Deforestation-driven food-web collapse linked to emerging tropical infectious disease, Mycobacterium ulcerans

Aaron Morris, Jean-François Guégan, Demetra Andreou, Laurent Marsollier, Kevin Carolan, Marie Le Croller, Daniel Sanhueza, Rodolphe Gozlan

To cite this version:
Aaron Morris, Jean-François Guégan, Demetra Andreou, Laurent Marsollier, Kevin Carolan, et al.. Deforestation-driven food-web collapse linked to emerging tropical infectious disease, Mycobacterium ulcerans. Science Advances, American Association for the Advancement of Science (AAAS), 2016, 2 (12), pp.e1600387. 10.1126/sciadv.1600387. inserm-01421033

HAL Id: inserm-01421033
https://www.hal.inserm.fr/inserm-01421033
Submitted on 21 Dec 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Deforestation-driven food-web collapse linked to emerging tropical infectious disease, *Mycobacterium ulcerans*

Aaron L. Morris,1,2* Jean-François Guégan,2,3 Demetra Andreou,1 Laurent Marsollier,4 Kevin Carolan,2 Marie Le Croller,2 Daniel Sanhueza,2,4 Rodolphe E. Gozlan1,2,5

*Corresponding author. Email: almorris@rvc.ac.uk

Generalist microorganisms are the agents of many emerging infectious diseases (EIDs), but their natural life cycles are difficult to predict due to the multiplicity of potential hosts and environmental reservoirs. Among 250 known human EIDs, many have been traced to tropical rain forests and specifically freshwater aquatic systems, which act as an interface between microbe-rich sediments or substrates and terrestrial habitats. Along with the rapid urbanization of developing countries, population encroachment, deforestation, and land-use modifications are expected to increase the risk of EID outbreaks. We show that the freshwater food-web collapse driven by land-use change has a nonlinear effect on the abundance of preferential hosts of a generalist bacterial pathogen, *Mycobacterium ulcerans*. This leads to an increase of the pathogen within systems at certain levels of environmental disturbance. The complex link between aquatic, terrestrial, and EID processes highlights the potential importance of species community composition and structure and species life history traits in disease risk estimation and mapping. Mechanisms such as the one shown here are also central in predicting how human-induced environmental change, for example, deforestation and changes in land use, may drive emergence.

INTRODUCTION

Despite a large percentage of the approximately 250 known human emerging infectious diseases (EIDs) that have been traced to tropical forests (1), precise mechanistic details of the association often remain elusive. A consequence of the rapid increase in human populations, settlements, and encroachments globally is the concomitant rapid decline in biological diversity, with significant shifts in species community composition and severe disruption to established food webs, which are directly tractable to land-use change and deforestation (2–4). The results include advanced effects on ecosystem functioning and the potential for marked changes in associated microbial dynamics (5–8), including impact on resident human populations.

Estimates suggest that 10,000 to 20,000 freshwater species are under severe threat of extinction (9–12). Riparian systems act as an interface between the environment and hosts and/or vectors of many human disease agents. Currently, the understanding of the consequences of changes in freshwater communities on the incidence and epidemiology of human EIDs is limited, and empirical evidence is still scarce (6, 13–15). The effects of deforestation and land use on potential aquatic infectious agent hosts have been described predominantly for established vector-transmitted diseases, notably malaria (16). These are clearly of interest because of the large number of human cases and known vectors or hosts that can be more readily surveyed and the data used as proxy for potential infection rates.

To potentially understand the more generalist EIDs, which often do not rely on a single specific host or vector species (17), extensive analysis of shifts in host/nonhost communities caused by natural habitat modifications and/or land-use change is required. For example, hosts of differing suitability will be dissimilarly adapted to modified environments; therefore, it is critical to quantify their relationship with an EID. In addition, land use or habitat change often results in a shift in trophic interactions, which can be a direct (for example, predation) or indirect (for example, extinction cascade) modification of the food-web structure. A severe change in a trophic network may have a significant impact on an EID, with promotion or decline of certain host species in the community network.

