Identifying Biological and Environmental Indicators of Emerging Infectious Diseases: The case of Buruli Ulcer

Morris. A. L.

This thesis has been submitted in partial fulfillment of the requirements of the degree of Doctor of Philosophy

Bournemouth University in collaboration with the Institut de recherche pour le développement (IRD)

October 2014

Abstract

Understanding disease ecology is vital in preventing future outbreaks of established infections and to predict the emergence of new pathogens. In recent decades there have been a number of high profile infectious diseases which have swept across countries and in some cases the world. Many of these begin as generalist emerging infections; such microbes are difficult to study in the wild due to their inherently ambiguous life histories and complex associations with numerous hosts and the environment. In this PhD a number of techniques are used to pinpoint and further understand the life history of one such pathogen *Mycobacterium ulcerans*, the causative agent of Buruli ulcer, in the hope that this data can be used to predict and prevent future outbreaks and can be applied to other emerging infections. The results of this study include the first identification of the pathogen in the environment for a whole new continent, South America. Further to this it has led to the discovery of the likely ecological niche of the bacilli by linking its presence to specific functional groups of organisms. In turn the occurrences of these groups have been related to anthropogenic conditions such as deforestation and human mediated land use. Finally complex links between climatic fluctuations and outbreaks of the disease in Southern America and Cameroon, central Africa help complete our understanding of this mysterious disease.

Table of contents

Contents

Abstract i

Table of contents ii

List of Figures vi

List of tables xi

Acknowledgements xiii

Declaration xiii

Chapter 1 Literature Review 1

1.1 Emerging infectious diseases 1

1.1.1 EID's and the environment 1

1.1.2 EID's and biodiversity 2

1.2 Objectives 5

1.3 Buruli ulcer 6

1.3.1 Epidemiology 8

1.3.1.1 Africa 8

1.3.1.2 Australia 9

1.3.1.3 Central and South America 9

1.3.2 Biology 10

1.3.3 MU evolution 11

1.3.4 Pathology 12

1.3.5 Ecology 13

1.3.6 Human impacts and landscape induced change on the infection rates of BU 18

1.4 Main country of study: French Guiana 20

1.5 References 21

Chapter 2 Site selection, field methodologies and site descriptions for French Guiana 31

2. 1 Initial survey to preliminary identify suitable sites for study 31

2.1.1 Environmental sample collection for sites 1-30 31

2.1.2 DNA extraction. 32

2.1.3 Non-quantitative PCR 32
2.2 Second survey 32
2.2.1 Site identification 32
2.2.2 qPCR protocols 34
2.3 Biotic and abiotic surveys of final analysis sites 36
2.3.1 Field method for sampling vertebrates and invertebrates 36
2.3.2 Abiotic readings 37
2.3.3 Organism identification cataloguing and measurement 37
2.3.4 Stable isotope analysis 38
2.3.5 DNA extraction for invertebrates and vertebrates 38
2.4 References 43

Chapter 3 First detection of *Mycobacterium ulcerans* DNA in environmental samples from South America 44

3.1 Summary 44
3.2 Introduction 45
3.3 Materials and Methods 46
3.3.1 Site identification and sampling methods 46
3.3.2 Statistical Analysis 46
3.3.2.1 Number of cases since the building of Petit-Saut dam 46
3.3.2.2 Differences between abiotic parameters at sites 48
3.4 Results 48
3.4.1 qPCR 48
3.4.2 Number of cases since the building of Petit-Saut Dam 52
3.5 Discussion 52
3.6 References 54

Chapter 4 Complex temporal climate signals drive the emergence of human water-borne disease. 58

4.1 Summary 58
4.2 Introduction 59
4.3 Material and methods 60
4.3.1 Environmental and disease data 60
4.3.1.1 French Guiana 60

4.3.1.2 Akonolinga 61
4.3.2 Ethical provisions 63
4.3.3 Statistical analysis 63
4.3.3.1 Singular spectrum analysis decomposition and reconstruction 63
4.3.3.2 Empirical Mode Decomposition 65
4.3.3.3 Trend coherence analysis 65
4.3.3.4 Seasonal coherence analysis 66
4.4 Results 67
4.4.1 French Guiana 67
4.4.1.1 SSA 67
4.4.1.2 Empirical mode decomposition (EMD) 71
4.4.1.3 Trend coherence analysis 71
4.4.1.4 Seasonal coherence analysis 74
4.4.2 Akonolinga 78
4.4.2.1 SSA 78
4.4.2.2 Empirical mode decomposition (EMD) 80
4.4.2.3 Seasonal coherence analysis 82
4.5 Discussion 85
4.6 References 90

Chapter 5 Functional diversity provides a new perspective for understanding the ecology of emerging infectious diseases 98

5.1 Summary 98
5.2 Introduction 99
5.3 Methods 103
5.3.1 Data collection 103
5.3.2 PCR analysis 103
5.3.3 Functional group identification 103
5.3.4 Analysis 104
5.4 Results 107
5.4.1 Indices of functional diversity 107
5.4.2 Functional group analysis 107
5.5 Discussion 110

5.6 References 114

Chapter 6 Land-use/deforestation-mediated food web collapse promotes infectious disease carrying species in tropical freshwater systems. 125

6.1 Summary 125

6.2 Introduction 126

6.3 Methods 128

6.3.1 Data collection 128

6.3.2 Landscape data 128

6.3.4 Trophic niche width and food web metric calculations of sites and aquatic invertebrates and vertebrates. 129

6.3.7 Biodiversity measures 130

6.3.8 Analysis 130

6.4 Results 132

6.5 Discussion 139

6.6 References 143

Chapter 7 Thesis discussion 151

7.1 Overview 151

7.2 Relating to other EID's 153

7.3 Potential for future disease in French Guiana 154

7.4 Future work 156

7.5 Conclusions 157

7.6 References 159

Appendices 162

Appendix I 162

Appendix II 171

Appendix III 185

Appendix IV 190

Appendix V 196

List of Figures

Figure 1.1 Global map showing countries reporting cases of BU (Merritt et al. 2010) 7

Figure 1.2 Cases of BU from 1969 to 2012 per 100,000 people, the dotted line represents the increasing population. 10

Figure 2.1 Map showing the 53 sampling locations. In green are the sites negative to both primers for MU, in yellow are the sites positive for IS2404 alone and in red are the sites positive for both IS2404 and KR. 35

Figure 2.2 Map of French Guiana showing the location of all 17 analysis sites where detailed surveys were undertaken. 36

Figure 3.1 Map showing the approximate location of the twenty-three sampling sites (Text S1). In green are the negative sites for both MU during this survey, in yellow are the site positive for IS2404 alone and in red are the sites positive to both IS2404 and KR. The Petit-Saut dam built in 1994 is indicated by a star (Lehner et al. 2006). 47

Figure 4.1 Map of French Guiana showing the location of 17 weather stations along the coast of French Guiana, including one station located on an off-shore island and the position of French Guiana within South America. 62

Figure 4.2 French Guiana, monthly time series data showing the decomposition, reconstruction and forecasting of datapoints using Singular Spectrum Analysis. A) Original rainfall time series average for 17 weather stations along the coast of French Guiana, from 1969 to 2012. B) Periodograms of the rainfall time series identifying significant repeating patterns once per year, twice per year and three times per year. SSA extracted component corresponding to periodogram spike of C) Three times per year, four month component, D) Twice per year, six month component, E) Once per year, twelve month component. F) The extracted rainfall trend. G) The reconstructed rainfall time series after the removal of stochastic noise. H) A second periodogram of the reconstructed rainfall series showing less stochastic noise around the 3 main repeating patterns. I) Forecasting of the rainfall trend to 2017 using sequential SSA. 68

Figure 4.3 French Guiana, three year period of extracted components (dashed line) plotted against the same period of the reconstructed rainfall (solid line), showing which parts of the rainfall series the components represent. A) Four month component corresponds to both rainy seasons and the dry season. C) Six month component corresponds to the two rainy seasons. C) Twelve month component represents the rainfall for the whole year. 69

Figure 4.4 French Guiana, ω-correlation matrix, the F values represent oscillating components within a year (i.e. F6 is the 6 month bi-annual component). The level of correlation can be found by finding the component of interest along the x-axis and looking up the y-axis to see where it corresponds with other rainfall components. Large values of ω-correlation between reconstructed components indicate that they should possibly be gathered into one group and correspond to the same component in SSA decomposition. The matrix uses a 20-grade grey scale from white to black corresponding to the absolute values of correlations from 0 to 1 (with 0 being no correlation and 1 being absolute correlation). 70

Figure 4.5 French Guiana, monthly time series and Empirical Mode Decomposition of BU cases per 100,000 people. A) Monthly time series of BU cases per 100,000 people from 1969 to 2012. B) First Intrinsic Mode Function (IMF). C) Second IMF. D) Third IMF. E) Fourth IMF. F) Fifth IMF. G) Sixth IMF. H) Seventh IMF. Periodograms for I) first IMF showing a high level of variation across the spectra and therefore should be considered white noise. J) Second IMF which has its highest power at two cycles per year. K) Third IMF with its highest power at one cycle per year. L) Fourth IMF with highest power approximately at one cycle every two years. M) Fifth IMF with highest power approximately at one cycle every 4 years. N) Sixth IMF with a low level of cycles representing very long-term trends. O) Seventh IMF also representing long-term trends. 72

Figure 4.6 French Guiana, wavelet coherence between A) BU incidences per 100,000 people and the rainfall trend obtained from Singular Spectrum Analysis (SSA). B) BU incidences per 100,000 people and El Niño-Southern Oscillation (ENSO). The colours are coded from dark blue to dark red with dark blue representing low coherence through to high coherence with dark red. The solid black lines around areas of red show the α=5% significance levels computed based on 2000 Monte Carlo randomisations. The dotted white lines represent the cone of influence; outside this area coherence is not considered as it may be influenced by edge effects. The black arrows represent the phase analysis and adhere to the following pattern: Arrows pointing to the right mean that rainfall and cases are in phase, arrows pointing to the left mean that they are in anti-phase., arrows pointing up mean that cases lead rainfall and arrows pointing down mean that rainfall leads cases. 73

Figure 4.7 French Guiana, cross correlation analysis between A) El Niño-Southern Oscillation (ENSO) and reconstructed rainfall time series from Singular Spectrum Anlaysis. B) ENSO and original rainfall time series. ENSO and C) First Intrinsic Mode Function (IMF) from BU cases. D) Second IMF. E) Third IMF. F) Fourth IMF. G) Fifth IMF. Reconstructed rainfall series and H) First IMF. I) Second IMF. J) Third IMF. K) Fourth IMF. L) Fifth IMF. Dashed horizontal blue lines in all panels represent the 95% confidence limit; black vertical lines which go beyond the dashed line can be considered non-random cohering oscillations between the two time series being assessed, with the lag period between an above average oscillation in the first

time series and a subsequent above average oscillation in the second shown on the x-axis. 75

Figure 4.8 French Guiana, cross correlation analysis between rainfall seasonal components extracted by Singular Spectrum Analysis and seasonal Intrinsic Mode Function (IMF) series of BU. Four month component and A) First IMF. B) Second IMF. C) Third IMF. Six month component and D) First IMF. E) Second IMF. F) Third IMF. Twelve month component and G) First IMF. H) Second IMF.I) Third IMF. Dashed horizontal blue lines in all panels represent the 95% confidence limit; black vertical lines which go beyond the dashed line can be considered non-random cohering oscillations between the two time series being assessed, with the lag period between an above average oscillation in the first time series and a subsequent above average oscillation in the second shown on the x-axis. 77

Figure 4.9 Akonolinga, monthly time series data showing the decomposition, reconstruction and forecasting of datapoints using Singular Spectrum Analysis. A) Original rainfall time series average for 17 weather stations along the coast of French Guiana, from 1969 to 2012. B) Periodograms of the rainfall time series identifying significant repeating patterns once per year, twice per year and three times per year. SSA extracted component corresponding to periodogram spike of C) Three times per year, four month component, D) Twice per year, six month component, E) Once per year, twelve month component. F) The extracted rainfall trend. G) The reconstructed rainfall time series after the removal of stochastic noise. H) A second periodogram of the reconstructed rainfall series showing less stochastic noise around the 3 main repeating patterns. 79

Figure 4.10 Akonolinga, ω-correlation matrix, the F values represent oscillating components within a year (i.e. F6 is the 6 month bi-annual component). The level of correlation can be found by finding the component of interest along the x-axis and looking up the y-axis to see where it corresponds with other rainfall components. Large values of ω-correlation between reconstructed components indicate that they should possibly be gathered into one group and correspond to the same component in SSA decomposition. The matrix uses a 20-grade grey scale from white to black corresponding to the absolute values of correlations from 0 to 1 (with 0 being no correlation and 1 being absolute correlation). 80

Figure 4.11 Akonolinga, monthly time series and Empirical Mode Decomposition of BU cases per 100,000 people. A) Monthly time series of BU cases per 100,000 people from 1969 to 2012. B) First Intrinsic Mode Function (IMF). C) Second IMF. D) Third IMF. E) Fourth IMF. F) Fifth IMF. G) Sixth IMF. H) Seventh IMF. Periodograms for I) first IMF showing a high level of variation across the spectra and therefore should be considered white noise. J) Second IMF which has its highest power at two cycles per year. K) Third IMF with its highest power at one cycle per year. L) Fourth IMF with highest power approximately at one cycle every two years. M) Fifth IMF with highest power approximately at one cycle every 4

years. N) Sixth IMF with a low level of cycles representing very long-term trends. O) Seventh IMF also representing long-term trends. 81

Figure 4.12 Akonlinga, cross correlation analysis between A) El Niño-Southern Oscillation (ENSO) and reconstructed rainfall time series from Singular Spectrum Anlaysis. B) ENSO and original rainfall time series. ENSO and C) First Intrinsic Mode Function (IMF) from BU cases. D) Second IMF. E) Third IMF. F) Fourth IMF. G) Fifth IMF. Reconstructed rainfall series and H) First IMF. I) Second IMF. J) Third IMF. K) Fourth IMF. L) Fifth IMF. Dashed horizontal blue lines in all panels represent the 95% confidence limit; black vertical lines which go beyond the dashed line can be considered non-random cohering oscillations between the two time series being assessed, with the lag period between an above average oscillation in the first time series and a subsequent above average oscillation in the second shown on the x-axis. 83

Figure 4.13 Akonolinga**,** Cross correlation analysis between rainfall seasonal components extracted by Singular Spectrum Analysis and seasonal Intrinsic Mode Function (IMF) series of BU. Four month component and A) First IMF. B) Second IMF. C) Third IMF. Six month component and D) First IMF. E) Second IMF. F) Third IMF. Twelve month component and G) First IMF. H) Second IMF.I) Third IMF. Dashed horizontal blue lines in all panels represent the 95% confidence limit; black vertical lines which go beyond the dashed line can be considered non-random cohering oscillations between the two time series being assessed, with the lag period between an above average oscillation in the first time series and a subsequent above average oscillation in the second shown on the x-axis. 84

Figure 5.1 Cluster dendrogram of species based on their functional traits, dendrogram created using Gower dissimilarity; taxon groups on branches which are closer together have a higher number of functional traits in common. See Appendix II Table S5.4 for list of species, traits and their corresponding grouping 106

Figure 5.2 Model fitted values for the proportion of environmental samples, which were positive for MU DNA against the proportional presence of functional groups related to the breakdown and consumption of organic material in lentic habitats (group 10) and the filtering of organic material and sediment in lotic habitats (group 9). Dashed line represents line of best fit through the fitted points. See Appendix II Table S5.4 for list of species, traits and their corresponding grouping. 109

Figure 6.1 Bar plot showing the land cover data for each site, the top chart shows the m^2 cover of Agricultural and Urban land around each site within the 2km (black), 1km (dark grey) and 500m (light grey) buffer zones. The bottom chart shows the m^2 cover of deforestation within the 3 years prior to the surveys. 129

Figure 6.2 The average δ ^{15}N and a low δ^{13}C in bi-plot space for all recorded host and non-host organisms from the 17 sites. δ ^{15}N and δ^{13}C were derived for each taxa using the data from the 3 sites which were analysed for stable isotopic readings. For each point the square root transformed mean number of bacteria of that taxa is represented by the size of the circle. 134

Figure 6.3 Plot showing the metrics of web networks that allow taxa, which on average carry a higher bacterial load to propagate. For all taxa, along the bottom axes are their mean regional food web metrics for vulnerability and generality (i.e. a measure of the food web metrics of sites at which they are most abundant), whilst the vertical axis shows their mean regional bacterial load of MU. 135

Figure 6.4 Plot showing the results of GAM's between the mean regional niche width of all organisms and their mean regional food web metrics; vulnerability and generality. The plot shows that as niche width decreases, vulnerability and generality of the food webs also decreas 136

