Received: 31 May 2022 | Accepted: 8 July 2022

DOI: 10.1111/1755-0998.13686

MOLECULAR ECOLOGY RESOURCES WILEY

NEWS AND VIEWS

PERSPECTIVE

# Drawing a line in the sand: Environmental DNA population genomics

**Taylor Matthew Wilcox**[1] | **Mads Reinholdt Jensen**[2]

[1]Rocky Mountain Research Station, USDA Forest Service, National Genomics Center for Wildlife and Fish Conservation, Missoula, Montana, USA

[2]Department of Biology, Aarhus University, Aarhus C, Denmark

**Correspondence**
Taylor Matthew Wilcox, Rocky Mountain Research Station, USDA Forest Service, National Genomics Center for Wildlife and Fish Conservation, 800 E Beckwith Ave, Missoula, MT, USA.
Email: taylor.wilcox@usda.gov

**Handling Editor:** Joanna Kelley

Environmental DNA (eDNA) sampling uses genetic material in the environment to infer species presence sight-unseen. The method has rapidly become a powerful tool for monitoring biodiversity. However, biological diversity, as per the Convention on Biological Diversity definition of "diversity within species, between species and of ecosystems" is more inclusive than most eDNA studies cover: The vast majority focus only on between-species and ecosystem-level biodiversity. However, a tantalizing prospect, as illustrated by Farrell et al. (2022) in this issue of *Molecular Ecology Resources*, is that we might also be able to unlock information about individual and population-level diversity via population genomic analysis of these environmental samples. Farrell et al. (2022) found that targeted samples of beach sand contained genetic material not just informative about sea turtle presence, but also indicated the presence of pathogens and genome-wide mitochondrial and nuclear sequences that could accurately infer individual turtle source population. Moving from proof-of-concept to robust, population genomic inference will require a growth of genomic resources for nonmodel organisms and careful study design considerations, some of which have already been pioneered by related fields.

**KEYWORDS**
high-throughput sequencing, HTS, individual-based sequencing, noninvasive genetics, pooled sequencing, study system, wildlife monitoring

The eDNA sampling community has been deeply interested in extracting population genomic inferences from environmental samples. Different sampling and analysis designs will be best suited to different types of questions. We foresee there being three general types of population-level inferences from eDNA samples: Metaphylogeography, noninvasive sampling of individuals, and targeted environmental sampling of populations.

Metaphylogeography: Environmental samples contain genetic material from potentially hundreds of taxa, opening the door to "metaphylogeographic" studies of biogeographic connectivity such as the effect of dispersal barriers and environmental gradients across many taxa and their pathogens simultaneously (Turon et al., 2020). For example, Shum and Palumbi (2021) used amplicon sequencing of 106 cobble samples from kelp forests to reveal small-scale ecological gradients in 527 species of animals and algae.

Noninvasive sampling of individuals: Studies may use environmental samples to noninvasively sample single individuals, which may be desirable for species that are of conservation concern or difficult to capture. This is essentially noninvasive genetics (Schwartz et al., 2007), but allows researchers to collect genetic material from sign alone, without requiring capture of a visible portion of the individual of interest (e.g., hair, scat). Farrell et al. (2022) were able to isolate trails left by individual turtles for sand sampling, similar to how scientists have extracted DNA from animal tracks in snow (e.g., Franklin et al., 2019). Both mother turtles and their hatchlings left genetic material behind

at nest sites, but Farrell et al. (2022) were able to leverage the maternal inheritance of mitochondrial DNA to make inferences about population membership of the mothers based on mitochondrial reads.