Here, we tested for the first time the links between the drivers of trophic community networks in tropical freshwater ecosystems and the emergence of a generalist microbial agent, *Mycobacterium ulcerans*, which causes the human skin disease, Buruli ulcer. A detailed structural presentation of the host and nonhost interaction under varying levels of anthropogenic stress related to *M. ulcerans* will be provided. The empirical relatedness of the structural changes in the host and nonhost community networks to the biomass of *M. ulcerans* will also be characterized.

RESULTS

More than 3600 invertebrates and fish were collected and cataloged from 17 freshwater sites across French Guiana, a French territory endemic for Buruli ulcer. A total of 78 different freshwater taxa were identified, of which 383 specimens representing 44 taxa tested positive for both IS2404 and KR (ketoreductase B) molecular markers (Table 1). *M. ulcerans* concentrations ranged from 6 to 10,385 bacillus cells per milligram of organism (mean, 1405; SD, ±2109). The highest concentrations of *M. ulcerans* were in species lower in the food chain, represented by low δ13N and δ13C stable isotope readings, suggesting a diet high in aquatic algae, detritus, diatoms, and similar food sources (Fig. 1). These taxa, in the upper quartile of the *M. ulcerans* load, were predominantly invertebrates in both adult and larval form, including Corixidae, Caenidae, Coenagrionidae, Baetidae, Velidae, Leptophlebiidae, and Simulidae, and two freshwater fish species, *Copella carsevenensis* and *Rivulus lungi*.
Deforestation and either agricultural or urban intrusion led to a decline in local trophic niche width (metric of total available niche space within a site) (adjusted $R^2 = 0.557$; deviance explained = 62.2%; deforestation: $P = 0.0008$, $F = 17.719$; agriculture/urban land use: $P = 0.0265$, $F = 6.228$) (Fig. 2A), resulting in decreased regional mean vulnerability of taxa (number of potential predator species each prey taxa has), coupled with a decrease in generality (number of potential prey species each predator taxa can consume) (vulnerability: adjusted $R^2 = 0.335$, deviance explained = 35.8%, $F = 20.67$, $P < 0.0001$; generality: adjusted $R^2 = 0.528$, deviance explained = 54.3%, $F = 29.31$, $P < 0.0001$) (Fig. 2B). This had an important, nonlinear effect on the potential local $M. ulcerans$ load. Host taxa, which on average carried a high level of $M. ulcerans$, were most abundant at sites where there was a very low level of vulnerability and a midlevel of generality (adjusted $R^2 = 0.482$; deviance explained = 36%; generality: $P = 0.0003$, $F = 7.303$; vulnerability: $P < 0.0001$, $F = 18.871$) (Fig. 2C).

### Table 1. Primers and probes for real-time PCR detection

Nucleotide position based on the first copy of the amplicon in pMUM001 (accession numbers: IS2404, AF003002; KR, BX649209). TF, forward primer; TR, reverse primer; TP, probe.

<table>
<thead>
<tr>
<th>Primer or probe</th>
<th>Sequence (5′-3′)</th>
<th>Nucleotide positions</th>
<th>Amplicon size (base pairs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IS2404 TF</td>
<td>AAAGCACCACGCGCATCT</td>
<td>27746–27762</td>
<td>59</td>
</tr>
<tr>
<td>IS2404 TR</td>
<td>AGCGACCCCGTGAGGATTG</td>
<td>27787–27804</td>
<td></td>
</tr>
<tr>
<td>IS2404 TP</td>
<td>6 FAM-CGTCCAACGCAGTG-MGBNFQ</td>
<td>27768–27781</td>
<td></td>
</tr>
<tr>
<td>KR TF</td>
<td>TCACCGCTTGATATGATCA</td>
<td>3178–3195</td>
<td>65</td>
</tr>
<tr>
<td>KR TR</td>
<td>TTGTGTGGGCACTGACGTTGAC</td>
<td>3222–3242</td>
<td></td>
</tr>
<tr>
<td>KR TP</td>
<td>6 FAM-ACCCCGAAGCACTG-MGBNFQ</td>
<td>3199–3212</td>
<td></td>
</tr>
</tbody>
</table>