Figure 6.5 Plot showing the relationship between local site niche widths and the land cover in m^2 of agricultural and urban landscape and deforestation in a 1km buffer zone. The change to niche width caused by deforestation appears to be a steady decline, whilst the presence of agricultural land causes a sharp drop from where there is no agriculture, before exhibiting a similar steady decline in niche width as with deforestation, albeit less extreme. 138

Figure S5.1: K-Means partitions comparison using Simple Structure Index (SSI) [1] shows a likely optimal number of functional groups to be 11. The values of the SSI criterion (right) are based on the maximum difference of each variable between the clusters of taxons on the dendrogram (Figure 5.1), the sizes of the most contrasting clusters and the deviation of a variable in the cluster centres compared to its overall mean. 171

Figure S6.1: The average δ ^{15}N and a low δ^{13}C in bi-plot space for each taxonomic group at each site (1 to 17), the dashed line represents the convex hull around the extent of the occupied niche space, niche width is calculated as the volume within this area. 185

List of tables

Table 2.1 Primers and probes for real-time PCR detection. TF = forward primer; TR = reverse primer, TP = probe. Nucleotide position based on the first copy of the amplicon in pMUM001, accession numbers: IS2404: AF003002, KR: BX649209. .. 34

Table 2.2 Table showing the site locations, and the initial environmental survey sample data, this includes both qPCR and PCR samples. Some field site numbers are missing due to the sites being recorded during the early searches for locations but considered unsuitable for further analysis. Sites used for subsequent biotic surveys are highlighted in bold. 40

Table 2.3 Mean abiotic data readings for each site, these are the mean from 3 readings taken near to the invertebrate and vertebrate survey locations. 42

Table 3.1 Site locations and the number and type of samples taken. In bold both IS2404 and KR identified. 49

Table 3.2: Details of samples with sites positive for IS2404 (CT>35). Abiotic parameters taken from where the sample was collected within the water body is also included. Bacteria per ml of 50ml water sample, or 0.25grams of solid sample. 50

Table 3.3: Details of positive samples for both IS2404 and KR (CT<35). Abiotic parameters taken from where the sample was collected within the water body is also included. Mycobacteria per ml of 50ml water sample, or 0.25grams of solid sample. 51

Table 5.1 Mixed effects model of functional indices at sites against the proportion of samples positive for MU DNA. Values of $p<0.05$ highlighted in bold. Models show how functional evenness and two measures of diversity do not vary between sites as the presence of MU DNA increases. Coefficient is the slope coefficient of regression. See Appendix II Table S5.1 for list of species, traits and their corresponding grouping. F.D. = Functional dispersion, F.E. = Functional evenness, R.Q. = Rao's quadratic entropy. 107

Table 5.2 Mixed effects model of functional groups at sites against the proportion of samples positive for MU. Values of $p<0.05$ highlighted in bold. Models show how functional groups 9 and 10 increase as the presence of MU DNA increases at sites in lentic and lotic habitats. See Appendix II Table S5.4 for list of species, traits and their corresponding grouping. 108

Table 5.3 Mixed effects model of biotic diversity within functional groups at sites against the proportion of samples positive for MU. Values of $p<0.05$ highlighted in bold. Models show how the diversity within functional groups 9 and 10 stays the

same as the presence of MU DNA increases at all sites. Model 9 was tested at lotic sites, as this is where it showed a positive relationship with MU, whilst similarly group 10 was tested at lentic sites. 110

Table 6.2 Results of multiple linear regressions between calculated niche widths for the entire population of organisms surveyed at each site and the square meter cover of deforestation and agricultural/urban land within buffer zones of 500m, 1km and 2km. Coefficient is the slope coefficient, a negative value indicates a decline in explanatory variable as the dependant variable (niche width) increases. 137

Table 6.3 Results of multiple linear regressions between calculated biodiversity indices, actual and theoretical species richness calculated via jacknife for the entire population of organisms surveyed at each site and the square meter cover of deforestation and agricultural/urban land within buffer zones of 500m, 1km and 2km. 137

Table 7.1 How knowledge of MU ecology has been furthered by the results of this PhD. 158

Table S5.1 Key for motility trait in table S5.2: 171

Table S5.2 Definitions of feeding groups 172

Table S5.3 Definitions of food types 173

Table S5.4 Table of species listed with their corresponding functional group and functional traits. 174

Table S5.5 The results of all mixed effects models of functional groups and taxonomic groups at lentic sites, against the number of samples positive for MU. P-values were estimated from the model using the pval.fnc function of the language R package with 10,000 Markov chain Monte Carlo samples. 178

Table S5.6 The results of all mixed effects models of functional groups and taxonomic groups at lotic sites, against the number of samples positive for MU. P-values were estimated from the model using the pval.fnc function of the language R package with 10,000 Markov chain Monte Carlo samples. 181

Table S6.1 The mean regional metrics for all taxa, calculated from the site metrics weighted by the abundance of each taxa at each site. Bacteria is calculated as number of bacilli per mg of organism per ml of supernatant during extraction. . 186

Table S6.2 Site location, biodiversity indices 189

Acknowledgements

This research was supported by a grant from the French Investissements d'Avenir Laboratory of Excellence LabEx CEBA (ANR-10-LABX-2501). AM, RG and JFG are funded by LabEx CEBA. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

I would like to extend thanks to my supervisors; Rodolphe Gozlan, Jean-François Guégan and Demetra Andreou for their help, support and advice throughout this PhD. Thanks (and minor apologies for all the hard work I made them do in the field and the labs) also go to Daniel Sanhueza, Kevin Carolan, Marie Le Croller and Rolland Ruffine. Final thanks to Alyson Pavitt for baked goods and a continual supply of sanity, Colin John Quinn for getting us out of the jungle alive, and to the Celestial Elks.

Declaration

I confirm that the work presented in this thesis is my own work, with the following exceptions:

The field data from Ghana used in chapter 5 was collected by Eric Benbow, funded by Grant Number R01TW007550 from the Fogarty International Centre through the NIH/NSF Ecology of Infectious Diseases Program.

Chapter 1 Literature Review

1.1 Emerging infectious diseases

Emerging infectious diseases (EID's) are perhaps one of the greatest threats to human civilisation. In recent decades several newly identified pathogens including H.I.V., swine flu, S.A.R.S. and Ebola, have emerged from the wild and swept across populations at both epidemic and pandemic scales. During this period numerous other less high-profile aetiological agents have emerged locally (Morens and Fauci 2013), their exact origins often unknown, and with each comes the possibility of a new uncontrollable contagion. As climate change progresses and land-use intensifies the number and frequency of EID's is likely to increase (Jones et al. 2008). In populations of impoverished people living on the boundaries of urbanisation there is a particularly high risk, with little money and poor sanitary conditions, established infections such as leptospirosis, influenza and cholera often take hold (Lau et al. 2010; Chowdhury et al. 2011). And it is in these melting pots of viruses and other pathogens, where humans, polluted water, domestic and wild animals live in close proximity, new infections and mutations are most likely to arise (Jones et al. 2008). For example the origins of H.I.V., Ebola and bird flu can potentially be traced back to such regions (Wolfe et al. 2007; Worobey et al. 2008), and yet despite their dangerous potential, they remain some of the most under-invested areas for disease control and surveillance (Jones et al. 2008).

1.1.1 EID's and the environment

The exact relationship between disease and anthropogenic modification to the surrounding environment is often poorly understood, with land use changes being a predominant driver in the emergence of many infections of humans, or infections

of economic importance (i.e. livestock and crops) (Patz et al. 2000). Land-use can encompass a number of categories, including; deforestation, agricultural practices, road building, dam construction, coastal zone degradation, wetland modification, mining and urbanisation, to name but a few. The intricate shifting dynamics of an infection that occur, can be attributed to the wide ranging implications land-use has on both the ecological aspect of the pathogen, in addition to the socio-economic factors of the people who come into contact with the aetiological agent, its hosts and vectors. Coupled with this, climatic changes including disturbance of seasonal patterns, extreme weather events and long-term temporal shifts often interact with the physical environment and change the behaviour of both people and livestock to disrupt or enhance the transmission of a disease through a biotic network. The mechanisms behind these changing networks depend on the aetiological agent of interest. For example in a simple scenario; many mosquito borne diseases are highly sensitive to fluctuating rainfall, with stagnant water-bodies being the preferred environment for the vectors larvae (Gagnon et al. 2001; Harvell et al. 2002; Chaves and Pascual 2006; Ruiz-Moreno et al. 2007; Lau et al. 2010; Chaves et al. 2012), these in turn are more abundant in deforested areas where there are less trees to absorb the water and an increase in people working on the land.

1.1.2 EID's and biodiversity

Often the biggest impact human encroachment has on an area is biodiversity loss. The implications it has for ecosystem services are thought to be hugely important, yet are largely unknown (Tilman 1999; Chapin et al. 2000; Pereira et al. 2012). A seminal review paper was published in 2010 in *Nature* presenting collated evidence that a loss of biodiversity predominantly equates to an increase in disease (Keesing et al. 2010). The mechanisms reported by which this occurs vary dependent on the pathogen and can be ascribed to the type of interactions it has with other organisms

and how these are influenced by corruption of species assemblages. The review described three pathways of emergence that were predominant within the literature:

1. The abundance of hosts and vectors being diluted by an increase in species diversity.

2. Changes in the behaviour of hosts, vectors or parasites in populations with higher diversity.

3. The physical condition of a host or a vector being influenced by surrounding biodiversity.

Further mechanisms exist and combinations are also possible, making precise predictions of which extinctions will cause changes to disease levels potentially difficult.

Theoretically this process can work both ways, e.g. the loss of biodiversity resulting in the decline of highly suitable host species could cause a decrease in disease prevalence; the authors in this case however found a strong correlation between low biodiversity and high levels of infection is consistently observed. Why this occurs they suggest results from a simple relationship; species that do well in disturbed environments are preferential host choices for pathogens. Such a relationship may be attributable to an overlap or trade-off in biological traits which are beneficial to both resilience to disturbances and susceptibility to disease. This fits well within the context of life history theory. For example; species which have evolved to allocate resources into somatic growth and high reproductive effort (*r* strategists) would have fewer resources to invest in their immune responses. This has been shown across a range of taxa; those with a high growth rate and reproductive output, have a lower risk of extinction and a competitive edge (Cardillo et al. 2005; Cardillo et al. 2008) but are also likely to have

weaker immune systems or other physiological weaknesses to pathogens (Nunn et al. 2003; Nunn and Heymann 2005; Martin Ii et al. 2006; Martin et al. 2007; Lee et al. 2008; Cronin et al. 2010). In addition, it is these species which predominantly live in large close-quartered groups, further facilitating the spread of disease (Nunn and Heymann 2005).

Caution however needs to be expressed when using a simple measure of biodiversity in a complex system. The transmission of a disease or parasite and its influence on a multifaceted food web must be considered carefully. Critiques of the above literature have highlighted the potential for simplification and misrepresentation of results in some of the above studies and argue that whilst biodiversity can be responsible for an increase in disease, there are also numerous situations where the opposite is true or where changes in biological communities, environmental, socio-economic factors are a greater driver of disease incidence (Roche and Guégan 2011; Randolph and Dobson 2012). It is important therefore within this context to also look closely at environmental and socio-economic factors, community composition of sites and the functional roles of species within a system and not to rely purely on an index of biodiversity as a measure of transmission potential.

1.2 Objectives

As human technology advances, our ability to fight infectious disease is one of the few areas of science which has begun to regress, antibiotics are rapidly losing their effectiveness (Laxminarayan 2014) and our options to combat both established and unknown pathogens are therefore becoming more limited. With this in mind, one of the best hopes for the future is an ability to accurately predict and prevent new diseases. By understanding the ecology and behaviour of emerging pathogens, there is a far greater potential to limit their spread or avoid their emergence altogether, thereby protecting at-risk populations and avoiding the next fatal pandemic. In this PhD a detailed examination of a model emerging infectious pathogen is undertaken. The aim is to use a variety of techniques to understand the ecological and environmental factors that can influence the more generalist microbes on the boundaries between benign and infectious.

1. Identify and map the environmental sources of the pathogen in French Guiana, South America.
2. Categorize relationships between outbreaks of the disease and climatic fluctuations.
3. Link the presence of the bacteria in the environment to specific indicator taxa in tropical freshwater systems and test the dilution theory on a generalist emergent pathogen.
4. Identify how changes in the aquatic communities of these systems can affect the presence of bacterial hosts and link these changes to anthropogenic factors.

1.3 Buruli ulcer

The focal disease of the PhD is Buruli Ulcer (BU), a necrotising skin infection typically found in tropical and sub-tropical regions and caused by the mycobacterium *Mycobacterium ulcerans* (MU). The history of documented human infection dates back to 1897 when Sir Albert Cook first described the symptoms in Uganda. However detailed analysis and identification of the pathogen was not published in the medical literature until 1948 by Professor Peter MacCallum (MacCallum et al. 1948). Since its discovery pathologies of BU have been identified in a number of countries, predominantly in Africa but also, Central and South America, Australia, China, and South East Asia (Figure 1.1).

Whilst there is still debate over the specifics of human transmission (and this may vary between countries); there is overwhelming evidence that BU is associated with numerous aquatic species (Marsollier et al. 2002; Marsollier et al. 2004a; Marsollier et al. 2004b; Merritt et al. 2005; Eddyani et al. 2008; Mosi et al. 2008; Williamson et al. 2008; Merritt et al. 2010; Willson et al. 2013; Garchitorena et al. 2014) and the level of reductive evolution in the genome suggests it is liable to be dependent on other organisms for survival (Stinear et al. 2007; Demangel et al. 2009). This implies there is likely to be a significant correlation between changes in species assemblages through biodiversity loss, climatic fluctuations, anthropogenic change and the emergent prevalence of the disease (Ramakrishnan et al. 1997; Merritt et al. 2005; Brou et al. 2008; Garchitorena et al. 2014).

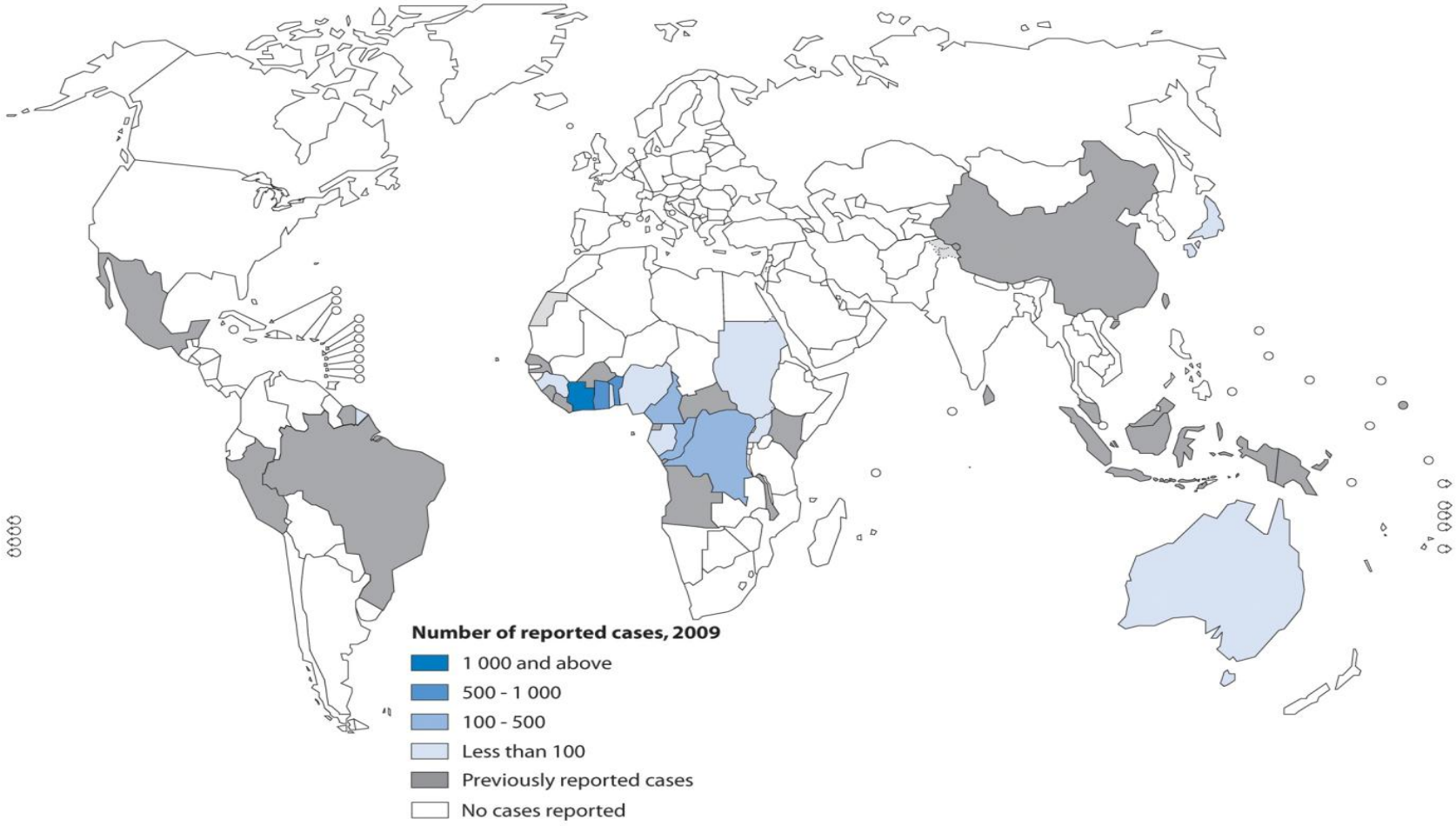


Figure 1.1 Global map showing countries reporting cases of BU (Merritt et al. 2010)

1.3.1 Epidemiology

1.3.1.1 Africa

After publication of the disease by MacCallum, reports of BU in Africa began to increase and numerous infections previously un-identified but bearing similar symptoms came to light. For example during the period 1942 to 1949, after Van Oye and Ballion (1950) identified the disease in a 6 year old boy living in the now Democratic Republic of Congo. Their associate, Professor P. G. Janssens noted on their report of the patient, that he had seen 81 cases of similar symptoms, predominantly in children from a 10km strip of savannah between Nzoro and Kibali rivers in the same country (Van Oye and Ballion 1950).

