

# Teasing apart the dilution effect by combining DNA metabarcoding and statistical modeling

Taal Levi<sup>1</sup> and Aimee Massey<sup>1</sup>

<sup>1</sup>Oregon State University

February 9, 2023

## Abstract

How changes in biodiversity affect disease, particularly in the face of large-scale land-use change, is a contentious topic in disease ecology that has implications for public health and conservation policy. The ‘dilution effect’ hypothesis argues that declines in biodiversity are associated with increased disease risk, but this can be challenging to demonstrate because many pathogens have complex life cycles such that changes to the species composition and abundance of hosts can influence the density and infection prevalence of vectors via multiple mechanisms. Key to addressing this debate is a quantification of interactions between hosts, vectors, and pathogens. In their recent study published in *Molecular Ecology*, Kocher et al. (2022) captured thousands of sandflies, some species of which are vectors for the *Leishmania* protozoan that causes Leishmaniasis, across a human footprint gradient in French Guiana (Fig. 1). By implementing DNA metabarcoding of vectors combined with an innovative modeling approach, they effectively quantified the nuanced relationships between changes in land-use, mammalian host diversity, vector abundance, and parasite prevalence. In support of the dilution effect hypothesis, Kocher et al. found that sites with higher mammal diversity were associated with lower relative abundance of reservoir hosts and higher *Leishmania* infection prevalence in sandflies. However, while infection prevalence was lower when mammal diversity was high, the density of sandfly vectors was higher, which resulted in a weak overall effect of mammal diversity on the density of infected vectors, the most important indicator of *Leishmania* transmission risk.

**Title:** Teasing apart the dilution effect by combining DNA metabarcoding and statistical modeling

Taal Levi<sup>1</sup>, Aimee L. Massey<sup>1</sup>

<sup>1</sup> Department of Fisheries, Wildlife, and Conservation Sciences, Oregon State University, Corvallis, OR 97331, USA

**How changes in biodiversity affect disease, particularly in the face of large-scale land-use change, is a contentious topic in disease ecology that has implications for public health and conservation policy. The ‘dilution effect’ hypothesis argues that declines in biodiversity are associated with increased disease risk, but this can be challenging to demonstrate because many pathogens have complex life cycles such that changes to the species composition and abundance of hosts can influence the density and infection prevalence of vectors via multiple mechanisms. Key to addressing this debate is a quantification of interactions between hosts, vectors, and pathogens. In their recent study published in *Molecular Ecology*, Kocher et al. (2022) captured thousands of sandflies, some species of which are vectors for the *Leishmania* protozoan that causes Leishmaniasis, across a human footprint gradient in French Guiana (Fig. 1). By implementing DNA metabarcoding of vectors combined with an innovative modeling approach, they effectively quantified the nuanced relationships between changes in land-use, mammalian host diversity, vector abundance, and parasite prevalence. In support of the dilution effect hypothesis, Kocher et al. found that sites with higher mammal diversity**

were associated with lower relative abundance of reservoir hosts and higher *Leishmania* infection prevalence in sandflies. However, while infection prevalence was lower when mammal diversity was high, the density of sandfly vectors was higher, which resulted in a weak overall effect of mammal diversity on the density of infected vectors, the most important indicator of *Leishmania* transmission risk.

Debate over the dilution effect hypothesis in disease ecology (Levi et al., 2016; Wood et al., 2014) has yet to be settled in part due to the challenge of directly measuring interactions between hosts, vectors, and pathogens. Parallel to this debate has been the advent of DNA metabarcoding, which allows for the sequencing of taxonomically informative genetic regions from a bulk DNA sample. More recently, emerging research has demonstrated that DNA metabarcoding of invertebrate-derived environmental DNA (iDNA) from hematophagous insects, including sandflies, can be used to inform locally-present vertebrate diversity (Kocher et al., 2017; Massey et al., 2022). Kocher and coauthors capitalize on the power of DNA metabarcoding to address the dilution effect using sandfly vectors as a source of DNA from the vector (the sandfly), hosts (the wildlife the sandflies feed upon), and the parasites (the *Leishmania* protozoa contracted from a vertebrate host). Kocher et al. captured and sorted 18,508 female sandflies (Fig. 2) and 855 individual blood-fed dipterans, 755 of which were sandflies. To profile sand fly diversity and estimate the prevalence of *Leishmania* parasites from across this landscape, they sequenced pools of sandflies and found *Leishmania* DNA in 26.3% of their pools. Separately, they sequenced individual, blood-fed sandflies (Fig. 3) to link sandfly vector species with specific mammalian hosts. Sandflies fed upon a diverse suite of mammal species (n=28 including 11 known host species of *Leishmania*) with host-feeding preferences varying by the genus of sandfly. Typically, the inference from sequencing individual blood meals would end there, but in an important contribution to the field Kocher et al. went on to develop an analytical framework to extract much more information from this iDNA data.