**Fig. 1.** The average $\delta^{15}N$ and a low $\delta^{13}C$ in biplot space for all recorded host and nonhost organisms from the 17 sites. $\delta^{15}N$ and $\delta^{13}C$ were derived for each host taxon using data from the three sites analyzed for stable isotopic readings and extrapolated to the additional 14 sites. For each point, the square root–transformed mean number of $M. ulcerans$ of the host is represented by the size of the circle.
The characterization of a relationship between the abundance of a bacterial pathogen, host species community structure, and land cover provides key information on the ecological pathways of an infectious agent’s emergence. Understanding such pathways is central to modeling the impact of global changes on EIDs in tropical areas and predicting infectious risk in developing countries.

The nonlinear increase in the potential abundance of *M. ulcerans* under the anthropogenic stresses of land-use change and deforestation can, in this instance, be attributed to an increase of preferred hosts [identified by the regional mean *M. ulcerans* load attained using quantitative polymerase chain reaction (qPCR)], represented by predominantly low–trophic level organisms (Fig. 2). At sites of low vulnerability, some basal host species flourish because of a decrease or disappearance of predation. However, if a site has a high level of generality, the number of basal host species is highly variable with a functionally more restrictive or fluctuating niche space. As the diversity of basal organisms begins to decrease (indicated by a loss in generality), the abundance of *M. ulcerans* hosts and thus the *M. ulcerans* load then start to increase, until the number of basal organisms becomes too low, leading again to a decrease in *M. ulcerans*.

As urbanization, agriculture, and deforestation intensify, notably in tropical regions of the world (which also hosts the largest number of developing countries on the planet), similar trends for other EIDs, as observed here for *M. ulcerans*, may become apparent. One hypothesis is that basal opportunistic and highly fecund species (that is, *r*-strategists) within a trophic food web could form a hotspot for many bacterial or viral pathogens. In both aquatic and terrestrial environments, many of the basal taxa live at the interfaces of soil and water or soil and vegetation, and their high numbers and potential reduced investment in sustainable land management practices could result in increased risks for EID emergence.
immunity (18–21) allow for many opportunities for emerging microbes to take hold or opportunistic microbes to make the leap from commensalism or symbiosis to pathogenesis (22). Although this hypothesis is plausible for some species, evidence within the literature is still debated and more studies in this field are required to prove a definitive link (13, 14).

Because human populations have greater contact with land-use change generated by global urbanization, it is not unreasonable to predict that emergence and reemergence of infectious diseases will become more common. We therefore recommend that more empirical evidence must be sought for mechanisms related to disease emergence in areas of anthropogenic disturbance, notably in tropical regions of the world where biodiversity (at both micro- and macroscales) can be high. Although this study identifies a potential mechanism for a specific (albeit important) emerging disease, the generalizability of these results needs further testing. The relationships EIDs have with other organisms and the environment are complex and are dictated by a multitude of factors, which must be individually explored. In doing so, we will increase our chances of predicting and understanding the risk posed by these pathogens.

MATERIALS AND METHODS
Study location
French Guiana is an extraperipheral territory of France, bordering Brazil to the east and south and Suriname to the west. Although large in area at 83,534 km², the population density is relatively low, with almost all inhabitants located in a thin strip along the coastline. The rest of the country is predominantly pristine tropical rain forest including 33,900 km² of the Guiana Amazonian National Park and a wealth of important ecosystems ranging from marshland to coastal mangroves. Since 1969, the number of Buruli ulcer cases has remained steady but unusually high in comparison to that of other countries in the surrounding region (25, 29). The disease is endemic in close to 30 tropical and subtropical countries, where it can affect health and lead to serious socioeconomic consequences (30).