It was in 1961 during extensive work on the disease by Clancey et al. (1962), in a region of Uganda near the Nile River called "Buruli", that BU got its official title. Before this a variety of nomenclature was associated with it dependent on the location of diagnosis (Clancey et al. 1962). Since this period, BU has been reported in 20 African countries, mainly in the western part of the continent, with Benin, Ghana and Côte d'Ivoire currently carrying the highest risks of infection. Epidemiological studies from these areas indicate the increasing significance BU has as a major contagion in African communities. In Benin, between 1997 and 2001 throughout the Zou region of the south, cases of MU were more common than the infamous mycobacterium infections Tuberculosis and Leprosy, with 21.5/100,000 infections being attributed to the disease (Debacker et al. 2004). Whilst in Côte d'Ivoire between 1978 and 1999 the number of cases was over 15,000 and prevalence in some communities was as high as 16% of the population. In Ghana this figure grows further with 6000 reports of infection in 1999 alone and all 10 regions of the country being identified as carrying risk (World Health Organisation 2001b).

1.3.1.2 Australia

Cases of the disease in Australia have predominantly been low, approximately 1-2 per year, however in the 1990's the number of reported infections increased dramatically. This increase can be attributed to a significant outbreak in Cowes, a small town on Phillip Island between the period of 1992 and 1995. The outbreak coincided with the opening of a new Golf resort leading many to believe infection was being spread through the resorts irrigation system, which used water sourced from the local environment. A study conducted by Ross et al. (1997) used an adaptation of a PCR-based assay, previously developed for clinical diagnosis to amplify DNA from samples of water taken from both the system and its source. Five samples provided positive results and it was the first time MU DNA had been amplified from an environmental sample. In addition to providing a strong case for the irrigation system spreading the disease, it gave further evidence to the theory that the mycobacteria were spread via aquatic systems (Ross et al. 1997). Since this period further outbreaks have occurred, predominantly in the south east Victoria area of Australia where the disease has also been diagnosed in number of different animals including: koalas, common ringtail possums, a common brushtail possum, a mountain brushtail possum, a long-footed potoroo, horses, dogs and alpacas and cats (Fyfe et al. 2010).

1.3.1.3 Central and South America

In most regions the number of cases in Central and South America is extremely low, with the exception of French Guiana, where the majority of this study was conducted. During the period of 1969 to 2011, 242 cases of BU were reported, the low population (274,652) makes this number of infections relatively high, with an average infection rate of 2.09/100,000 persons per year (Figure 1.2). This may however be a result of better medical infrastructure leading to improved diagnostic screening in this French overseas territory.

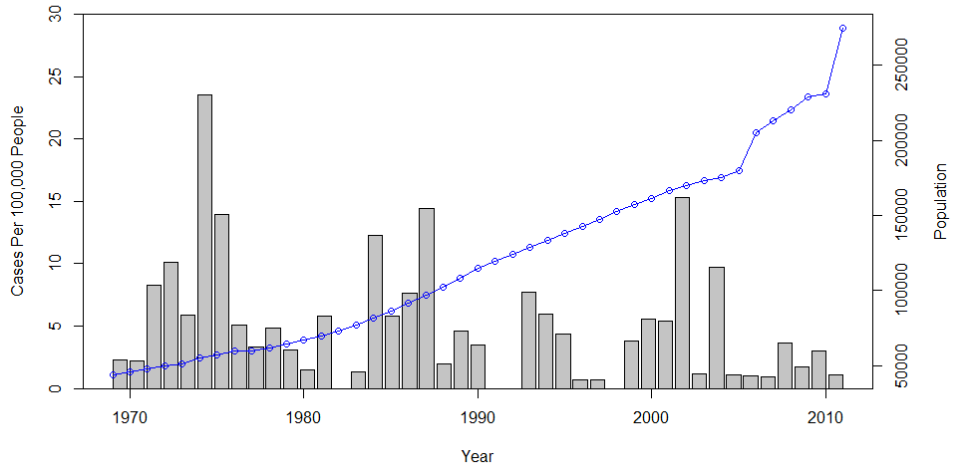


Figure 1.2 Cases of BU from 1969 to 2012 per 100,000 people, the dotted line represents the increasing population.

1.3.2 Biology

MU belongs to the *Mycobacteria*, a genus of *Actinomycobacteria* in the suborder of *Corynemycobacteria*. Mycobacteria are generally slow growing and all are acid-fast, meaning they retain dyes following an acid-alcohol decolourisation test. Within the Mycobacteria MU is part of the *Mycobacterium marinum* complex, a group of very closely related Mycobacteria which all share a common ancestry (Käser et al. 2009). Like most Mycobacteria it has a slow generation time of about 20 hours and primary cultures may take between 6 to 8 weeks to develop. Cultures are most successfully grown on Lowenstein-Jensen medium at temperatures lower than most other mycobacteria (29-33°C), and have also shown enhanced growth under micro-aerophilic conditions of 2.5-5% oxygen (World Health Organisation 2001a).

The predominant mechanism by which MU expresses trauma to its host is through the necrosis of skin tissue and underlying fat. This process is the result of a secretion of a recently identified toxin, mycolactone. Mycolactone is a polyketide-derived 12 membered ring macrolide with the chemical formulae: $C_{44}H_{70}O_9$

(Merritt et al. 2010). It causes necrosis through the induction of apoptosis in cells and also acts as an immunosuppressant. This mode of cell destruction, may be the reason why infection by MU is not accompanied by an acute inflammatory response, often making the site of infection painless (George et al. 2000).

Our understanding of human infection from mycobacteria predominately stems from two species, *M. tuberculosis* and *M. leprae*. Most other mycobacteria are classed as "opportunistic" or "occasional" pathogens. These less virulent species often require certain conditions to facilitate infection of humans, such as sites of trauma or a weakened immune system (a serious concern as cases of HIV continue to increase in endemic areas) (Portaels 1995). Whilst often considered an opportunistic species, as the mode of transmission for MU remains uncertain, it is difficult to identify how contagious the bacilli are. The level of reductive evolution within the genome, and the presence of genes for the synthesis of mycolactone would suggest however, that MU is highly dependent on a host and more virulent than other opportunistic relations. This would also suggest it is less likely to be free living within the environment (Stinear et al. 2007; Demangel et al. 2009).

1.3.3 MU evolution

Complete genome sequencing suggests MU recently evolved from *M. marinum* through lateral gene transfer and significant reductive evolution (Doig et al. 2012). There are three described ecovars of MU which have been isolated from humans, fish and amphibians in a number of locations, *Mycobacterium shinshuense, Mycobacterium pseudoshottsii*, and "*Mycobacterium liflandii"* (Doig et al. 2012). Using 8-gene multi locus sequence typing, it is possible to identify that they evolved from an *M. marinum* ancestor before subsequently diverging into two principal lineages. These lineages are often termed mycolactone-producing Mycobacteria, as like MU they produce the cytotoxin mycolactone (Yip et al.

2007), although they are not typically found in humans, there has been suggestion that they should still be labled as *M. ulcerans* rather than distinct ecovars (Doig et al. 2012).

The divergence has involved the accumulation of multiple copies of the insertion sequences IS2404 and IS2606, 771 pseudogenes, two bacteriophages and multiple DNA deletions (Stinear et al. 2007). This evolution suggests it has begun to exploit a new environmental niche, differing from closely related species. The pathogenic nature of MU suggests that this new niche is either directly as a parasite of other species or a symbiote which can become pathogenic under certain conditions. Recent analysis suggests that this adaptation is still ongoing, mediated by changes in IS2404 and IS2606, with 23 different African insertion sequence element single nucleotide polymopism (ISE-SNP) types dominating in different areas of close proximity in one study alone (Vandelannoote et al. 2014). These variations have likely arisen due to geographical barriers between water basins, limiting gene flow between bacteria. Because of this genetic plasticity, it has been suggested that the reason for less virulence of BU outside of Africa is due to differing ecological pressures which have created a strain which has less need to become pathogenic. This however is not backed up by any empirical evidence and there needs to be more work done in this area.

1.3.4 Pathology

The incubation time for MU can be between two to three months and BU can present itself on a patient in four main clinical forms (World Health Organisation 2001a):

- **Nodule**: A firm raised lesion located in the subcutaneous tissue, usually attached to the skin and approximately 1 – 2 cm in diameter. Skin covering

the lesion is often hypopigmented and whilst painless, nodules can cause itching.

- **Papule:** A similarly painless, raised skin lesion surrounded by reddened skin and less than 1cm across.
- **Plaque:** A raised, painless, hard area of skin with irregular boundaries, often with a degree of de-pigmentation or spotted erythema. They can grow large sometimes reaching 15cm in diameter and often develop into irregularly shaped ulcers.
- **Oedematous:** An extensive diffuse, non-pitted swelling with irregular margins involving part of or an entire limb, firm and can be painful. This form of infection is also often associated with fever.

Approximately 70% of all infections occur in children with an equal ratio of male and females affected. Most lesions occur on the extremities; however other areas of the body can be afflicted, including the trunk and the head/face (World Health Organisation 2001a). The most common progression of the disease begins with the forming of a nodule before rupturing to become an ulcer, with a raised undermined edge and a cotton-wool like necrotic slough at the base. The greatest risk to patients is through the loss of organs such as the eye, amputations of limbs, disability caused from scarring and poor re-growth of the skin or illness through secondary infections. The long gestation period and the slow growing nature of the ulcers mean that reporting of the disease is often delayed making identification of the source of infection extremely difficult (World Health Organisation 2001a).

1.3.5 Ecology

MU is typically associated with aquatic environments and incidences of the disease are significantly higher in areas during and after flooding events, or where people come into continual contact with rivers, ponds, swamps and lakes (Merritt et al.

2005; Merritt et al. 2010; Garchitorena et al. 2014; Landier et al. 2014), see also chapters 3 and 4. Since the identification of the IS2404 region of DNA as an accurate tool for identifying MU through Polymerase Chain Reaction (PCR) (Stinear et al. 1999), DNA and cultures of the mycobacteria have been associated with numerous aquatic species. Aquatic plants have been found to facilitate the formation of biofilms of MU on their surface, with extracts of *Rhizoclonium* sp. and *Hydrodictyon reticulatum* causing a doubling of the usual generation time (Marsollier et al. 2004a). Freshwater snails have also been tested positive for the DNA and have been experimentally infected after feeding on plant biofilms containing the mycobacteria, suggesting they may play a prominent role as disease reservoirs (Marsollier et al. 2004b). Further identification of DNA has come via extracts of detritus, fish, molluscs and amphibians (Merritt et al. 2010). This has added evidence to a hypothetical transmission model originally described by Portaels et al. (1999) and further adapted by Marsollier et al. (2005) to include species more recently identified as testing positive for IS2404. The model describes how MU present in detritus, soil and in biofilms on plants, is ingested by intermediate reservoirs such as freshwater snails which are subsequently fed upon by predaceous Hemiptera (a family of aquatic water bugs containing species such as water-boatman and pond-skaters, some of which are known to bite humans). After inoculation through feeding, these predators are then able to mechanically transmit the mycobacteria into the skin of a human. The snails along with other aquatic organisms such as fish are also hypothesised to concentrate MU in their faeces, cycling it back into the system with faecal providing a second route in which Hemiptera can become infected.

Providing definitive evidence of this transmission route has however proved difficult, as whilst DNA extraction is relatively easy, successful isolation of viable mycobacteria from natural sources has been largely unsuccessful. In 2002, isolation

of viable bacilli from the environment was achieved for the first time by Marsollier et al (2002), via extracts of saliva from a family of aquatic Hemiptera capable of biting humans, the Naucoridae. The insects were experimentally infected by feeding on inoculated larvae and the mycobacteria were observed growing and multiplying within the salivary glands. There were no apparent adverse effects on the Naucoridae by MU and the mycobacteria was able to be extracted and re-cultured within the lab. Subsequently, in vitro infection of mice through biting by the Naucoridae was achieved, providing further evidence the mycobacteria were transmissible and able to propagate within the insect. During the same study, wild caught Naucoridae from the BU endemic region of Daloa in Côte d'Ivoire were also identified as carrying likely strains of MU. Successful cultivations of mycobacteria were achieved from 2 of 80 wild caught insects, with PCR analysis of these cultures showing specific genetic sequences found in MU. In addition genetic identification of MU with IS2404 was achieved in 5 specimens. Further experimentation in 2007 has also identified mechanisms by which the mycobacteria could enter the salivary glands (Marsollier et al. 2007)

Caution however needs to be expressed in suggesting water bugs as a definitive and significant vector of BU, with the implications of these findings being so high; criticisms have been raised over the validity of the experiment (Benbow et al. 2008; Portaels et al. 2008). The isolations from wild caught insects were mycobacterial strains related to MU, showing shared DNA sequences from the IS2404 region. This region on its own, whilst accurate is not wholly reliable, as strains not causative of BU in humans may also test positive (Yip et al. 2007). For increased validation recent studies have also looked for additional DNA sequences such as those for the coding of mycolactone (Williamson et al. 2008). Further to this, the presence of DNA does not equate to a viable and transmissible organism. In order to definitively identify the presence of living human-infecting strains of

MU, extensive phenotypic analysis is often sought (Vincent Lévy-Frébault and Portaels 1992). In this study transmission of infection to mice through biting was successful; the source of this inoculation however came from the experimentally infected insects, not wild insects, which were fed on prey uncommon to them in nature (blowfly larvae). In addition the insects were infected with a strain of MU from a differing geographical location (Water bugs from France and a South American strain of MU from French Guiana). These factors may have caused a response that may not be seen in the wild. Despite the study being more proof of concept, it provided the first substantial evidence that BU could be carried and transmitted by an insect vector.

In 2008 Mosi, et al. (2008) sought to further validate the claims a predaceous insect could be infected through feeding. Using an infection model based on naturally occurring predator-prey relationships, giant water bugs (Belostomidae) collected from Africa were fed on infected mosquito larvae. The results confirmed Marsollier's findings and the insects became infected, although in this study mycobacteria were found to accumulate on the exoskeleton rather than within salivary glands (Mosi et al. 2008). Also in 2008 the first isolation of MU from an environmental source, identified through both extensive genetic and phenotypic analysis was achieved (Portaels et al. 2008). The mycobacteria were extracted and cultured from a species of water-boatman (*Gerris sp.* not associated with biting humans), taken from BU endemic regions of Benin and Togo. The culture tested positive for IS2404 with a banding pattern identical to that of African strains of MU. In addition nucleotide sequences of the 16 sRNA genes as described by Portaels et al. (1996) were consistent with an African type strain of MU, except for a single substitution at position 1317. Thin layer chromatography of the mycolactone produced by the isolate also showed the major lipid species to be A/B which, according to mycolactone heterogeneity analysis by Mve-Obiant et al.

(2003) also suggests the mycobacteria was a strain of African MU. Phenotypic analysis involved the inoculation of mice and histopathological observations using mycobacteria from wild caught specimens, both of which produced observations consistent with strains that have been described in mice and humans, in particular the spread of the disease to the bones. Despite the source of mycobacteria being from a species not known to bite humans, it is closely related to other more aggressive Hemiptera and therefore those species could feasibly have a similar ability to carry the mycobacteria. More recently a study by Marion et al (2010) also provided additional evidence of the presence of MU in the salivary glands of various wild caught biting Hemiptera, through both DNA analysis and the successful inoculation of mice using saliva.