As the practice of using DNA metabarcoding of iDNA for monitoring biodiversity increases, there is substantial need to develop analytical frameworks to pull greater inference from the data. Kocher et al. developed an innovative hierarchical Bayesian modeling framework for data containing species of vertebrate host nested within sandfly genus nested within a sampling site. Each blood-fed sandfly contains the DNA from a random draw of vertebrates at a site contingent on both the feeding preferences of that sandfly species and the relative abundance of hosts at that site. The key challenge is distinguishing between feeding preferences and host abundance. Kocher et al's innovation was to make feeding preferences and relative abundances distinguishable by sampling across vectors with distinct preferences (including leveraging information from dipterans other than sandflies) and across sites with distinct host communities. They recognized that the number of vector bloodmeals coming from each vertebrate host is a multinomial process resulting in a joint likelihood composed of the product of multinomial likelihoods across all vector and sites. The resulting posterior distribution would typically describe a vector of probabilities, but Kocher et al. made these probabilities a sensible function of relative host abundances and sandfly feeding preferences. The posteriors for relative host abundances could then be used to calculate the Shannon diversity index at each site directly from metabarcoding of iDNA, which could then be linked to the density and infection prevalence of sandflies. A really wonderful and powerful approach!

This is an exciting moment for combining statistical and molecular approaches to facilitate more efficient and effective biodiversity data collection in remote environments. Kocher et al. showed that vector-derived DNA metabarcoding can measure host, vector, and pathogen communities directly, and that with a large enough sample size these data can be used to statistically model species interactions, biodiversity, and landscape epidemiology. This framework should have great utility for future research both in landscape epidemiology and for the growing field of biodiversity monitoring using iDNA. Because identifying individually blood-fed sandflies under a dissecting microscope is extremely laborious, perhaps further advances in statistical modeling would make use of iDNA metabarcoding of sandfly pools rather than individual blood-fed vectors. Such an approach would substantially increase throughput and reduce costs but presents additional statistical challenges.

Kocher et al. used their methodological advance to make an important scientific contribution to our understanding of the dilution effect. Key assumptions underlying the dilution effect in this system are that (1) diversity declines with anthropogenic disturbance, (2) more competent hosts for *Leishmania* have higher relative abundance at degraded sites, and (3) the vectors are sufficiently generalist feeders that bloodmeals can be diverted away from competent hosts. Kocher et al. find support for all three of these assumptions. However, the overall risk of contracting *Leishmania* also depends on the abundance of sandfly vectors, which they found to increase with mammal diversity counter to the dilution effect hypothesis. Vector abundance can increase with host diversity either because sandflies are host-limited and higher diversity sites have higher host abundance, or because sandflies respond negatively to the changing environmental conditions in degraded habitats. We find the latter to be plausible because edge effects in degraded tropical forests increase desiccation, which may reduce the availability of humid habitats with decaying organic matter in which sandflies reproduce (Moncaz, Faiman, Kirstein, & Warburg, 2012). Thus land-use change may be associated with increased risk of contracting *Leishmania* if bitten by a sandfly (higher infection prevalence), but reduced risk of being bitten because sandflies are less common (lower abundance). The result was that diluting and amplifying forces combined to produce no apparent effect of mammal diversity on the density of infected sandflies. This acts as an important reminder that teasing apart mechanisms is critical to understanding the effects of biodiversity and land use change on disease risk.



Figure 1: One of the sampling sites in the Amazon rainforest, Counami, French Guiana. (credit: B. de Thoisy)



Figure 2: Light trap used for the collection of phlebotomine sand flies (credit: A. Kocher)



Figure 3: Content of a light trap viewed with a stereo microscope where blood-fed sandflies can be individually sorted. Under the forceps: a phlebotomine sand fly. (credit: A. Kocher)

## References

Kocher, A., de Thoisy, B., Catzeffis, F., Valière, S., Bañuls, A. L., & Murienne, J. (2017). iDNA screening: Disease vectors as vertebrate samplers. *Molecular ecology*, *26* (22), 6478-6486.

Levi, T., Massey, A. L., Holt, R. D., Keesing, F., Ostfeld, R. S., & Peres, C. A. (2016). Does biodiversity protect humans against infectious disease? Comment. *Ecology*, *97* (2), 536-542.

Massey, A. L., Bronzoni, R. V. d. M., da Silva, D. J. F., Allen, J. M., de Lázari, P. R., dos Santos-Filho, M., . . . Levi, T. (2022). Invertebrates for vertebrate biodiversity monitoring: Comparisons using three insect taxa as iDNA samplers. *Molecular Ecology Resources*, *22* (3), 962-977.

Moncaz, A., Faiman, R., Kirstein, O., & Warburg, A. (2012). Breeding sites of *Phlebotomus sergenti*, the sand fly vector of cutaneous leishmaniasis in the Judean Desert. *PLOS Neglected Tropical Diseases*, *6* (7), e1725.

Wood, C. L., Lafferty, K. D., Deleo, G., Young, H. S., Hudson, P. J., & Kuris, A. M. (2014). Does biodiversity protect humans against infectious disease? *Ecology*, *95* (4), 817-832.