Data collection
Seventeen locations with freshwater sites in French Guiana suitable for study were identified as endemic or having the potential to harbor M. ulcerans according to previously published data (29). These sites are empirically similar but have gradients of land use from urban through agricultural to pristine rain forest. Each site was surveyed in the rainy season twice, during February and June 2013. Samples of aquatic invertebrates, fish, and other small aquatic vertebrates were collected and analyzed as follows.

Invertebrate samples were collected at four locations, spatially distributed within each site to include a range of microhabitats. Nets were used to vigorously collect samples from the bottom sediment and the surface of the water column within a meter square area. All collected samples were placed on a white tray for preliminary sorting in the field and fixed in 70% ethanol for subsequent identification and storage in the laboratory.

Fish were collected using three standard baited fish traps and spatially distributed as the dip net locations (see above). Traps were left in place for 1.5 hours before retrieval. Fish were collected and euthanized using a 400-mg to 1-liter solution of clove oil. The solution was added gradually, and fish remained submerged for 10 min until death, following the guidelines set out by the American Veterinary Medical Association as advocated by the Institutional Animal Care and Use Committee (31). Fish were then preserved in 70% alcohol for sorting and identification back at the laboratory. Fish were collected under a U.K. personal license for regulated animal procedures held by R. E. Gozlan and adhered to the French and European laws. The project was approved by the Bournemouth University Ethics Committee and was performed under the Home Office project license number PPL80/1979.

Identification and data collection
The length of each invertebrate measured, and identification was conducted predominantly to family level where possible, using a number of literature sources (32–36). Specimens of the same family were pooled by site and collection date and preserved in 70% alcohol for qPCR analysis. Invertebrates containing more than 10 individuals were pooled in series of 10. Dry weights in milligrams were calculated using length to mass regression equations (37–39). Fish were identified to species level (40, 41), and for each individual, a standardized length measurement was taken. Fish weights were also calculated using length to weight regression equations (42); however, because there is no direct conversion available in the literature for fish to establish dry weights for all species, an average wet to dry weight conversion of 20% was used. For qPCR, the fish were pooled in groups of three individuals per species per sample (dip net or trap), an important step for the analysis of trapped fish because they are likely to congregate in very large numbers within a trap. Therefore, this threshold was applied for the abundance to be relative to a small sample area for invertebrates.

Quantitative PCR
Pooled organisms were homogenized in 50 mM NaOH solution using the Qiagen TissueLyser II. Tissues were heated at 95°C for 20 min. DNA from the homogenized tissue was purified using the QIAquick 96 PCR Purification Kit (Qiagen) according to the manufacturer’s recommendations. The 10% negative controls were included for extraction and purification.

Oligonucleotide primer and TaqMan probe sequences were selected from GenBank IS2404 sequences (43), a more generic marker for mycolactone-producing mycobacteria and KR domain of the mycolactone polyketide synthase (mks) gene from the plasmid pMUM001 (Table 1) (44), which is highly specific to M. ulcerans. qPCR mixtures contained 5 µl of template DNA, 0.3 µM concentration of each primer, 0.25 µM concentration of the probe, and Brilliant II QPCR Master Mix with Low ROX (Agilent Technologies) in a total volume of 25 µl. Amplification and detection were performed with Thermocycler (Chromo4, Bio-Rad) using the following program: 1 cycle at 50°C for 2 min, 1 cycle at 95°C for 15 min, 40 cycles at 95°C for 15 s and at 60°C for 1 min. DNA extracts were tested in duplicate, and the 10% negative controls were included in each assay. Quantitative readout assays were established on the basis of an external standard curve with M. ulcerans (strain 1G897) DNA serially diluted over 5 logs (from 10⁶ to 10² U/ml). Samples were considered positive only if both the gene sequence encoding the KR domain of the mks and IS2404 sequence were detected, with threshold cycle (Ct) values strictly <35 cycles [see the studies of Garchitorena et al. (26) and Morris et al. (29)]. The number of bacilli in the sample was estimated from the PCR data of KR because it is more specific to M. ulcerans.