To confirm the viability of biting Hemiptera as a significant source of infection, a large scale field experiment was conducted. The study involved 15 BU endemic and 12 non-endemic areas of Ghana, Africa, however the results found no significant relationship (Benbow et al. 2008). This suggests that whilst biting Hemiptera may carry and feasibly transmit BU, their role as vectors is perhaps minor. This major study, the first of its size, necessitated a serious re-think of the previously described infection model and highlights the urgent need for further study in this area. This study was followed up by a second in 2013 where again similar results were found, with little firm evidence of a specific vector or host being identified (Benbow et al. 2013).

Other modes of transmission have recently been suggested, particularly in areas outside of Africa where prolonged contact with aquatic environments is significantly lower. For example in Australia where the use of water is often limited to the artificial systems present in people homes and gardens. The disease here, unlike Africa, is characterised by sudden, large outbreaks within a small geographical area and it may be how water is being used that is key to explaining

this difference. The most famous of these outbreaks was perhaps in Phillip Island, as described in detail in section 1.2.1. Evidence traced the source of infection to an irrigation system at a newly built golf resort, it was suggested MU in drifting aerosols were the cause and subsequent modification of the system saw a dramatic decline in the disease. Further outbreaks however have been unable to find links with sources of aerosols.

There has also been significant interest in the possibility of MU being transmitted by mosquitoes after some anecdotal evidence of patients complaining the disease appeared at the location of previous mosquito bites. This has led to some preliminary work on mosquitoes as possible vectors. An important recent find, was the identification of MU DNA in large numbers of *Aedes camptorhynchus* from BU endemic regions in Australia (Johnson et al. 2007). A recently described mode of transmission model involving mammals as reservoirs and mosquito vectors (Fyfe et al. 2010) has also been postulated, although the evidence for this remains inconclusive. Despite this, there remains the potential for low level basal organisms such as mosquito larvae to be key hosts of the bacteria; theoretical studies on food webs have shown the importance of oligochaeta in the persistence of the bacterium in the environment (Roche et al. 2013). It is not unfeasible that if the bacilli can utilise such taxa that other highly similar taxa such as mosquitoe and midge larvae could also be a suitable host.

1.3.6 Human impacts and landscape induced change on the infection rates of BU

BU is heavily linked with human modification of aquatic systems and has been prolific in areas of deforestation, mining and agriculture, particularly when slow moving or stagnant water is present (Merritt et al. 2005; Brou et al. 2008). As the mechanisms of BU infection remain inconclusive, it is not possible to identify the

optimum conditions necessary for the disease to spread. It is possible however speculate on preferential environments, based on the current evidence and physiological traits. Sequencing of the genome reveals a lack of *crt*I which is responsible for the production of phytoene dehydrogenase, an enzyme that within *M. marinum* is necessary for the synthesis of light-inducible carotenoid pigments. The loss of need for these pigments, which give protection against UV-induced damage, suggests that MU occupies a niche where light levels are low (Ramakrishnan et al. 1997). In addition MU has lost the genes for anaerobic respiration, which would suggest that its occupied niche also has low levels of dissolved oxygen. This has been confirmed in vitro where the mycobacteria have shown a preference for a low oxygen environment (Palomino et al. 1998). These traits therefore suggest that the mycobacteria have a preference for aquatic systems that have low light and oxygen levels. This gives an indication as to the possible reasons why human socio-economic activities increase infection rates. For example, rivers and ponds around which have suffered deforestation and subsequent soil erosion are often areas which have high numbers of cases. Deforestation may cause increases in water temperature to optimal conditions through lack of shading. Further to this, increases in sedimentation and eutrophication from agriculture may lower UV penetration and dissolved oxygen facilitating mycobacterial growth. This increase, in areas which are high in sedimentation and run-off also makes a logical connection to the higher prevalence of the disease after heavy rains and flooding. Agriculture (such as rice fields, notably in the Côte d'Ivoire region) can further compound the issue by utilising irrigation systems which use water from these sources, thereby spreading the mycobacteria into areas of high human activity.

1.4 Main country of study: French Guiana

French Guiana is a French ultra-peripheral territory within South America, bordering the countries of Brazil to the east and south and Suriname to the west. It is 83,534km^2 with a very low population density; most inhabitants reside in a 50km wide strip along the coastline. The rest of the country is almost purely, pristine primary tropical rain forest and contains the 33,900 square kilometres Guiana Amazonian national park as well as a wealth of important ecosystems ranging from marshland to coastal mangroves.

First attempts at colonisation of the territory occurred in 1763, although were largely unsuccessful due to disease and the harsh climate. Subsequently large numbers of prisoners were sent to the penal colonies set up on the mainland and also on adjacent islands, including the notorious Île du Diable. Population growth did not principally begin to rise however until the mid 1900's; this was largely due to the lack of exportable goods and difficulty in undertaking standard agricultural practices owing to the nutrient poor soil. Since this period, growth has been largely attributed to gold mining, fishing and more recently the space industry, with the European Space Agency's main launch facilities now being based in Kourou. Despite economic growth, the territory is still largely dependent on France for subsidies and imports and there is a significant amount of poverty, primarily in illegal immigrants who cross the difficult to police jungle borders.

1.5 References

Benbow, M. E., Kimbirauskas, R., McIntosh, M. D., Williamson, H., Quaye, C., Boakye, D., Small, P. L. C. and Merritt, R. W., 2013. Aquatic Macroinvertebrate Assemblages of Ghana, West Africa: Understanding the Ecology of a Neglected Tropical Disease. *EcoHealth*, 1-16.

Benbow, M. E., Williamson, H., Kimbirauskas, R., McIntosh, M. D., Kolar, R., Quaye, C., Akpabey, F., Boakye, D., Small, P. and Merritt, R. W., 2008. Aquatic Invertebrates as Unlikely Vectors of Buruli Ulcer Disease. *Emerging Infectious Diseases*, 14 (8), 1247-1254.

Brou, T., Broutin, H., Elguero, E., Asse, H. and Guegan, J-F., 2008. Landscape Diversity Related to Buruli Ulcer Disease in Côte d'Ivoire. *PLoS Neglected Tropical Diseases*, 2 (7), e271.

Cardillo, M., Mace, G. M., Gittleman, J. L., Jones, K. E., Bielby, J. and Purvis, A., 2008. The predictability of extinction: biological and external correlates of decline in mammals. *Proceedings of the Royal Society B: Biological Sciences*, 275 (1641), 1441-1448.

Cardillo, M., Mace, G. M., Jones, K. E., Bielby, J., Bininda-Emonds, O. R. P., Sechrest, W., Orme, C. D. L. and Purvis, A., 2005. Multiple Causes of High Extinction Risk in Large Mammal Species. *Science*, 309 (5738), 1239-1241.

Chapin, F. S., Zavaleta, E. S., Eviner, V. T., Naylor, R. L., Vitousek, P. M., Reynolds, H. L., Hooper, D. U., Lavorel, S., Sala, O. E., Hobbie, S. E., Mack, M. C. and Diaz, S., 2000. Consequences of changing biodiversity. *Nature*, 405 (6783), 234-242.

Chaves, L. F. and Pascual, M., 2006. Climate Cycles and Forecasts of Cutaneous Leishmaniasis, a Nonstationary Vector-Borne Disease. *PLoS Medicine*, 3 (8), e295.

Chaves, L. F., Satake, A., Hashizume, M. and Minakawa, N., 2012. Indian Ocean Dipole and Rainfall Drive a Moran Effect in East Africa Malaria Transmission. *Journal of Infectious Diseases*, 205 (12), 1885-1891.

Chowdhury, F., Rahman, M. A., Begum, Y. A., Khan, A. I., Faruque, A. S. G., Saha, N. C., Baby, N. I., Malek, M. A., Kumar, A. R. and Svennerholm, A.-M., 2011. Impact of rapid urbanization on the rates of infection by Vibrio cholerae O1 and enterotoxigenic Escherichia coli in Dhaka, Bangladesh. *PLoS Neglected Tropical Diseases*, 5 (4), e999.

Clancey, J., Dodge, R. and Lunn, H., 1962. Study of a mycobacterium causing skin ulceration in Uganda. *Annales de la Société Belge de Médecine Tropicale*, (4), 585–590.

Cronin, J. P., Welsh, M. E., Dekkers, M. G., Abercrombie, S. T. and Mitchell, C. E., 2010. Host physiological phenotype explains pathogen reservoir potential. *Ecology Letters*, 13 (10), 1221-1232.

Debacker, M., Aguiar, J., Steunou, C., Zinsou, C., Meyers, W. M., Guédénon, A., Scott, J. T., Dramaix, M. and Portaels, F., 2004. Mycobacterium ulcerans Disease (Buruli ulcer) in Rural Hospital, Southern Benin, 1997–2001. *Emerging Infectious Diseases*, 10 (8).

Demangel, C., Stinear, T. P. and Cole, S. T., 2009. Buruli ulcer: reductive evolution enhances pathogenicity of *Mycobacterium ulcerans*. *Nature Reviews Microbiology*, 7 (1), 50-60.

Doig, K. D., Holt, K. E., Fyfe, J. A. M., Lavender, C. J., Eddyani, M., Portaels, F., Yeboah-Manu, D., Pluschke, G., Seemann, T. and Stinear, T. P., 2012. On the origin of *Mycobacterium ulcerans*, the causative agent of Buruli ulcer. *BMC genomics*, 13 (1), 258.

Eddyani, M., De Jonckheere, J. F., Durnez, L., Suykerbuyk, P., Leirs, H. and Portaels, F., 2008. Occurrence of free-living amoebae in communities of

low and high endemicity for Buruli ulcer in southern Benin. *Applied and environmental microbiology*, 74 (21), 6547-6553.

Fyfe, J. A. M., Lavender, C. J., Handasyde, K. A., Legione, A. R., O'Brien, C. R., Stinear, T. P., Pidot, S. J., Seemann, T., Benbow, M. E., Wallace, J. R., McCowan, C. and Johnson, P. D. R., 2010. A Major Role for Mammals in the Ecology of *Mycobacterium ulcerans*. *PLoS Neglected Tropical Diseases*, 4 (8), e791.

Gagnon, A. S., Bush, A. B. G. and Smoyer-Tomic, K. E., 2001. Dengue epidemics and the El Niño Southern Oscillation. *Climate Research*, 19 (1), 35-43.

Garchitorena, A., Roche, B., Kamgang, R., Ossomba, J., Babonneau, J., Landier, J., Fontanet, A., Flahault, A., Eyangoh, S. and Guégan, J.-F., 2014. *Mycobacterium ulcerans* Ecological Dynamics and Its Association with Freshwater Ecosystems and Aquatic Communities: Results from a 12-Month Environmental Survey in Cameroon. *PLoS Neglected Tropical Diseases*, 8 (5), e2879.

George, K. M., Pascopella, L., Welty, D. M. and Small, P. L. C., 2000. A *Mycobacterium ulcerans* Toxin, Mycolactone, Causes Apoptosis in Guinea Pig Ulcers and Tissue Culture Cells. *Infection and Immunity.*, 68 (2), 877-883.

Harvell, C. D., Mitchell, C. E., Ward, J. R., Altizer, S., Dobson, A. P., Ostfeld, R. S. and Samuel, M. D., 2002. Climate Warming and Disease Risks for Terrestrial and Marine Biota. *Science*, 296 (5576), 2158-2162.

Johnson, P. D. R., Azuolas, J., Lavender, C. J., Wishart, E., Stinear, T. P., Hayman, J. A., Brown, L., Jenkin, G. A. and Fyfe, J. A. M., 2007. *Mycobacterium ulcerans* in Mosquitoes Captured during Outbreak of Buruli Ulcer, Southeastern Australia. *Emerging Infectious Diseases*, 13 (11).

Jones, K. E., Patel, N. G., Levy, M. A., Storeygard, A., Balk, D., Gittleman, J. L. and Daszak, P., 2008. Global trends in emerging infectious diseases. *Nature*, 451 (7181), 990-993.

Keesing, F., Belden, L. K., Daszak, P., Dobson, A., Harvell, C. D., Holt, R. D., Hudson, P., Jolles, A., Jones, K. E., Mitchell, C. E., Myers, S. S., Bogich, T. and Ostfeld, R. S., 2010. Impacts of biodiversity on the emergence and transmission of infectious diseases. *Nature*, 468 (7324), 647-652.

Käser, M., Hauser, J., Small, P. and Pluschke, G., 2009. Large Sequence Polymorphisms Unveil the Phylogenetic Relationship of Environmental and Pathogenic Mycobacteria Related to *Mycobacterium ulcerans*. *Applied and Environmental Microbiology*, 75 (17), 5667-5675.

Landier, J., Gaudart, J., Carolan, K., Seen, D. L., Guégan, J.-F., Eyangoh, S., Fontanet, A. and Texier, G., 2014. Spatio-temporal Patterns and Landscape-Associated Risk of Buruli Ulcer in Akonolinga, Cameroon. *PLoS Neglected Tropical Diseases*, 8 (9), e3123.

Lau, C. L., Smythe, L. D., Craig, S. B. and Weinstein, P., 2010. Climate change, flooding, urbanisation and leptospirosis: fuelling the fire? *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 104 (10), 631-638.

Laxminarayan, R., 2014. Antibiotic effectiveness: Balancing conservation against innovation. *Science*, 345 (6202), 1299-1301.

Lee, K. A., Wikelski, M., Robinson, W. D., Robinson, T. R. and Klasing, K. C., 2008. Constitutive immune defences correlate with life-history variables in tropical birds. *Journal of Animal Ecology*, 77 (2), 356-363.

MacCallum, P., Tolhurst, J. C., Buckle, G. and Sissons, H. A., 1948. A new mycobacterial infection in man. *The Journal of Pathology and Bacteriology*, 60 (1), 93-122.

Marion, E., Eyangoh, S., Yeramian, E., Doannio, J., Landier, J., Aubry, J., Fontanet, A., Rogier, C., Cassisa, V., Cottin, J., Marot, A., Eveillard, M.,

Kamdem, Y., Legras, P., Deshayes, C., Saint-André, J.-P. and Marsollier, L., 2010. Seasonal and Regional Dynamics of *M. ulcerans* Transmission in Environmental Context: Deciphering the Role of Water Bugs as Hosts and Vectors. *PLoS Neglected Tropical Diseases*, 4 (7), e731.

Marsollier, L., André, J.-P. S., Frigui, W., Reysset, G., Milon, G., Carbonnelle, B., Aubry, J. and Cole, S. T., 2007. Early trafficking events of *Mycobacterium ulcerans* within Naucoris cimicoides. *Cellular Microbiology*, 9 (2), 347-355.

Marsollier, L., Robert, R., Aubry, J., Saint Andre, J.-P., Kouakou, H., Legras, P., Manceau, A.-L., Mahaza, C. and Carbonnelle, B., 2002. Aquatic Insects as a Vector for *Mycobacterium ulcerans*. *Applied and Environmental Microbiology.*, 68 (9), 4623-4628.

Marsollier, L., Stinear, T., Aubry, J., Saint André, J. P., Robert, R., Legras, P., Manceau, A.-L., Audrain, C., Bourdon, S., Kouakou, H. and Carbonnelle, B., 2004a. Aquatic Plants Stimulate the Growth of and Biofilm Formation by *Mycobacterium ulcerans* in Axenic Culture and Harbor These Bacteria in the Environment. *Applied and Environmental Microbiology*, 70 (2), 1097-1103.

Marsollier, L., Sévérin, T., Aubry, J., Merritt, R. W., Saint André, J.-P., Legras, P., Manceau, A.-L., Chauty, A., Carbonnelle, B. and Cole, S. T., 2004b. Aquatic Snails, Passive Hosts of *Mycobacterium ulcerans*. *Applied and Environmental Microbiology*, 70 (10), 6296-6298.

Martin Ii, L., Hasselquist, D. and Wikelski, M., 2006. Investment in immune defense is linked to pace of life in house sparrows. *Oecologia*, 147 (4), 565-575.

Martin, L. B., Weil, Z. M. and Nelson, R. J., 2007. Immune defense and reproductive pace of life in peromyscus mice. *Ecology*, 88 (10), 2516-2528.

Merritt, R. W., Benbow, M. E. and Small, P. L. C., 2005. Unraveling an Emerging Disease Associated with Disturbed Aquatic Environments: The Case of Buruli Ulcer. *Frontiers in Ecology and the Environment*, 3 (6), 323-331.

Merritt, R. W., Walker, E. D., Small, P. L. C., Wallace, J. R., Johnson, P. D. R., Benbow, M. E. and Boakye, D. A., 2010. Ecology and Transmission of Buruli Ulcer Disease: A Systematic Review. *PLoS Neglected Tropical Diseases*, 4 (12), e911.

Morens, D. M. and Fauci, A. S., 2013. Emerging infectious diseases: threats to human health and global stability. *PLoS Pathogens*, 9 (7), e1003467.

Mosi, L., Williamson, H., Wallace, J. R., Merritt, R. W. and Small, P. L. C., 2008. Persistent Association of *Mycobacterium ulcerans* with West African Predaceous Insects of the Family Belostomatidae. *Applied and Environmental Microbiology*, 74 (22), 7036-7042.

