly, DNA was extracted using a silica-based method, where each individual bone or spine was incubated overnight under motion at 55°C in 500 uL extraction buffer (0.45 M EDTA, 0.1 M UREA, 100 mg proteinase K). Each sample was then centrifuged at 2300 rpm for 5 min and the supernatant was collected and concentrated and purified using a QIAGEN MinElute spin column (QIAGEN, Valencia, CA, USA).

### Single-stranded library build

Double-stranded Illumina libraries had been previously constructed using the blunt-end single tube (B.E.S.T.) method[37] and those libraries were then sequenced to saturation for the Kirch et al. study,[19] resulting in coverage of 16,923,179 bp of the threespine stickleback genome. For this study, single-stranded DNA libraries were constructed from using 20 μl of the DNA extracts from Kirch et al.[19] using the method in Kapp et al.[38] The adapters and single-stranded binding proteins (SSB) were prepared and diluted following the recommendation for concentration defined by the DNA input tier system described in the protocol published by Kapp et al.[38] As a conservative approach to avoid adapter dimer, we used tier 5 in all reactions. Libraries were purified with the MinElute reaction clean-up kit (Qiagen), and were eluted in 35 μl EBT. A volume of 15 μl of DNA library template was used for dual-indexing PCR amplification following a AmpliTaq Gold (ThermoFisher Scientific) protocol at 15 cycles for each sample. Each 50μl PCR reaction contained 15μl of library, 1μl of AmpliTaq Gold polymerase, 0.5μM dNTP, 1x PCR buffer, 0.25 mM MgCl2, 0.4μl of Bovine Serum Albumin (BSA), and was made up to 50μl with molecular grade water. PCR temperature profile included an activation step at 95°C for 5 min, followed by 15 cycles of denaturation at 95°C for 30 s, annealing at 55°C for 30 s and elongation at 72°C for 1 min, with a final extension step at 72°C for 7 min. An additional dual-indexed PCR per library was setup using 5μl of library, 1μl of PFU Turbo Cx polymerase (Agilent), 5μl of buffer, 0.5μM dNTP, 0.4μl of BSA and made up to 50μl with molecular grade water. PCR products were then purified using 1.8× HighPrep PCR beads (MagBio Genomics). Extraction, library build and index PCR blanks were included to evaluate potential contamination during the library building process. Multiple PCRs were performed on each library to maximize complexity for downstream analyses.

### Sediment aDNA extraction

After first removing approximately 0.3g for an automated extraction process (see below), the remaining sediment in the falcon tube was weighed on a balance. Extraction weights ranged from 21.4-29.9g after subtracting the weight of the falcon tube. Lysis buffer was prepared as per Rohland et al.,[39] but scaled up to individual reaction size of 5ml of lysis buffer per sediment sample. Each 5ml aliquot lysis buffer contained 4.5 ml of 0.5 M EDTA (pH 8.0) combined with 0.372 ml of water, 2.5 μl of Tween 20 and 125 μl of 10 mg/ml proteinase K. The lysis buffer was added to the sediment sample and mixed by 10 seconds of light vortexing, followed by flicking and inverting the tube. Each tube was then sealed with parafilm and then gently mixed by continuously rotating for 24 hours at 18 r.p.m. whilst incubating at 37°C. The 50ml centrifuge tubes containing the lysate were then centrifuged at 4000 r.p.m. for 10 mins to separate the sediment and lysate. The lysate was transferred to a new 15 ml falcon tube. The amount of lysate varied slightly among samples. When possible we removed 4ml, but for some samples, even after centrifugation, it was not possible to remove the full 4ml without disturbing the pelleted sediment, however we took >3ml from all samples. In all cases we added molecular grade water to make all samples up to 5 ml.