Stable isotope analysis
Stable isotope analysis was conducted on all invertebrates and vertebrates from three sites for both collection dates for the stable isotope
signatures of $^{13}$C and $^{15}$N. This covered the majority of the taxa present at all 17 sites and was applied to the same taxa from other sites (table S1). For invertebrate analysis, a limb of the organism was used, or if it was too small and many were present, whole organisms were used. For fish, caudal fin clippings were used. A minimum amount of material was taken to preserve organisms for qPCR analysis. In most cases, three samples of each taxonomic group were analyzed for each site. However, because of the limited number of available organisms one or two samples were used in some cases. For any taxa that were not analyzed due to abundance limitations, stable isotope data from a closely related and functionally highly similar taxon were recorded for use in the statistical analysis. The material was dried at 40°C for 8 hours to remove all traces of moisture. After drying, the material was homogenized, and the stable isotope signatures of $^{13}$C and $^{15}$N were identified using a Finnigan MAT Delta Plus isotope ratio mass spectrometer.

**Trophic niche width and food-web metric calculations of sites and aquatic invertebrates and vertebrates**

Trophic niche width (a metric of total available niche space within a site) (45) was calculated from the stable isotope data as the total convex hull area of the mean $^{13}$C and $^{15}$N isotopic readings in biplot space for the entire surveyed population of all organisms within a site (fig. S1). As a measure of available niche space, niche width can be seen as a useful metric for quantifying how capable a community is at exploiting its current environment, something that will be directly affected by land use and/or deforestation.

The allometric diet breadth model (ADBM) was used to predict links between nodes in each food web, and these were subsequently tested for accuracy from previous knowledge. Any biologically inaccurate or missing links were updated accordingly. The ADBM is based on the mass of an organism and, in some cases, assumes a larger organism to prey on a smaller one. However, without providing biological and behavioral information for all taxa, in some instances, this is not precise. Therefore, all food webs must be checked for missing or inaccurate links.

Once a network was estimated, a number of average local metrics were calculated for each site (table S2). This included the connectance (that is, a measure of the number of node connections in a food web) or link density, the generality (a measure of the average number of different organisms a taxon can feed on), and the vulnerability (a measure of the average number of species by which an organism is preyed upon).

**Landscape data**

Landscape data were extracted from two maps. The first was the 25-m-resolution French Ministère de l’Écologie, du Développement Durable et de l’Énergie CORINE Land Cover 2006 based on the CORINE 2006 European land cover map using satellite-derived data from 2005 and 2008. The second comprised the Hansen deforestation maps, which consists of a high-resolution map of global forest change to 30-m resolution, with individual data for each year from 2001 to 2012 (46). Each site location was plotted on the land cover maps, and a 1-km buffer zone was drawn around each. From within the buffer zone, land cover data in square meters were extracted for the CORINE map and deforestation data for the years 2010 to 2012 for the Hansen’s map (table S3). Data extraction was conducted in Quantum GIS using the LecoS landscape data analysis plug-in (47). For two sites, which were considered “pristine” and located outside the small villages further into the forest, the CORINE Land Cover map did not cover the area. Therefore, data from the GlobCover 2009 Global Land Cover Map were used (48). This map, although lower resolution for French Guiana, was used only for two sites, which were buffered by almost entirely forest. The retrieved data were considered sufficient for comparisons to be made.

**Analysis**

For each taxon identified during the surveys, the regional mean $M.\ ulcerans$ load was calculated from the qPCR data, encompassing all sites, survey dips, and dates, that is, representing a taxonomic group average bacterial load if found in the wild, based on information gathered for each site. In addition, each taxon’s regional mean niche width, vulnerability, connectance, and generality were also derived and weighted by the abundance at each site (that is, if a species occurred predominantly at sites with low generality, its average generality metric would be lower).