Mve-Obiang, A., Lee, R. E., Portaels, F. and Small, P. L. C., 2003. Heterogeneity of Mycolactones Produced by Clinical Isolates of *Mycobacterium ulcerans*: Implications for Virulence. *Infection and Immunity.*, 71 (2), 774-783.

Nunn, C. and Heymann, E., 2005. Malaria infection and host behavior: a comparative study of Neotropical primates. *Behavioral Ecology and Sociobiology*, 59 (1), 30-37.

Nunn, C. L., Gittleman, J. L. and Antonovics, J., 2003. A comparative study of white blood cell counts and disease risk in carnivores. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270 (1513), 347-356.

Palomino, J. C., Obiang, A. M., Realini, L., Meyers, W. M. and Portaels, F., 1998. Effect of Oxygen on Growth of *Mycobacterium ulcerans* in the BACTEC System. *Journal of Clinical Microbiology*, 36 (11), 3420-3422.

Patz, J. A., Graczyk, T. K., Geller, N. and Vittor, A. Y., 2000. Effects of environmental change on emerging parasitic diseases. *International Journal for Parasitology*, 30 (12), 1395-1405.

Pereira, H. M., Navarro, L. M. and Martins, I. S., 2012. Global biodiversity change: the bad, the good, and the unknown. *Annual Review of Environment and Resources*, (37), 25 -50.

Portaels, F., 1995. Epidemiology of mycobacterial diseases. *Clinics in Dermatology*, 13 (3), 207-222.

Portaels, F., Elsen, P., Guimaraes-Peres, A., Fonteyne, P.-A. and Meyers, W. M., 1999. Insects in the transmission of *Mycobacterium ulcerans* infection. *The Lancet*, 353 (9157), 986.

Portaels, F., Fonteyene, P., de Beenhouwer, H., de Rijk, P., Guedenon, A., Hayman, J. and Meyers, M., 1996. Variability in 3' end of 16S rRNA sequence of *Mycobacterium ulcerans* is related to geographic origin of isolates. *Journal of Clinical Microbiology*, 34 (4), 962-965.

Portaels, F., Meyers, W. M., Ablordey, A., Castro, A. G., Chemlal, K., de Rijk, P., Elsen, P., Fissette, K., Fraga, A. G., Lee, R., Mahrous, E., Small, P. L. C., Stragier, P., Torrado, E., Van Aerde, A., Silva, M. T. and Pedrosa, J., 2008. First Cultivation and Characterization of *Mycobacterium ulcerans* from the Environment. *PLoS Neglected Tropical Diseases*, 2 (3), e178.

Ramakrishnan, L., Tran, H. T., Federspiel, N. A. and Falkow, S., 1997. A crtB homolog essential for photochromogenicity in Mycobacterium marinum: isolation, characterization, and gene disruption via homologous recombination. *Journal of Bacteriology*, 179 (18), 5862-5868.

Randolph, S. E. and Dobson, A. D. M., 2012. Pangloss revisited: a critique of the dilution effect and the biodiversity-buffers-disease paradigm. *Parasitology*, 139 (07), 847-863.

Roche, B., Benbow, M. E., Merritt, R., Kimbirauskas, R., McIntosh, M., Small, P. L. C., Williamson, H. and Guégan, J.-F., 2013a. Identifying the Achilles heel of multi-host pathogens: the concept of keystone 'host'species illustrated by *Mycobacterium ulcerans* transmission. *Environmental Research Letters*, 8 (4), 045009.

Roche, B. and Guégan, J.-F., 2011. Ecosystem dynamics, biological diversity and emerging infectious diseases. *Comptes Rendus Biologies*, 334 (5–6), 385-392.

Ross, B., Johnson, P., Oppedisano, F., Marino, L., Sievers, A., Stinear, T., Hayman, J., Veitch, M. and Robins-Browne, R., 1997. Detection of *Mycobacterium ulcerans* in environmental samples during an outbreak of ulcerative disease. *Applied and Environmental Microbiology*., 63 (10), 4135-4138.

Ruiz-Moreno, D., Pascual, M., Bouma, M., Dobson, A. and Cash, B., 2007. Cholera Seasonality in Madras (1901–1940): Dual Role for Rainfall in Endemic and Epidemic Regions. *EcoHealth*, 4 (1), 52-62.

Stinear, T., Ross, B. C., Davies, J. K., Marino, L., Robins-Browne, R. M., Oppedisano, F., Sievers, A. and Johnson, P. D. R., 1999. Identification and Characterization of IS2404 and IS2606: Two Distinct Repeated Sequences for Detection of *Mycobacterium ulcerans* by PCR. *Journal of Clinical Microbiology*, 37 (4), 1018-1023.

Stinear, T. P., Seemann, T., Pidot, S., Frigui, W., Reysset, G., Garnier, T., Meurice, G., Simon, D., Bouchier, C., Ma, L., Tichit, M., Porter, J. L., Ryan, J., Johnson, P. D. R., Davies, J. K., Jenkin, G. A., Small, P. L. C., Jones, L. M., Tekaia, F., Laval, F., Daffé, M., Parkhill, J. and Cole, S. T., 2007. Reductive evolution and niche adaptation inferred from the genome of *Mycobacterium ulcerans*, the causative agent of Buruli ulcer. *Genome Research*, 17 (2), 192-200.

Tilman, D., 1999. The Ecological Consequences Of Changes In Biodiversity: A Search For General Principles101. *Ecology*, 80 (5), 1455-1474.

Vandelannoote, K., Jordaens, K., Bomans, P., Leirs, H., Durnez, L., Affolabi, D., Sopoh, G., Aguiar, J., Phanzu, D. M. and Kibadi, K., 2014. Insertion Sequence Element Single Nucleotide Polymorphism Typing Provides Insights into the Population Structure and Evolution of *Mycobacterium ulcerans* across Africa. *Applied and Environmental Microbiology*, 80 (3), 1197-1209.

Van Oye, E. and Ballion, M., 1950. *Faudra-t-il tenir compte d'une nouvelle affection* a *bacilles* acido-resistants en Afrique? *Annales de la Société Belge de Médecine Tropicale*, **30**, 619–627.

Vincent Lévy-Frébault, V. and Portaels, F., 1992. Proposed Minimal Standards for the Genus Mycobacterium and for Description of New Slowly Growing Mycobacterium Species†. *International Journal of Systematic Bacteriology*, 42 (2), 315-323.

Williamson, H. R., Benbow, M. E., Nguyen, K. D., Beachboard, D. C., Kimbirauskas, R. K., McIntosh, M. D., Quaye, C., Ampadu, E. O., Boakye, D., Merritt, R. W. and Small, P. L. C., 2008. Distribution of *Mycobacterium ulcerans* in Buruli Ulcer Endemic and Non-Endemic Aquatic Sites in Ghana. *PLoS Neglected Tropical Diseases*, 2 (3), e205.

Willson, S. J., Kaufman, M. G., Merritt, R. W., Williamson, H. R., Malakauskas, D. M. and Benbow, M. E., 2013. Fish and amphibians as potential reservoirs of *Mycobacterium ulcerans*, the causative agent of Buruli ulcer disease. *Infection Ecology and Epidemiology*, (3).

Wolfe, N. D., Dunavan, C. P. and Diamond, J., 2007. Origins of major human infectious diseases. *Nature*, 447 (7142), 279-283.

World Health Organisation, 2001a. *Buruli Ulcer. Diagnosis of Mycobacterium ulcerans disease. A manual for health care providers*. Geneva. Available from: http://www.who.int/iris/handle/10665/67000#sthash.cioj71jV.dpuf.

World Health Organisation, 2001b. *Buruli Ulcer Management of Mycobacterium Ulcerans Disease*. Geneva. Available from: http://whqlibdoc.who.int/hq/2001/WHO_CDS_CPE_GBUI_2001.3.pdf.

Worobey, M., Gemmel, M., Teuwen, D. E., Haselkorn, T., Kunstman, K., Bunce, M., Muyembe, J.-J., Kabongo, J.-M. M., Kalengayi, R. M. and Van Marck, E., 2008. Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. *Nature*, 455 (7213), 661-664.

Yip, M. J., Porter, J. L., Fyfe, J. A. M., Lavender, C. J., Portaels, F., Rhodes, M., Kator, H., Colorni, A., Jenkin, G. A. and Stinear, T., 2007. Evolution of *Mycobacterium ulcerans* and Other Mycolactone-Producing Mycobacteria from a Common Mycobacterium marinum Progenitor. *Journal of Bacteriology*, 189 (5), 2021-2029.

Chapter 2 Site selection, field methodologies and site descriptions for French Guiana

2. 1 Initial survey to preliminary identify suitable sites for study

In total, 53 water-bodies were identified across French Guiana as potentially suitable for the study (table 2.1). These sites were located across the length of the territory, in both endemic and non-endemic regions. They were all similarly small in size and had no or very little water flow; large rivers, lakes and reservoirs were not considered due to the increased complexity of recording consistent biotic communities. Initially, environmental samples from 30 sites (sites 1-30 table 2.2) were taken and non quantitative PCR used to identify the presence of MU.

2.1.1 Environmental sample collection for sites 1-30

Samples included: water, soil/sediment, detritus and when present dominant aquatic plant species, algae, biofilms and samples of the semi-aquatic plant *Montrichardia arborescens* (Araceae), a plant species which is characteristic of Amazonian swamps (Table 3.1) These represent a range of previously described habitats for MU where positive samples have been identified in other continents (Marsollier et al. 2004a; Fyfe et al. 2007; Williamson et al. 2008). In certain cases when water samples were taken the location contained dense aquatic vegetation, biofilms were unavoidably disturbed from the leaves and stems and it was not possible to take a sample without also capturing these, it is clearly stated in cases where these samples are positive that they contain both water and biofilms.

Water was collected (50ml) from various sampling points at each site with an attempt to cover all meso-habitats, this included: bank-side, the centre of the water body, water from within aquatic vegetation, shaded areas and exposed areas. Water was taken to the laboratory and samples were filtered through 1.6µmglass

microfiber filters before being filtered through 0.4µm cellulose nitrate membrane and the residue collected and frozen at -20°C for future analysis.

2.1.2 DNA extraction.

DNA extraction was carried out using PowerSoil DNA extraction kit according to the manufacturers protocols (Mo Bio Lab., Carlsbad, USA), on all samples.

2.1.3 Non-quantitative PCR

Two primers were used to positively identify the bacteria (Williamson et al. 2008). The first primer pair amplifies an insertion sequence (IS2404) which is specific to MU and some other closely related mycolactone producing bacteria and is present in high copy numbers (Stinear et al. 1999). The second primer pair is part of the enoyl reductase region of the giant plasmid which codes for mycolactone (ER). Positive identification is not normally assumed until both primers test positive, however ER is present in much lower copy numbers of 1-2 per bacteria. Three sites tested positive for IS2404, two of which were water-filtrate and one was of algae. When tested with the ER primers however, all sites were negative. This could potentially be due to the lower copy numbers of ER being present and further samples may be needed (Williamson et al. 2012).

2.2 Second survey

2.2.1 Site identification

A further 21 sites (Sites 31-43 and 45-52 table 2.2) were tested for the presence of MU, environmental samples again included: water, soil/sediment and detritus and when present dominant aquatic plant species, algae, biofilms and samples of the semi-aquatic plant *Montrichardia arborescens* (Araceae). Samples were taken and processed in the same manner as detailed in chapter 2.1.1 Sites in this instance

were explicitly targeted by looking for water-bodies in areas where the disease was prevalent and had a high chance of human contact. Indications of this included: trails near or through the site, presence of fishing or boating activities or other recreational uses and close proximity to human settlement. In addition, similar sites in areas where there have been no cases of disease and sites which had little human contact were also identified. Preference was given to sites which remained present during the dry season, as this was indicative of permanence (See chapter 3).

DNA was extracted from these samples as detailed in chapter 2.1.2, however PCR was conducted using quantitative PCR (qPCR), the protocols of which are detailed in chapter 2.2.2. The qPCR analysis also included IS2404 positive samples from the previous survey in chapter 2.1, making a total of 163 samples from 23 sites. For detailed results of this survey see chapter 3.

From this information and the data from chapter 2.1, 18 sites were chosen for further study ("Analysis sites", Figure 2.2, see Appendix I for site pictures), these sites were typologically similar and represented a gradient of surrounding land use from pristine forest, through agricultural to heavily urbanised. In order to make comparisons more accurate, each site also had to be considered permanent (i.e. it did not dry up completely during the dry season). Two of the proposed sites had no environmental samples previously tested for MU however they were however the right typography for comparison to other sites. From the 18, one was subsequently removed from the analysis due to exceptionally low levels of recorded organisms during biotic surveys (n=8 from 2 surveys), potentially due to heavy pollution from an adjacent cement factory.

2.2.2 qPCR protocols

Oligonucleotide primer and TaqMan probe sequences were selected from the GenBank IS*2404* sequence (Stinear et al. 1999) and the ketoreductase B (KR) domain of the mycolactone polyketide synthase (mls) gene from the plasmid pMUM001 (Fyfe et al. 2007) (Table 2.1). 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 Low Rox (Agilent Technologies) in a total volume of 25 µl. Amplification and detection were performed with Thermocycler (Chromo 4, Bio-Rad) using the following program: 1 cycle of 50°C for 2 min, 1 cycle of 95°C for 15 min, 40 cycles of 95°C for 15 s and 60°C for 1 min. DNA extracts were tested at least in duplicates and the 10% negative controls were included in each assay. Quantitative readout assays were set up, based on an external standard curve with MU (strain 1G897), DNA was serially diluted over 5 logs (from 10^6 to 10^2 U/ml). Samples were only considered positive only if both the gene sequence encoding the ketoreductase B domain (KR) of the mycolactone polyketide synthase and IS*2404* sequence were detected, with threshold cycle (Ct) values <35 cycles)

Table 2.1 Primers and probes for real-time PCR detection. TF = forward primer; TR = reverse primer, TP = probe. Nucleotide position based on the first copy of the amplicon in pMUM001, accession numbers: IS2404: AF003002, KR: BX649209.

Primer or probe	Sequence (5'-3')	Nucleotide positions	Amplicon size (bp)
IS*2404* TF	AAAGCACCACGCAGCATCT	27746-27762	59
IS*2404* TR	AGCGACCCCAGTGGATTG	27787-27804	
IS*2404* TP	6 FAM CGTCCAACGCGATC MGBNFQ	27768-27781	
KR TF	TCACGGCCTGCGATATCA	3178-3195	65
KR TR	TTGTGTGGGCACTGAATTGAC	3222-3242	
KR TP	6 FAM-ACCCCGAAGCACTG-MGBNFQ	3199-3212	

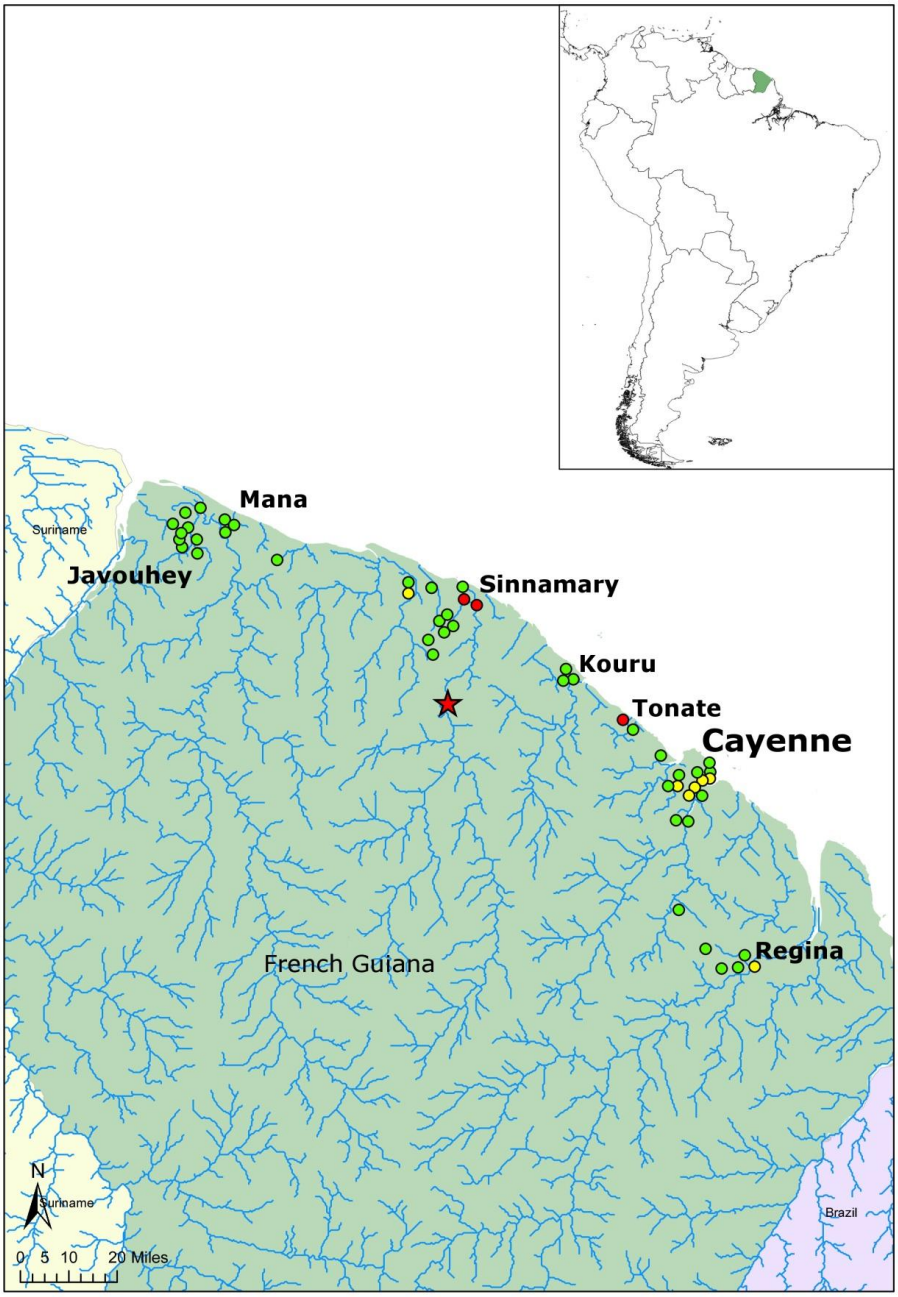


Figure 2.1 Map showing the 53 sampling locations. In green are the sites negative to both primers for MU, in yellow are the sites positive for IS2404 alone and in red are the sites positive for both IS2404 and KR.