To bind the DNA to the silica membrane of the spin column, we first modified stock 500 ml of Qiagen PB binding buffer by adding 2.5ml 5M NaCl and 15ml of 3M NaOAc.[40] The modified binding buffer was then added to the lysate at a ratio of 5x the volume of lysate and mixed with pipetting and gentle vortexing. This resulted in a total volume of 30ml of lysate from each sediment sample plus

binding buffer mix, each of which was split into two 15ml falcon tubes (Figure 2). Aliquots of 600ul lysate plus binding buffer mix were added to individual Qiagen MinElute spin columns and spun for 10 minutes a 10,000 r.p.m. The silica membrane in the spin column was then subject to two wash steps each using 500 ul of Qiagen PE buffer and centrifuged for 1 minute at 10,000 r.p.m., with an additional drying centrifuge step of 1 minute at 13,300 r.p.m. DNA was then eluted by adding 15ul per spin column of Qiagen EB buffer pre-heated to 37°C, which was then incubated at 37°C for 15 minutes before being spun down in a centrifuge for 1 min at 8,000 r.p.m. These steps were repeated, each time in a new set of spin columns, until all the lysate plus binding buffer mix for each sample had been filtered. Aliquots of eluted DNA from the same sediment layer sample were then combined and gently mixed by pipette. Aliquots of extract and unamplified libraries were taken to the modern DNA laboratory for QC checks using the both the TapeStation (Agilent Technologies) and either the Fragment Analyzer (Agilent Technologies) or 2100 Bioanalyzer (Agilent Technologies), in each case using high sensitivity chips for visualization of fragment length and DNA yield.

The sedaDNA extracts were then built into single-stranded libraries using the method given above and amplified using both PCR protocols (PFU Turbo Cx polymerase and AmpliTaq Gold). PCR products were purified using 1.8× HighPrep PCR beads (MagBio Genomics). QC checks on the purified post-PCR product were performed as above.

In addition to the bulk-sediment lab work we submitted samples to an automated ancient eDNA workflow. Here, DNA extraction and library preparation were performed in dedicated clean rooms at the Globe Institute, University of Copenhagen, Denmark. Sediment subsamples (~0.2 grams) were transferred to SAFE® 2D barcoded tubes (LVL technologies) and added 200 μL Omni 0.1 mm as well as 200 μL Omni 0.5 mm Ceramic Bulk Beads. The DNA was hereafter extracted on a Tecan Fluent DreamPrep 780 using the Qiagen® MagAttract® Power Soil Pro kit following manufacturer's protocol with the following modifications: The LVL Bead Tube Rack was vortexed using a FastPrep-96™ at 1600 rpm for 30 seconds, paused for 30 seconds, and then added another 1600 rpm for 30 seconds. Hereafter, all samples were incubated shaking at a speed 1000 rpm overnight at 37°C. The DNA lysate volume was reduced to 240 μL for downstream purification, and adjusting the binding buffer ratio to 1:3 lysate:binding buffer. Lastly, all extracts were converted to dual-indexed Illumina double-stranded DNA libraries following the Meyer and Kircher[40] method, and sequenced on a NovaSeq 6000 XP lane running 100 Bp paired-end.

### Target enrichment capture

To maximize the coverage at informative marine-freshwater divergent sites in the genome that could be used in downstream ancient DNA analyses, we performed target capture enrichment experiments.[23] Using genetic coordinates of regions of the genome previously identified as being consistently differentiated between marine and freshwater populations in a global dataset,[12] we then screened these regions for transversions. We included filters specifying the probability that any SNP was polymorphic should be $P < 0.000001$, have minimum mapping quality and minimum base quality of 20, a minimum minor allele frequency of 0.25, and be in a region uniquely mapping within the stickleback genome. This screening identified 10,149 transversions. A total of 40,596 biotinylated RNA baits[23] consisting of 60 nucleotide probes were then designed by Arbor Daicel Biosciences, so that two probes covered each SNP, one for each allele, and two additional probes were immediately upstream and downstream of the SNP. From this set, 38,391 passed QC checks using BLAST to screen for redundancy, repeats and hits to the mitochondrial genome. This reduced our targeted SNPs to 10,026; of these 9,177 retained all four probes, 384 retained 3 probes, 66 retained 2 probes, and 399 targeted SNPs retained one probe. Baits are commercially available through MyBaits, Daciel Arbor Biosciences, specifying Design ID: D1029010KNT and Ref#: 220126-900. A subset of aliquots of amplified libraries were subject to enrichment following the manufacturer's High Sensitivity protocol, but with a single hybridization step at 60°C and no post-capture PCR. Enriched and non-enriched libraries were then sequenced across a lane of an Illumina NovaSeq 6000 S4 flow cell.