Targeted environmental sampling of populations: Finally, eDNA studies may retrieve nuclear data from a group of individuals simultaneously by targeting aggregations of a taxon for environmental sampling and thus mimic a pooled sequencing approach (e.g., Ferretti et al., 2013). For example, Andres et al. (2021) showed how microsatellite loci could be informative in estimating allele frequencies and identifying numbers of contributing individuals using water samples from experimental mesocosms and natural settings. In a natural setting, Jensen et al. (2021) used capture enrichment to sequence environmental samples for nuclear whale shark (*Rhincodon typus*) DNA and achieved 0.55% on-target reads, covering 61.6% of the targeted regions, albeit with low coverage. Capture efficacy was limited by cocapture of highly abundant mackerel tuna, leaving room for optimization of molecular methods and more targeted sampling approaches, taking closely related, co-occurring taxa into account.

Of these three approaches, metaphylogeography is an application-ready tool and suitable data sets are already available from metabarcoding, but scaling up to population genomics from environmental samples will require some careful thought and more extensive genomic resources (Figure 1).

A particularly important consideration for population genomics from environmental samples will be correct mapping of rare target sequences within the larger suite of DNA found in environmental samples. Reads might map to the target taxon genome, but actually be derived from a closely-related taxa or even a distantly-related taxa if they occur at either a highly conserved or highly repetitive locus, particularly when working with short reads. This was a consideration highlighted by Jensen et al. (2021) where, even with capture enrichment, careful attention was required to ensure that mapped reads only came from the species of interest. In Farrell et al. (2022) only mitochondrial reads were interpreted, but the authors suggest the potential to leverage thousands of reads mapping to the nuclear genome to make even more powerful population inferences. Farrell et al. (2022) may not have had the genomic resources necessary to confidently filter and map these reads, but genomic resources are on the horizon. Large-scale efforts such as the Earth Biogenome Project (Lewin et al., 2020) are generating the multiple, well-curated genomes that we need for making environmental population genomic inferences.

**FIGURE 1** Sand track and nesting site left on a beach in northern Florida, USA.
Photo credit: Dr. Lucas Meers.

Where these genomic resources exist, ancient DNA studies have pioneered the potential to make robust population genomic inferences just from a bit of dirt. For example, Pedersen et al. (2021) retrieved low-coverage environmental bear genomes (0.03–0.04x) from cave sediments dated 14–16 thousand years BP, and used newly generated bear genomes to separate sequences of closely related species copreserved in environmental samples. With these data the authors were able to make new inferences about the evolutionary relationships and ancient dispersal of both extant and extinct bears. Studies such as Farrell et al. (2022) suggest that, with careful study design, modern eDNA samples may be an equally viable source of genetic material for population genomic inference on contemporary populations. Researchers looking to leverage this contemporary DNA may benefit from the tools, approaches, and best practices pioneered by ancient DNA researchers, who have spent the last decade(s) developing tools for such inferences from complex sample mixtures. Further, careful consideration will be required before attempting to apply these tools. Beyond novelty, focus should remain on applying the best tool for the specific research question. In some cases noninvasive environmental genetics may match or outperform invasive tissue sampling. In others, eDNA-based approaches may not provide the same resolution as is possible from invasive tissue sampling. When conceptualizing a new study, researchers should carefully consider potential trade-offs between the data resolution needed for their question and the conservation costs of invasive sampling, particularly for species of special conservation concern.

## AUTHOR CONTRIBUTIONS

Taylor Matthew Wilcox and Mads Reinholdt Jensen conceived and wrote the manuscript.

## CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

## ORCID

*Taylor Matthew Wilcox* https://orcid.org/0000-0003-3341-7374
*Mads Reinholdt Jensen* https://orcid.org/0000-0001-8240-1083

## REFERENCES

Andres, K. J., Sethi, S. A., Lodge, D. M., & Andrés, J. (2021). Nuclear eDNA estimates population allele frequencies and abundance in experimental mesocosms and field samples. *Molecular Ecology*, *30*(3), 685–697. https://doi.org/10.1111/mec.15765