2.3 Biotic and abiotic surveys of final analysis sites

2.3.1 Field method for sampling vertebrates and invertebrates.

Each "analysis site" was surveyed in the rainy season on two separate occasions during February and June of 2013. Sites were sampled for aquatic invertebrates, fish and other small aquatic vertebrates.

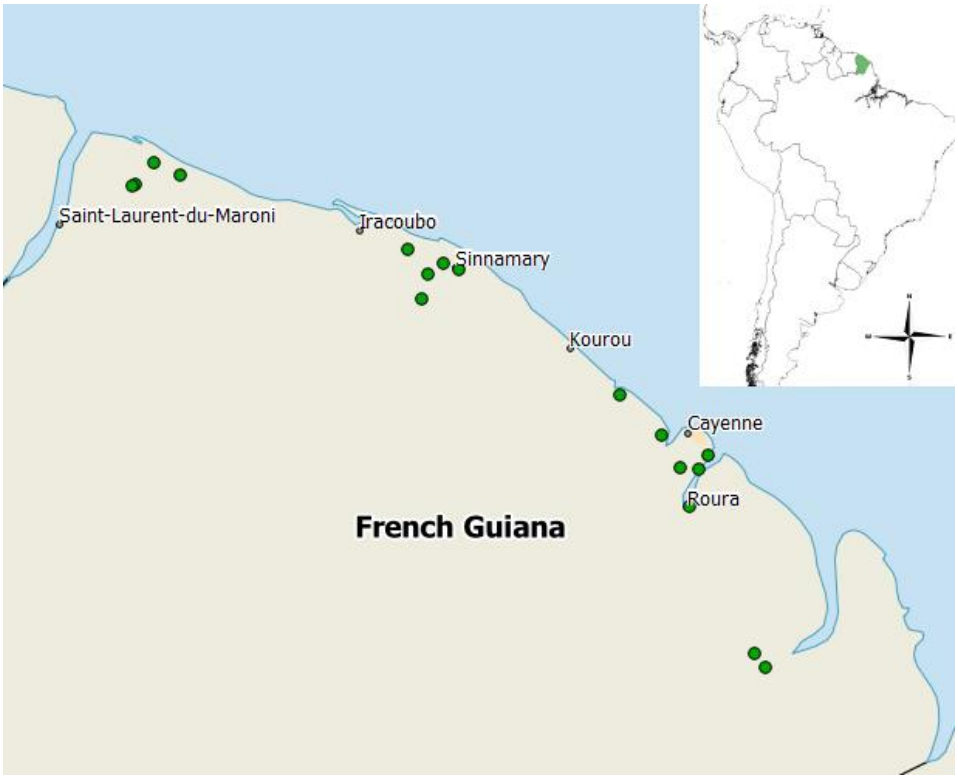


Figure 2.2 Map of French Guiana showing the location of all 17 analysis sites where detailed surveys were undertaken.

Invertebrate samples were taken using dip nets at four locations spatially spread within each site, to include a range of microhabitats. Each dip was carried out using a net 25cm x 45cm with a mesh size of 500 microns and consisted of vigorously sampling from the bottom substrate to the top of the water column within a meter squared area. All collected organisms were placed in a white tray for in-field preliminary sorting and were collected in 70% ethanol for identification and

storage. All substrate and remaining invertebrates were then collected in jars with 70% ethanol and taken back to the lab for final sorting and identification.

Fish were collected using three standard baited fish traps and spatially spread as with the dip net locations (see above). Fish nets comprised a rectangular netted tunnel with two funnelled openings, the trap was 50cm x 22cm x 22cm with an entrance diameter of 7cm, and mesh diameter was 0.3cm . Traps were left for 1 hour and 30 minutes before being recovered; all fish were collected and euthanized with an overdose of anaesthetic, before being preserved in 70% alcohol for sorting and identification in the lab. Fish were collected under a European license for animal experimentation held by Rodolphe Gozlan and adhered to French laws.

2.3.2 Abiotic readings

At each location within sites, near to where fish and invertebrates were collected, abiotic readings were taken; these included: dissolved oxygen, temperature, *pH* and conductivity. The readings were taken with a short time delay before fish and invertebrate sampling, so as to keep disturbance to a minimum (Table 2.3).

2.3.3 Organism identification cataloguing and measurement

Each invertebrate was measured in length and identification was conducted predominantly to family level where possible, using a number of literature sources (Merritt and Cummins 1996; Domínguez 2006; Heckman 2006, 2008, 2011). Organisms in the same family were pooled by site and collection date, before being stored in 70% alcohol for qPCR analysis, any invertebrate containing more than 10 organisms were pooled in series' of 10. Dry weights in milligram were calculated using length to mass regression equations from a number of literature sources (Smock 1980; Benke et al. 1999; Miserendino 2001). Fish were identified down to species (Keith et al. 2000a; Keith et al. 2000b) and for each individual a standard

length measurement was taken. Weights were calculated using length to weight regression equations (Froese et al. 2014), as there is no direct conversion in the literature for fish to dry weights for all species, an average wet to dry weight conversion of 20% was used. For qPCR fish were pooled in maximum groups of three per species per sample (dip net or trap).

2.3.4 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 $\delta^{13}C$ and $\delta^{15}N$. This covered the majority of taxa present throughout the 17 sites. For the invertebrate analysis a limb of the organism was taken, or if too small and many were present, whole organisms. For fish, caudal fin clippings were taken. A minimum amount of material was used in order to preserve organisms for qPCR analysis. In most cases three samples of each taxonomic group were analysed for each site, however due to limits in the number of available organisms, in some cases one or two samples were used. For any taxa that were unable to be analysed due to abundance limitations, stable isotope data from a closely related and functionally, highly similar taxon was recorded for use in the statistical analysis (See Appendix III, Table S6.1). The material was dried at 40°C for 8 hours to remove all traces of moisture. After drying the material was homogenised and the stable isotope signatures of $\delta^{13}C$ and $\delta^{15}N$ were identified using a Finnigan MAT Delta Plus isotope ratio mass spectrometer.

2.3.5 DNA extraction for invertebrates and vertebrates

For biotic samples (invertebrates and vertebrates), pooled organisms were ground together and homogenized in 50 mM NaOH solution using a QIAGEN Tissue Lyser II. Homogenates of the tissue were then heated at 95°C for 20 min. DNA

from the homogenized tissue was then purified using QIAquick 96 PCR Purification Kit (QIAGEN), according to manufacturer's recommendations. 10% negative controls were included for extraction and purification. qPCR was performed as detailed in chapter 2.2.

Table 2.2 Table showing the site locations, and the initial environmental survey sample data, this includes both qPCR and PCR samples. Some field site numbers are missing due to the sites being recorded during the early searches for locations but considered unsuitable for further analysis. Sites used for subsequent biotic surveys are highlighted in bold.

Field Site No.	Analysis Site No.	Northing	Easting	Location	Total Samples	Soil	Water	Algae/Biofilms	Plant material
1	-	5.62043	-53.6939	Mana	2	0	1	0	0
2	**1**	5.630706	-53.7081	**Mana**	**2**	**0**	**0**	**0**	**0**
3	-	5.594756	-53.7072	Between Mana and Iracoubo	2	0	1	0	0
5	-	4.887111	-52.2641	Lac de Rémire-Montjoly	2	0	1	0	0
6	-	4.894368	-52.2585	Route des plages Montjoly	6	2	1	3	0
7	**2**	4.860284	-52.2572	**Montjoly**	**12**	**1**	**4**	**4**	**1**
8	-	4.858214	-52.2749	Montjoly	6	2	2	2	0
9	-	4.728838	-52.3244	Roura	4	1	1	1	0
10	**3**	4.737259	-52.327	**Roura**	**2**	**1**	**0**	**0**	**0**
11	**4**	4.832641	-52.3486	**Matoury**	**17**	**4**	**5**	**4**	**3**
12	-	4.832641	-52.3486	Matoury	2	1	1	0	0
15	-	4.297617	-52.1748	Régina (on road to)	1	0	0	0	0
16	-	4.294523	-52.2192	Régina (on road to)	1	1	0	0	0
17	-	4.35743	-52.2693	Régina (on road to)	1	0	1	0	0
18	-	4.470109	-52.3499	Route de l'Este (road to Régina)	1	0	0	0	0
19	-	5.298387	-53.0501	Piste St Elie	1	1	0	0	0
20	-	5.298387	-53.0501	Piste St Elie (further down road)	2	1	0	1	0
21	**5**	5.317791	-53.045	**Piste St Elie**	**2**	**2**	**0**	**0**	**0**
22	-	5.317791	-53.045	Piste St Elie	1	0	0	1	0
23	**6**	5.340026	-53.036	**Road to Piste St Elie**	**1**	**0**	**0**	**0**	**0**
24	-	5.36331	-53.0327	Javouhey	2	0	1	0	1
25	-	5.610736	-53.8173	Javouhey	2	0	1	0	0

26	-	5.610621	-53.8202	Javouhey	2	0	1	0	1
27	-	5.609506	-53.8238	Javouhey	2	1	1	0	0
28	**7**	5.609506	-53.8238	**Javouhey**	**3**	**0**	**1**	**0**	**1**
29	-	5.605557	-53.8279	Javouhey	4	1	1	0	1
31	-	5.629661	-53.8707	Kourou	7	0	6	0	1
32	-	5.155477	-52.6641	Kourou	5	0	3	2	0
33	-	5.173087	-52.6588	Kourou	5	1	2	1	1
34	**8**	5.177678	-52.6615	**Matoury**	**21**	**6**	**7**	**5**	**3**
35	-	4.833778	-52.3018	Matoury	12	3	8	1	0
36	-	4.833778	-52.3018	Matoury	6	1	2	1	0
37	-	4.832085	-52.299	Matoury	8	2	3	3	0
38	**9**	4.83778	-52.3498	**Route JoJo**	**6**	**0**	**4**	**2**	**0**
39	-	5.392771	-52.9934	Route **JoJo**	4	0	4	0	0
40	-	5.403585	-52.9959	Halfway between Kourou and Iracoubo crossroads	7	1	5	1	0
41	**10**	5.428044	-53.0851	**Roura**	**12**	**4**	**7**	**1**	**0**
42	-	5.428044	-53.0851	Tonate	5	1	4	0	0
43	**11**	5.008112	-52.4888	**Sinnamary**	**18**	**5**	**11**	**1**	**0**
44	**12**	5.376385	-52.9541	**Régina**	**0**	**0**	**0**	**0**	**0**
45	**13**	4.331787	-52.1533	**Régina**	**6**	**3**	**2**	**1**	**0**
46	**14**	4.298854	-52.1413	**Near Tonate**	**5**	**0**	**3**	**2**	**0**
47	**15**	5.033737	-52.5166	**Macouria**	**7**	**0**	**5**	**2**	**0**
48	-	4.929222	-52.4038	Montjoly	3	0	3	0	0
49	**16**	4.858214	-52.2749	**Mana**	**8**	**2**	**6**	**0**	**0**
50	-	5.667646	-53.7796	Mana	4	0	3	1	0
51	-	5.651339	-53.8226	Near Iracoubo	5	1	3	1	0
52	-	5.443163	-53.1585	Near Iracoubo	5	2	3	0	0
53	**17**	5.443163	-53.1585	**Javouhey**	**0**	**0**	**0**	**0**	**0**

Table 2.3 Mean abiotic data readings for each analysis site, these are the mean from 3 readings taken near to the invertebrate and vertebrate survey locations.

Site	Mean Conductivity µS/cm	Mean Temperature °C	Mean Dissolved Oxygen mg/L	Mean *pH*
1	87.32	27.22	0.86	5.60
2	359.71	30.68	0.79	5.62
3	508.86	26.10	5.26	4.74
4	27.57	25.64	4.76	5.07
5	30.43	24.85	1.79	4.55
6	29.08	25.50	4.59	5.30
7	67.52	25.30	0.90	5.50
8	103.03	28.67	4.35	6.75
9	31.96	27.34	2.43	5.14
10	37.29	25.01	0.54	5.06
11	198.74	24.85	0.67	5.49
12	13.94	25.08	3.31	5.02
13	17.66	26.67	1.33	5.31
14	34.01	25.64	1.54	5.09
15	519.20	29.95	0.65	6.80
16	335.16	27.07	2.64	5.81
17	37.99	26.78	2.83	5.08

2.4 References

Stinear, T., Ross, B. C., Davies, J. K., Marino, L., Robins-Browne, R. M., Oppedisano, F., Sievers, A. and Johnson, P. D. R., 1999. Identification and Characterization of IS2404 and IS2606: Two Distinct Repeated Sequences for Detection of *Mycobacterium ulcerans* by PCR. *Journal of Clinical Microbiology*, 37 (4), 1018-1023.

Williamson, H. R., Benbow, M. E., Campbell, L. P., Johnson, C. R., Sopoh, G., Barogui, Y., Merritt, R. W. and Small, P. L. C., 2012. Detection of *Mycobacterium ulcerans* in the Environment Predicts Prevalence of Buruli Ulcer in Benin. *PLoS Neglected Tropical Diseases*, 6 (1), e1506.

Williamson, H. R., Benbow, M. E., Nguyen, K. D., Beachboard, D. C., Kimbirauskas, R. K., McIntosh, M. D., Quaye, C., Ampadu, E. O., Boakye, D., Merritt, R. W. and Small, P. L. C., 2008. Distribution of *Mycobacterium ulcerans* in Buruli Ulcer Endemic and Non-Endemic Aquatic Sites in Ghana. *PLoS Neglected Tropical Diseases*, 2 (3), e205.

Chapter 3 First detection of *Mycobacterium ulcerans* DNA in environmental samples from South America

Reference: First Detection of *Mycobacterium ulcerans* DNA in Environmental Samples from South America. *PLoS Neglected Tropical Diseases*, 8 (1), e2660.

See Appendix IV for full publication.

3.1 Summary

Although MU has been previously diagnosed in South America, until now the presence of its DNA in the wild has only been identified in Australia where there have been significant outbreaks and in western and central regions of Africa where the disease is persistent. Here for the first time, the DNA has been identified in environmental samples from South America. The DNA was positively identified using qPCR on 163 environmental samples, taken from 23 freshwater bodies in French Guiana, using primers for both IS2404 and for the ketoreductase-B domain of the MU mycolactone polyketide synthase genes (KR). Five samples out of 163 were positive for both primers from three different water bodies. A further nine sites had low levels of IS2404 close to a standard CT of 35 and could potentially harbour MU. The majority of the positive samples (8/14) came from filtered water. These results also reveal the Sinnamary River as a potential source of infection to humans despite a decline in cases in the 18 years since the Petit-Saut Dam construction was completed.

3.2 Introduction

Since the discovery of specific PCR primers, which are sensitive enough to detect MU from the environment, it has been found on, or within numerous aquatic species and in environmental samples from aquatic systems (Marsollier et al. 2002; Marsollier et al. 2004a; Marsollier et al. 2004b; Mosi et al. 2008; Merritt et al. 2010), see chapter 1.3.4. As the current knowledge of MU's life history remains inconclusive however, it is not possible to definitely advocate the optimum aquatic conditions necessary for the disease to flourish. It is possible however speculate on preferred habitats, based on the current evidence and physiological traits. As discussed in chapter 1.3.2, sequencing of the genome reveals a lack of *crt*I which is responsible for the production of phytoene dehydrogenase, an enzyme that within *M. marinum* is necessary for the synthesis of light-inducible carotenoid pigments (Ramakrishnan et al. 1997). The loss of need for these pigments, which give protection against UV-induced damage, suggests that MU lives in conditions where it does not require this ability. Culturing *in vitro* has also shown the mycobacterium to have a preference for a low oxygen environment (Palomino et al. 1998). These traits therefore indicate that the mycobacteria have a preference for environments with low light and oxygen levels.