### Mapping, filtering and masking of sedaDNA data

AdapterRemoval[26] was used to remove adapters and trim both Ns and low-quality bases from the reads. Trimmed sequencing data were aligned against the reference genome gasAcu1[12] by using the Burrows-Wheeler Alignment Tool (BWA) with the aln algorithm,[28] disabling seeding (option -l 1024) to turn-off seeding thereby increasing mapped data by including reads with post mortem damage at the read-ends.[41] The resulting bam files were sorted and merged using Samtools.[31] Duplicate reads in all sorted bam files were identified by Samtools markdup. Masked regions encompassing interspersed repeats and low complexity DNA sequences detected by Repeat-Masker[30] and covering 3.72% of the stickleback genome, and highly repetitive DNA sequences detected by WindowMasker[32] from the NCBI C++ toolkit covering 25.59% of the stickleback genome (constructed using -sdust true as setting) were removed from the bam files using bedtools.[34] **See** Table S1 for summary statistics of mapping. Furthermore, due to potential reference bias towards freshwater alleles when mapping to the reference assembly, which was generated from a Pacific freshwater stickleback,[12] we created an alternative marine reference from a present-day Altafjord sample by using the -dofasta option implemented in ANGSD.[27] Sequence data were re-mapped to this alternative reference using the steps above. Analyses duplicated using the alternative assemblies provided qualitatively similar results (See also Figure S3).

## QUANTIFICATION AND STATISTICAL ANALYSIS

### Assessing postmortem DNA damage and contamination

Analyses of potential nucleotide misincorporations using PMDtools[29] to compare with the modern reference genome revealed that sequencing reads exhibited characteristic post-mortem damage patterns.[20–22] Damage found in DNA extracted from the sediment

layer was estimated from a single-strand library, we therefore see just C>T changes at both the 5 and 3 read termini (see also Figure S1 and https://doi.org/10.5061/dryad.z8w9ghxkj). In contrast, single- and double-strand libraries were built on DNA extracted from the bony plate and spine, then merged and analyzed for DNA damage. Hence, we also see an excess of G>A (the reverse complement of C>T) changes at the 3 read termini for the bony plate and spine. Therefore, except where otherwise stated, only transversions were considered in downstream analyses that included the ancient samples.

## Metagenomic analyses

We first investigated the metagenomic content of our sedaDNA libraries through competitively mapping our sequencing data against the RefSeq mitochondrial database (version 92). We found no reads competitively mapped to the threespine stickleback mitochondrial genome (NC_041244.1), nor to the two closest related species in the database, the ninespine stickleback (*Pungitius pungitius*) and blackspotted stickleback (*Gasterosteus wheatlandi*), indicating that threespine stickleback DNA would be at best at low frequencies in our genomic libraries. When we compared different read groups, we found that the highest numbers of unmated paired-end reads mapped to the horse-fly (*Haematopota turkestanica*). However, we found that reads stacked up over a small region of 476 bp and attribute this to the small size of the unmated paired-end reads stacking up in regions of low mapability. When we sorted taxa by the number of covered bases, the green algae (*Nannochloropsis limnetica*) mitochondrial reference (NC_022256.1) was consistently the top hit, due to coverage from the longer collapsed pair-end reads. This observation combined with our observations of high percentage of unmated reads mapping to the threespine stickleback relative to collapsed paired-end reads, yet not increasing the number of covered bases, made us exclude the unmated reads from our final analyses of adaptive genotypes. The statistical outputs from competitively mapping to the RefSeq mitochondrial database are accessible in the dryad respository:
https://doi.org/10.5061/dryad.z8w9ghxkj.