To identify whether those host taxa with a higher mean $M.\ ulcerans$ load were most common at sites with a certain type of food-web structure (for example, high levels of connectance, generality, or vulnerability), general additive models (GAMs) were run, with a tensor product smoothing performed between the regional $M.\ ulcerans$ load and their mean regional food-web metrics (connectance, vulnerability, and generality). Nonsignificant metrics were removed until a final best-fit model was found. Because many taxa had no $M.\ ulcerans$ across all sites, a Tweedie’s probability distribution was used, as this distribution is well suited to zero-inflated models and provided the best-fit to the data (49). To identify how generality, connectance, and vulnerability responded to a reduction in niche width of the population, GAMs of each metric (that is, vulnerability, generality, and connectance as a function of niche width) were run separately.

The relationships between land cover data and niche width were identified at the local site level using GAMs of the level of deforestation in square meters within each buffer zone in addition to urban and agricultural cover (square meters). For deforestation, data from the preceding 3 years of the survey date were used; this was done to allow for enough data to be available for each site but not to be too far from the survey as to be unlikely to have an effect. This period was also 1 year preceding 2008, when the last SPOT-5 satellite data for the agricultural and urban land cover were collected, so it is possible to be confident that this land type was present when deforestation occurred.

In all models, the data were appropriately transformed to adhere to model assumptions, and the residuals of each model were tested for spatial autocorrelation using both Moran’s $I$ test (95% confidence interval) and CorreLog smooths with a spline function (95% point-wise bootstrap confidence intervals) (50). Further model checking was performed using $q-q$ plots and residual plots against fitted values, with any significant outliers removed. This constituted the following: Polymictarciidae, a family of mayflies with only one individual found from all sites, with a very high level of $M.\ ulcerans$ (10,385 per milligram), a possible PCR anomaly; a single individual of the order Trichoptera found with very low regional convex hull measurements; Cleithracara maronii, a species of fish, also a single individual found with a very low regional convex hull measurement; and a single sample of Moenkhausia hemigrammnomides from a single site found to be a significant outlier when model checking.

**SUPPLEMENTARY MATERIALS**

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/2/12/e1600387/DC1

fig. S1. The average $^{15}$N and a low $^{13}$C in biplot space for each taxonomic group at each of the 17 sites.
REFERENCES AND NOTES


Acknowledgments: We would like to thank V. O. Ezenwa and R. Colwell for their comments and suggestions on an earlier draft of the manuscript. We would also like to thank O. Petchey
for use of his R script for calculating the ADBM. **Funding:** This work has benefited from a 3-year Bournemouth University PhD fellowship grant to A.L.M. A.L.M., J.-F.G., M.L.C., K.C., D.S., and R.E.G. have benefited from an “Investissement d’Avenir” grant managed by Agence Nationale de la Recherche (CEBA; reference ANR-10-LABX-2501). J.-F.G. is sponsored by the Institut de Recherche pour le Développement (IRD) and the Centre National de la Recherche Scientifique, and R.E.G. is sponsored by IRD. L.M. and J.-F.G. are funded by an Equipe Inserm Avenir “ATOMycA” delivered to L.M. D.S. was funded by BECAS CHILE (CONICYT PAI/INDUSTRIA 79090016). The content is solely the responsibility of the authors and does not necessarily represent the official views of the ANR or ATIPe Avenir INSERM. **Author contributions:** A.L.M., R.E.G., and J.-F.G. designed the experiment. A.L.M., K.C., and M.L.C. carried out data collection. A.L.M., L.M., and D.S. carried out genetic analysis. A.L.M. performed the statistical analysis, A.L.M., J.-F.G. and R.E.G. wrote the paper with contributions from all authors. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Submitted 23 February 2016
Accepted 10 November 2016
Published 7 December 2016
10.1126/sciadv.1600387