Identification of the bacteria in the environment has been generally isolated to parts of Australia where there have been significant outbreaks (Fyfe et al. 2007; Johnson et al. 2007; Fyfe et al. 2010) or tropical regions of Africa where the disease is persistent (Marston et al. 1995; Marsollier et al. 2002; Debacker et al. 2004; Mosi et al. 2008; Portaels et al. 2008; Williamson et al. 2008; Williamson et al. 2012), see chapter 1.3.1 and 1.3.4. Whilst in South America human cases of the disease have been definitively present since 1969 as for example in French Guiana (Guerra et al. 2008; McGann et al. 2009; Merritt et al. 2010) and previous data suggests the presence of MU DNA in environmental water sources (Guerra et al. 2008) also see

chapter 1.3.1, the DNA has never before been definitively identified in the environment. Therefore this chapter aims to monitor a range of freshwater habitats to detect MU's presence from environmental samples in South America. Further to this an analysis was undertaken to test if the Petit-Saut Dam, a hydro-electric installation upstream from an endemic BU area of French Guiana had in the years following its construction impacted the number of cases downstream near the city of Sinnamary.

3.3 Materials and Methods

3.3.1 Site identification and sampling methods

Details of French Guiana can be found in chapter 1.3. Methods for site identification were as detailed in chapter 2.2.1. For a detailed description of environmental sampling and protocols for DNA extraction and qPCR see chapter 2.2.2. Abiotic readings: *pH,* temperature, dissolved oxygen and conductivity were recorded near to where each environmental sample was taken within each site.

3.3.2 Statistical Analysis

To help identify the biological niche of MU in French Guiana a series of statistical tests were to identify whether the DNA was predominantly present within certain abiotic conditions and whether the building of the Petit-Saut dam had a significant influence on the number of downstream cases.

3.3.2.1 Number of cases since the building of Petit-Saut dam

To account for non-normal distributions of data, a Wilcoxon-rank sum test was performed on the number of cases in Sinnamary district per month, per 100,000 people for the 18 years since the Dam was constructed and impoundment of water

started in January 1994 against the 18 years prior to this period. This was repeated for comparison for cases across the whole of French Guiana.

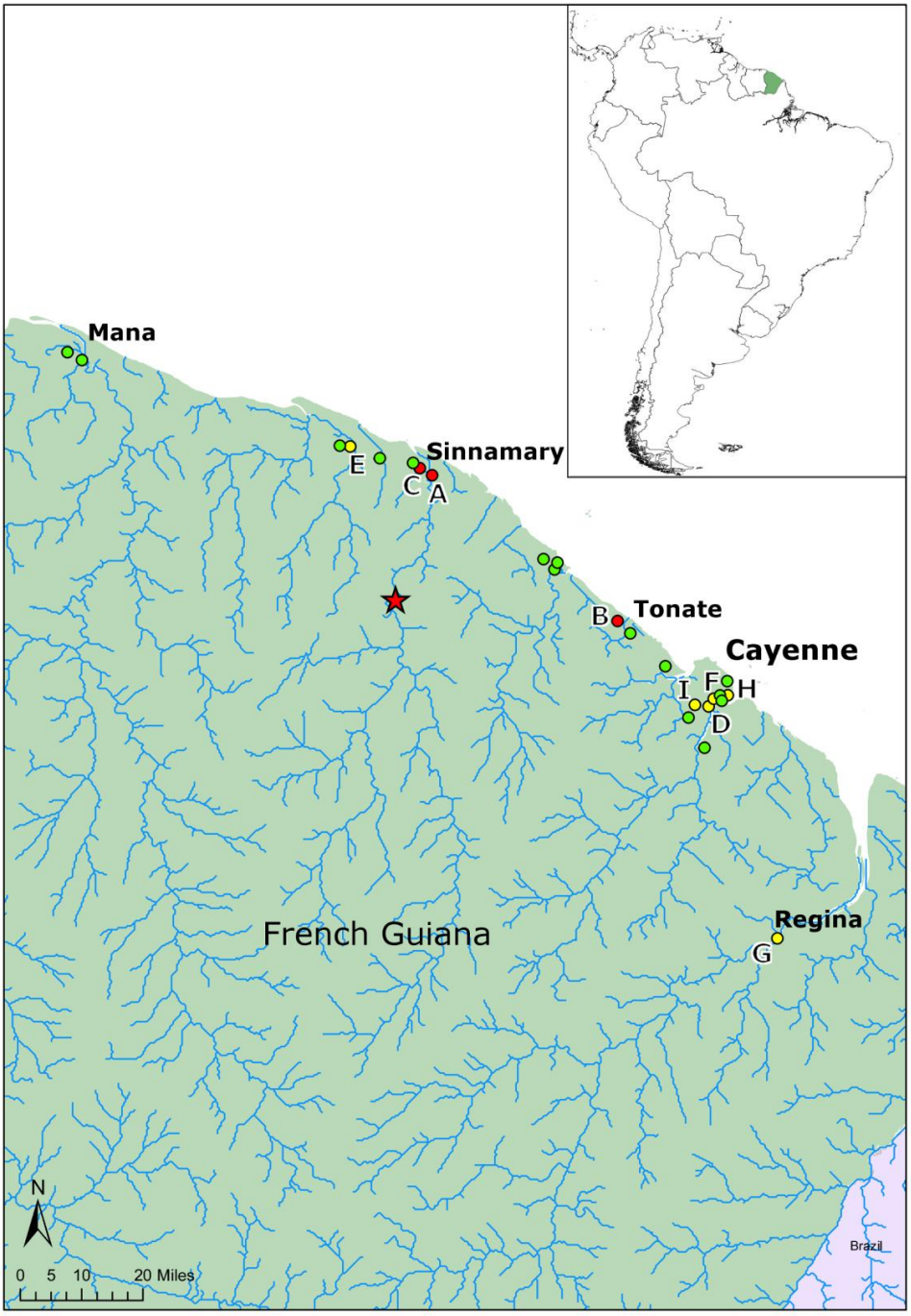


Figure 3.1 Map showing the approximate location of the twenty-three sampling sites (Text S1). In green are the negative sites for both MU during this survey, in yellow are the site positive for IS2404 alone and in red are the sites positive to both IS2404 and KR. The Petit-Saut dam built in 1994 is indicated by a star (Lehner et al. 2006).

3.3.2.2 Differences between abiotic parameters at sites

Wilcoxon tests were further used to determine statistical differences between abiotic measurements (*pH*, conductivity, dissolved oxygen and water temperature) at MU positive sites against the same measurements at negative sites. These tests were also repeated to identify any statistical differences in abiotic factors between sites which were positive for MU.

3.4 Results

3.4.1 qPCR

Five samples out of 163 were identified as positive for both IS2404 and KR (CT<35) from three different water bodies in French Guiana (table 3.1 and Figure 3.1); Site A (lat 5.3772, long -52.953883) in Sinnamary (1/19 samples positive), site B (lat 5.3941, long -52.992017) near Tonate (3/5 samples positive) and site C (lat 5.03535, long -52.516483) on Route JoJo a road just outside Sinnamary (1/6 samples positive) (see Table 3.2). A further nine sites had low levels of IS2404 close to a standard CT of 35 (Table 3.3) and could potentially harbour MU, however it is not possible to conclude this definitively as they were negative for KR.

All three sites (five samples) positive for both IS2404 and KR were typologically similar, with areas that were highly stagnant forming shallow water bodies and with high levels of *M. arborescens* plant growth (Table 3.2). Dissolved oxygen levels at sites A and B, at the locations where the mycobacterium was found were low (<1mg/l), at site C dissolved oxygen was higher (1.93 mg/l); however other locations within the site and within a few meters were also similarly low to sites A and C (<1mg/l). There was no statistical significance for differences in *pH*, conductivity, dissolved oxygen and temperature between positive sites and negative sites.

Table 3.1 Site locations and the number and type of samples taken. In bold both IS2404 and KR identified.

Site	Latitude	Longitude	Location	IS2404	KR	Total Samples	Soil	Water	Algae/ Biofilms	Plant Material	*M. arborescens*	Detritus	Insect
43	**5.3772**	**-52.953883**	**Sinnamary**	**+**	**+**	**19**	**3**	**13**	**1**	**-**	**-**	**1**	**1**
46	**5.03535**	**-52.516483**	**Nr Tonate**	**+**	**+**	**5**	**-**	**3**	**2**	**-**	**-**	**-**	**-**
38	**5.3941**	**-52.992017**	**Route Jojo**	**+**	**+**	**6**	**-**	**6**		**-**	**-**	**-**	**-**
35	4.834467	-52.3021	Matoury	+	-	6	-	6	-	-	-	-	-
52	5.44505	-53.158183	Nr Iracoubo	+	-	5	2	3	-	-	-	-	-
34	4.83365	-52.3004	Matoury	+	-	18	4	8	3	-	-	3	-
45	4.300417	-52.13995	Regina	+	-	6	3	2	1	-	-	-	-
7	4.8608	-52.25675	Montjoly	+	-	9	1	3	2	1	1	1	-
11	4.838083	-52.35325	Matoury	+	-	15	4	4	3	3	1	-	-
49	5.6666	-53.7799	Mana	-	-	8	2	6	-	-	-	-	-
50	5.652183	-53.825	Mana	-	-	4	-	3	1	-	-	-	-
51	5.44505	-53.165	Nr Iracoubo	-	-	5	1	3	1	-	-	-	-
40	5.428733	-53.088717	Nr. Sinnamary	-	-	3	-	3	-	-	-	-	-
41	4.729714	-52.318397	Roura	-	-	12	4	7	1	-	-	-	-
31	5.1561	-52.665233	Kouru	-	-	3	-	3	-	-	-	-	-
32	5.172417	-52.658467	Kouru	-	-	5	-	3	2	-	-	-	-
33	5.180617	-52.66165	Kouru	-	-	4	1	2	1	-	-	-	-
39	5.403217	-52.99575	Route Jojo	-	-	1	-	1	-	-	-	-	-
42	5.00615	-52.4869	Tonate	-	-	5	1	4	-	-	-	-	-
47	4.929067	-52.403817	Macouria	-	-	7	-	5	2	-	-	-	-
6	4.893783	-52.257883	Monjoly	-	-	3	-	3	-	-	-	-	-
36	4.833183	-52.2993	Matoury	-	-	5	1	2	-	-	-	2	-
37	4.837767	-52.349483	Matoury	-	-	6	2	2	2	-	-	-	-
48	4.860267	-52.2753	Monjoly	-	-	3	-	3	-	-	-	-	-

Table 3.2: Details of samples with sites positive for IS2404 (CT>35). Abiotic parameters taken from where the sample was collected within the water body is also included. Bacteria per ml of 50ml water sample, or 0.25grams of solid sample.

Site	Sample Type	Locality	CT Value (IS204)	CT Value (KR)	Bact/ml	Dissolved Oxygen (mg/l)	*pH*	Conductivity (µS/m)	Water Temp (°C)	Season
43	Water & Biofilms	Sinnamary	34.70	33.69	18.33	0.74	5.55	197	24.9	Dry
46	Water filtrand	Nr Tonate	31.93	33.48	138.9	0.93	6.237	150.8	28.9	Dry
46	Water filtrand	Nr Tonate	29.77	30.92	674.6	0.93	6.237	150.8	28.9	Dry
46	Water filtrand	Nr Tonate	32.11	34.5	121.8	0.03	5.896	92.4	26.4	Dry
48	Water filtrand	Route JoJo	34.14	34.47	27.49	1.93	5.342	32	26.8	Dry

Table 3.3: Details of positive samples for both IS2404 and KR (CT<35). Abiotic parameters taken from where the sample was collected within the water body is also included. Mycobacteria per ml of 50ml water sample, or 0.25grams of solid sample.

Site	Sample Type	Locality	CT Value	Bact/ml	Dissolved Oxygen (mg/l)	*pH*	Conductivity (µS/m)	Water Temp (°C)	Season
43	Water filtrand	Sinnamary	36.2	10.53	0.46	5.3	201	24.7	Dry
52	Soil/sediment	Nr Iracoubu	35.41	17.9	NA	NA	NA	NA	Dry
52	Soil/sediment	Matoury	37.2	4.877	4.94	6.764	99.50	28.8	Dry
34	Water filtrand	Matoury	37.43	2.48	0.05	6.442	14.32	26.6	Dry
45	Water filtrand	Regina	36.33	5.553	3.23	5.50	15.6	27.3	Dry
45	Water filtrand	Regina	37.39	2.56	0.95	5.35	13.5	26.0	Dry
45	Soil/sediment	Regina	35.96	7.297	1.26	5.27	13.9	26.2	Dry
7	*M. arborescens* skin	Monjoly	35.53	15.39	0.57	5.62	280	30	Dry
11	Filamentous algae	Matoury	36.61	7.33	4.3	4.87	34.8	25.6	Dry

3.4.2 Number of cases since the building of Petit-Saut Dam

The number of cases in the 18 years after the construction of the Petit-Saut Dam were significantly lower in Sinnamary with 0.6 cases per 100,000 people per year when compared to 10.1 cases before (Test statistic (W) = 210, p-value = 0.0296). The number of cases in the whole of French Guiana remained statistically similar before and after construction (Test statistic (W) = 132.5, p-value = 0.359).

3.5 Discussion

It is difficult yet to draw definitive conclusions about potential abiotic or biotic factors affecting mycobacterium levels in French Guiana from the current information, whilst the results showed no statistical differences. This may be because of too few positive sites to draw reliable statistical comparisons. The results extend the range of geographical distribution of environmental MU to another continent, i.e. south America, where previously the DNA has not been identified in the environment. Similarly it was found that the majority of our positive samples came from filtered water (8/14), which has been the case in other countries (Williamson et al. 2008). Whilst it could be concluded that the mycobacteria are more prevalent in the water column, this may be an artefact of the sampling methods used. When sampling water it is possible to concentrate a large volume onto filters for extraction, whereas with soil, biofilms and plants, the limitations of the extraction kits mean only a relatively small fraction of material from a site (0.25g) can be utilised. In addition the freshwater bodies where positive samples were found, in this area were on floodplains, suggesting this category of environment constitutes a source for environmentally-persistent mycobacteria or a "receptacle" concentrating these bacilli from further upstream. From the results it is possible to also recommend the importance of taking multiple abiotic readings from a single site, because the ecology of an organism that is living in a

microscopic environment is being assessed, the large variation in abiotic parameters within a few meters or less in the same water-body must be considered (for example Site B, Tables 2 and 3).

The low levels of cases combined with a low population density in a territory such as French Guiana suggested that the identification of MU from the environment would be difficult. However, it was possible to find positive sites with less than 200 samples, which in themselves were relatively small components of a system. This would reinforce the possibility that MU is a fairly common and widely distributed mycobacterium, and it is other factors, e.g. socio-economic (i.e. levels of human contact with water, sanitation etc), transmission related (i.e. presence of potential vectors, hosts or reservoirs), or habitat modifications (deforestation, dam construction, etc), that might be the primary drivers of cases.

The finding of the majority of positive sites from the area around Sinnamary River downstream to the dam is of interest, the first cases of BU in French Guiana were recorded here and approximately 10% of human cases of BU in French Guiana concern the inhabitants of the Sinnamary. In the Sinnamary region the number of cases has been very low since 1994, with significantly less cases in the 18 years after 1994, despite an increasing human population. Whilst changes in the behaviour of people may have had an influence, the building of the Petit-Saut Dam (Figure 3.2) may also have been playing a role. The dam has profound effects on the level of water which comes into the area, possibly reducing flooding or regulating water flows, or potentially limiting mycobacteria being brought upstream from the rainforest and riverine swamp areas.

3.6 References

Brou, T., Broutin, H., Elguero, E., Asse, H. and Guegan, J-F., 2008. Landscape Diversity Related to Buruli Ulcer Disease in Côte d'Ivoire. *PLoS Neglected Tropical Diseases*, 2 (7), e271.

Debacker, M., Aguiar, J., Steunou, C., Zinsou, C., Meyers, W. M., Guédénon, A., Scott, J. T., Dramaix, M. and Portaels, F., 2004. *Mycobacterium ulcerans* Disease (Buruli ulcer) in Rural Hospital, Southern Benin, 1997–2001. *Emerging Infectious Diseases*, 10 (8).