## Principal Component Analysis

We used pseudo-haploid genotype calls of present-day marine and freshwater sticklebacks and the ancient samples. We compared covariance within freshwater-marine divergence associated regions among samples using ecology informative markers. The ancient samples were included in the PC computations and not projected onto PCs of modern samples, which has the advantage of providing a quality control measure. For example, if the ancient samples were impacted by sequencing- or sequence data processing errors, the samples would appear as outliers in the PCA. Additional filtering steps included in these analyses were the removal of regions of poor mapping quality ($Q < 20$), removal of sites with low base quality scores ($q < 20$), calling only SNPs inferred with a likelihood ratio test (LRT) of $P < 0.000001$, a minimum allele frequency of 0.25, specifying uniquely mapping reads only, and the removal of transitions to avoid bias from C to T and G to A DNA damage patterns. The eigenvectors from the covariance matrix were generated with the R function ''eigen'', and significance was determined with a Tracy-Widom test[42] performed in the R-package AssocTest[43] to evaluate the statistical significance of each principal component. See also Figure S2.

## Comparison of ancient and modern genotypes

To investigate the marine and/or freshwater origin of adaptive alleles carried by environmental genomes we assigned the ancestral state of each allele, conditioning on allele frequencies in present-day marine (Altafjord) and freshwater (Jossavannet) populations using the method from Kirch et al.[19] In contrast to Kirch et al., we did not include base quality re-calibrated transitions, as the re-calibration can be biased by paired-end read data. We therefore conservatively restricted our analyses to transversions that were targeted by the enrichment capture probes. We also estimate genotype probability at the per-site level, rather than per-window as in Kirch et al.[19]

We defined two possible ancestries for the ancient genome, $A_{marine}$ and $A_{fresh}$.

Where the allele carried by the ancient genome shares most recent common ancestry with other marine and freshwater fish respectively. The probability of observing an allele (0,1) in the ancient genome given a specific ancestral state can therefore be calculated as:

*Pr(allele|Ancestry):*

$$P(0|A_{marine}) = f_{marine}(0)$$

$$P(1|A_{marine}) = f_{marine}(1)$$

$$P(0|A_{fresh}) = f_{fresh}(0)$$

$$P(1|A_{fresh}) = f_{fresh}(1)$$

We assume uniform prior probabilities $P(A_{marine})$ and $P(A_{fresh})$, and calculate the posterior probability of each possible ancestral state as:

$$P(Ancestry|allele) = \frac{P(allele|Ancestry)P(A)}{\sum_i P(Ancestry_i|allele)}$$

## Metagenomic Damage analysis

We analyzed the eukaryotic fraction of the shotgun sequenced metagenomes by compiling a database consisting of all NCBI RefSeq mitochondria and plastid databases (downloaded November 6th 2022), including the GTDB (release 207) as microbial decoy database. The database were hereafter indexed using bowtie2[44] and all reads competitively mapped against it setting the parameters: –end-to-end (-D 15 -R 2 -N 0 -L 22 -i S,1,1.15). The resulting alignments were hereafter sorted by readID using SAMtools[31] and parsed for taxonomic profiling and DNA damage estimated using metaDMG-A.[33] The metaDMG output were next parsed to R,[35] filtering for taxa with more > 100 reads, a lambda likelihood ratio > 3 and mean length of reads mapped to each taxonomic node > 35Bp using tidyverse[45] and plotting key taxa using the ggplot[46] package.