Farrell, J. A., Whitmore, L., Mashkour, N., Rollinson Ramia, D. R., Thomas, R. S., Eastman, C. B., Burkhalter, B., Yetsko, K., Mott, C., Wood, L., Zirkelbach, B., Meers, L., Kleinsasser, P., Stock, S., Libert, E., Herren, R., Eastman, S., Crowder, W., Bovery, C., … Duffy, D. J. (2022). Detection and population genomics of sea turtle species via noninvasive environmental DNA analysis of nesting beach sand tracks and oceanic water. *Molecular Ecology Resources*. https://doi.org/10.1111/1755-0998.13617

Ferretti, L., Ramos-Onsins, S. E., & Pérez-Enciso, M. (2013). Population genomics from pool sequencing. *Molecular ecology*, *22*(22), 5561–5576. https://doi.org/10.1111/mec.12522

Franklin, T. W., McKelvey, K. S., Golding, J. D., Mason, D. H., Dysthe, J. C., Pilgrim, K. L., Squires, J. R., Aubry, K. B., Long, R. A., Greaves, S. E., Raley, C. M., Jackson, S., MacKay, P., Lisbon, J., Sauder, J. D., Pruss, M. T., Heffington, D., & Schwartz, M. K. (2019). Using environmental DNA methods to improve winter surveys for rare carnivores: DNA from snow and improved noninvasive techniques. *Biological Conservation*, *229*, 50–58. https://doi.org/10.1016/j.biocon.2018.11.006

Jensen, M. R., Sigsgaard, E. E., Liu, S., Manica, A., Bach, S. S., Hansen, M. M., Møller, P. R., & Thomsen, P. F. (2021). Genome-scale target capture of mitochondrial and nuclear environmental DNA from water samples. *Molecular Ecology Resources*, *21*(3), 690–702. https://doi.org/10.1111/1755-0998.13293

Lewin, H. A., Richards, S., Lieberman Aiden, E., Allende, M. L., Archibald, J. M., Bálint, M., Barker, K. B., Baumgartner, B., Belov, K., Bertorelle, G., Blaxter, M. L., Cai, J., Caperello, N. D., Carlson, K., Castilla-Rubio, J. C., Chaw, S. M., Chen, L., Childers, A. K., Coddington, J. A., … Zhang, G. (2020). The earth BioGenome project 2020: Starting the clock. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(4), e2115635118. https://doi.org/10.1073/pnas.2115635118

Pedersen, M. W., De Sanctis, B., Sarem, N. F., Sikora, M., Puckett, E. E., Gue, Z., Moon, K. L., Kapp, J. D., Vinner, L., Vardanyan, Z., Ardelean, C. F., Arroyo-Cabrales, J., Cahill, J. A., Heintzman, P. D., Zazula, G., RDE, M. P., Shapiro, B., Durbin, R., & Willerslev, E. (2021). Environmental genomics of late Pleistocene black bears and Giant 1 short-faced bears. *Current Biology*, *31*(12), 2728–2736.e8. https://doi.org/10.1016/j.cub.2021.04.027

Schwartz, M. K., Luikart, G., & Waples, R. S. (2007). Genetic monitoring as a promising tool for conservation and management. *Trends in Ecology & Evolution*, *22*(1), 25–33. https://doi.org/10.1016/j.tree.2006.08.009

Shum, P., & Palumbi, S. R. (2021). Testing small scale ecological gradients and intraspecific differentiation for hundreds of kelp forest species using haplotypes from metabarcoding. *Molecular Ecology*, *30*(13), 3355–3373. https://doi.org/10.1111/mec.15851

Turon, X., Antich, A., Palacín, C., Præbel, K., & Wangensteen, O. S. (2020). From metabarcoding to metaphylogeography: Separating the wheat from the chaff. *Ecological Applications*, *30*(2), e02036. https://doi.org/10.1002/eap.2036

**How to cite this article:** Wilcox, T. M., & Jensen, M. R. (2022). Drawing a line in the sand: Environmental DNA population genomics. *Molecular Ecology Resources*, *00*, 1–3. https://doi.org/10.1111/1755-0998.13686