Fyfe, J. A. M., Lavender, C. J., Handasyde, K. A., Legione, A. R., O'Brien, C. R., Stinear, T. P., Pidot, S. J., Seemann, T., Benbow, M. E., Wallace, J. R., McCowan, C. and Johnson, P. D. R., 2010. A Major Role for Mammals in the Ecology of *Mycobacterium ulcerans*. *PLoS Neglected Tropical Diseases*, 4 (8), e791.

Fyfe, J. A. M., Lavender, C. J., Johnson, P. D. R., Globan, M., Sievers, A., Azuolas, J. and Stinear, T. P., 2007. Development and Application of Two Multiplex Real-Time PCR Assays for the Detection of *Mycobacterium ulcerans* in Clinical and Environmental Samples. *Applied and Environmental Microbiology*, 73 (15), 4733-4740.

Guerra, H., Palomino, J. C., Falconí, E., Bravo, F., Donaires, N., Van Marck, E. and Portaels, F., 2008. *Mycobacterium ulcerans* disease, Peru. *Emerging infectious diseases*, 14 (3), 373.

Johnson, P. D. R., Azuolas, J., Lavender, C. J., Wishart, E., Stinear, T. P., Hayman, J. A., Brown, L., Jenkin, G. A. and Fyfe, J. A. M., 2007. *Mycobacterium ulcerans* in Mosquitoes Captured during Outbreak of Buruli Ulcer, Southeastern Australia. *Emerging Infectious Diseases*, 13 (11).

Lehner, B., Verdin, K. and Jarvis, A., 2006. *HydroSHEDS*: World Wildlife Fund US, Washington, DC. Available from: http://hydrosheds.cr.usgs.gov/index.php

Marion, E., Eyangoh, S., Yeramian, E., Doannio, J., Landier, J., Aubry, J., Fontanet, A., Rogier, C., Cassisa, V., Cottin, J., Marot, A., Eveillard, M., Kamdem, Y., Legras, P., Deshayes, C., Saint-André, J.-P. and Marsollier, L., 2010. Seasonal and Regional Dynamics of *M. ulcerans* Transmission in Environmental Context: Deciphering the Role of Water Bugs as Hosts and Vectors. *PLoS Neglected Tropical Diseases*, 4 (7), e731.

Marsollier, L., Robert, R., Aubry, J., Saint Andre, J.-P., Kouakou, H., Legras, P., Manceau, A.-L., Mahaza, C. and Carbonnelle, B., 2002. Aquatic Insects as a Vector for *Mycobacterium ulcerans*. *Applied and Environmental. Microbiology.*, 68 (9), 4623-4628.

Marsollier, L., Stinear, T., Aubry, J., Saint André, J. P., Robert, R., Legras, P., Manceau, A.-L., Audrain, C., Bourdon, S., Kouakou, H. and Carbonnelle, B., 2004a. Aquatic Plants Stimulate the Growth of and Biofilm Formation by *Mycobacterium ulcerans* in Axenic Culture and Harbor These Bacteria in the Environment. *Applied and Environmental Microbiology*, 70 (2), 1097-1103.

Marsollier, L., Sévérin, T., Aubry, J., Merritt, R. W., Saint André, J.-P., Legras, P., Manceau, A.-L., Chauty, A., Carbonnelle, B. and Cole, S. T., 2004b. Aquatic Snails, Passive Hosts of *Mycobacterium ulcerans*. *Applied and Environmental Microbiology*, 70 (10), 6296-6298.

Marston, B. J., Diallo, M. O., Horsburgh, C. R., Diomande, I., Saki, M. Z., Kanga, J. M., Patrice, G., Lipman, H. B., Ostroff, S. M. and Good, R. C., 1995. Emergence of Buruli ulcer disease in the Daloa region of Cote d'Ivoire. *The American journal of tropical medicine and hygiene*, 52 (3), 219-224.

McGann, H., Stragier, P., Portaels, F., Gascoyne-Binzi, D., Collyns, T., Lucas, S. and Mawer, D., 2009. Buruli ulcer in United Kingdom tourist returning from Latin America. *Emerging infectious diseases*, 15 (11), 1827.

Merritt, R. W., Benbow, M. E. and Small, P. L. C., 2005. Unraveling an Emerging Disease Associated with Disturbed Aquatic Environments: The Case of Buruli Ulcer. *Frontiers in Ecology and the Environment*, 3 (6), 323-331.

Merritt, R. W., Walker, E. D., Small, P. L. C., Wallace, J. R., Johnson, P. D. R., Benbow, M. E. and Boakye, D. A., 2010. Ecology and Transmission of Buruli Ulcer Disease: A Systematic Review. *PLoS Neglected Tropical Diseases*, 4 (12), e911.

Mosi, L., Williamson, H., Wallace, J. R., Merritt, R. W. and Small, P. L. C., 2008. Persistent Association of *Mycobacterium ulcerans* with West African Predaceous Insects of the Family Belostomatidae. *Applied and Environmental Microbiology*, 74 (22), 7036-7042.

Palomino, J. C., Obiang, A. M., Realini, L., Meyers, W. M. and Portaels, F., 1998. Effect of Oxygen on Growth of *Mycobacterium ulcerans* in the BACTEC System. *Journal of Clinical Microbiology*, 36 (11), 3420-3422.

Portaels, F., Meyers, W. M., Ablordey, A., Castro, A. G., Chemlal, K., de Rijk, P., Elsen, P., Fissette, K., Fraga, A. G., Lee, R., Mahrous, E., Small, P. L. C., Stragier, P., Torrado, E., Van Aerde, A., Silva, M. T. and Pedrosa, J., 2008. First Cultivation and Characterization of *Mycobacterium ulcerans* from the Environment. *PLoS Neglected Tropical Diseases*, 2 (3), e178.

Ramakrishnan, L., Tran, H. T., Federspiel, N. A. and Falkow, S., 1997. A crtB homolog essential for photochromogenicity in Mycobacterium marinum: isolation, characterization, and gene disruption via homologous recombination. *Journal of Bacteriology*, 179 (18), 5862-5868.

Williamson, H. R., Benbow, M. E., Campbell, L. P., Johnson, C. R., Sopoh, G., Barogui, Y., Merritt, R. W. and Small, P. L. C., 2012. Detection of *Mycobacterium ulcerans* in the Environment Predicts Prevalence of Buruli Ulcer in Benin. *PLoS Neglected Tropical Diseases*, 6 (1), e1506.

Williamson, H. R., Benbow, M. E., Nguyen, K. D., Beachboard, D. C., Kimbirauskas, R. K., McIntosh, M. D., Quaye, C., Ampadu, E. O., Boakye, D., Merritt, R. W. and Small, P. L. C., 2008. Distribution of *Mycobacterium ulcerans* in Buruli Ulcer Endemic and Non-Endemic Aquatic Sites in Ghana. *PLoS Neglected Tropical Diseases*, 2 (3), e205.

Chapter 4 Complex temporal climate signals drive the emergence of human water-borne disease.

Reference: Morris, A., Gozlan, R. E., Hassani, H., Andreou, D., Couppié, P. and Guégan, J.-F., 2014b. Complex temporal climate signals drive the emergence of human water-borne disease. *Emerging Microbes and Infections*, 3 (8), e56

See appendix V for full publication

4.1 Summary

Like many tropical diseases, associations with climate have been previously observed and could help identify the causative agent's ecological niche. In this chapter links between changes in rainfall and outbreaks of BU in French Guiana and Akonolinga a central district in the central African country Cameroon, were identified using a combination of statistical tests based on singular spectrum analysis, empirical mode decomposition and cross wavelet coherence analysis. From this it was possible to postulate for the first time that outbreaks of BU can be triggered by combinations of rainfall patterns occurring on a long (i.e. several years) and short (i.e. seasonal) temporal scale, in addition to stochastic events driven by the El Niño Southern Oscillation that may disrupt or interact with these patterns. Seasonal patterns were common to both French Guiana and Akonlinga with both exhibiting two peaks in disease cases corresponding to two rainy seasons. Long-term forecasting of rainfall trends further suggests the possibility of an upcoming outbreak of BU in French Guiana.

4.2 Introduction

The identification of cohering patterns between climate and infectious disease using time series analysis is an important component in understanding the ecological niche of disease causing agents, and in predicting future outbreaks. Such correlations can occur with both local and large scale climatic oscillations (Pascual et al. 2000; Gagnon et al. 2001; Rodó et al. 2002; Zhou et al. 2004; Chaves and Pascual 2006; Pascual et al. 2008; Hanf et al. 2011; Chaves et al. 2012; Mahamat et al. 2013). The mechanisms behind these relationships often vary and have been attributed both to the direct and indirect effects of changing climate, notably for vector-borne and reservoir-borne diseases for which a component of their life-cycle may be highly sensitive to any rainfall or temperature variation. For example, decreases in precipitation can create pools of stagnant water which are breeding grounds for vectors (Gagnon et al. 2001; Gagnon et al. 2002), flooding may cause contamination of surface water and wells through overflow of sewage systems and the failure of septic tanks (Lipp et al. 2002), or the loss of crops or water supplies may cause habitual changes or immunological deterioration in the population (Patz et al. 2005).

A key problem with time series analysis in long-term datasets is the separation of signals and stochastic noise. Noise can hide cohering patterns, whilst within a series there may be a number of competing signals of varying strength. For example rainfall measures over time have a number of seasonal changes, over a long period an ecological process such as disease outbreaks may only be linked to changes in one of these components. Here a number of statistical techniques were used to separate both the rainfall and BU cases time series data into different signals, including seasonal patterns, stochastic elements and long term trends. Further analysis was then undertaken to identify cohering patterns between rainfall events and BU cases.

4.3 Material and methods

4.3.1 Environmental and disease data

4.3.1.1 French Guiana

The only accurate long-term (decadal) dataset for cases of BU coupled with accurate weather data is from French Guiana in South America, with records going back to 1969. For details on French Guiana see chapter 1.4. The numbers of BU cases were obtained from Cayenne central hospital records dating from1969 until 2012, with identification based on a combination of histo-pathological, clinical and genetic analysis. This dataset is the most accurate long-term data on BU to our knowledge, which can be used for coherence with climatic factors. A potential issue with lesion causing diseases is the variation in time between appearance of symptoms and seeking of medical attention, reflected in lesion size. As French Guiana is part of the European Union and is a low population French territory, case reporting and assessment of lesions incurred a minimal delay; active surveillance of the disease was being undertaken with health care professionals who are trained to recognize BU being present in all towns and villages. Disease cases are distributed across the territory in line with the distribution of the population and are present almost ubiquitously where there are people. Rainfall data was obtained from Météo-France and recorded as the average rainfall in millimeters per month from 17 weather stations (Figure 4.1) across the populated coastal area of the territory from 1969-2012. Due to the restricted range of inhabited areas, a small human population and therefore a relatively low number of cases in each locality, an average rainfall reading from the stations along the coastal area was taken and compared to data on all BU cases across French Guiana. ENSO data for the period was taken from the American National Climatic Data Center and measured as the Sea Surface Temperature (SST) of the equatorial Pacific Ocean.

4.3.1.2 Akonolinga

Akonolinga is a town in the central province of the African country Cameroon. It lies along the river Nyong and has a population of approximately 21,300 people. The area is dominated by tropical rainforest and has a similar climate to that of French Guiana.

Data on BU cases in Cameroon were recorded from the BU management intervention at Akonolinga District Hospital; these cases were confirmed positive by the Centre Pasteur du Cameroun (CPC) (see Landier et al. 2014). Data on rainfall for Akonolinga was obtained from the NASA Tropical Rainfall Measuring Mission. Whilst this data is potentially less accurate than the data from French Guiana, particularly pertaining to the monitoring of cases and therefore the time between infection and reporting, if a similar response to rainfall is exhibited, this will help reinforce any BU-climate relationships.

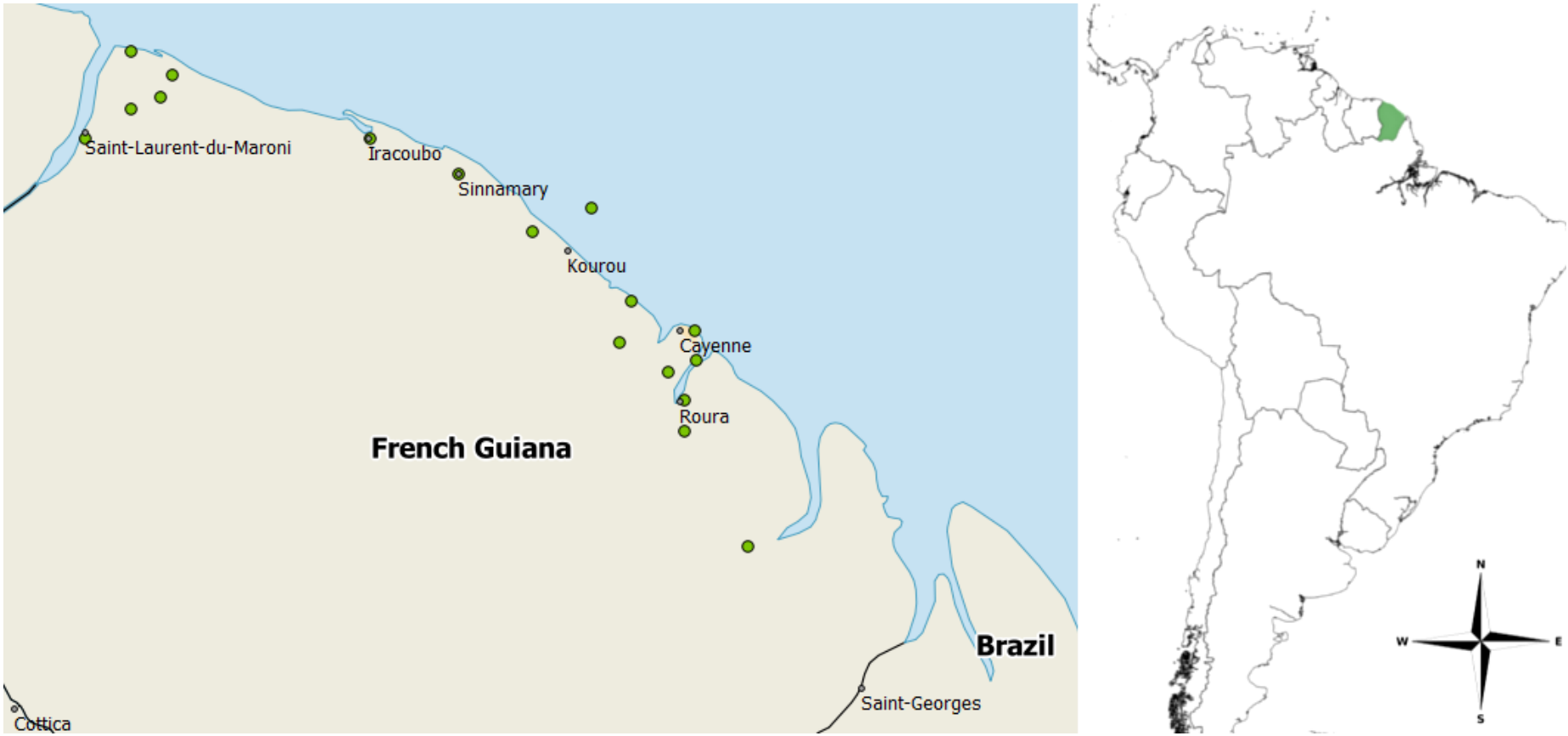


Figure 4.1 Map of French Guiana showing the location of 17 weather stations along the coast of French Guiana, including one station located on an off-shore island, and the position of French Guiana within South America.

4.3.2 Ethical provisions

Written consent for participation to the study was obtained from patients in all instances. BU cases in French Guiana received treatment appropriately according to the French laws in public health, which are also under application in this territory. BU cases in Akonolinga were treated in accordance with local laws. The study protocol was authorized by Cayenne General Hospital authorities according to French ethical rules. The database for French Guiana was declared to the Commission National Informatique et Libertés (CNIL number 3X#02254258) following French law requirements. The databases from both Akonolinga and French Guiana did not include names or any variable that could allow the identification of patients.

4.3.3 Statistical analysis

For both French Guiana and Akonolinga, the following analysis was carried out, however forecasting and coherence analysis between the long-term Akonolinga BU cases trend and rainfall trend was omitted due to the dataset being temporally too short:

4.3.3.1 Singular spectrum analysis decomposition and reconstruction

Since the introduction of SSA by Broomhead and King (Broomhead and King 1986a; Broomhead and King 1986b), it has been applied successfully to several economic, financial and industrial time series (Hassani 2007; Hassani et al. 2009; Hassani and Thomakos 2010) and has also been used previously in the analysis of coherence between disease and climate (Pascual et al. 2000; Rodó et al. 2002; Koelle et al. 2005; Chaves and Pascual 2006; Chaves et al. 2012). For example consider a real-valued nonzero time series $Y_T = (\gamma_1,...,\gamma_T)$ of sufficient length T.

The main purpose of SSA is to decompose the original series into a sum of series, so that each component in this sum can be identified as either a trend, periodic or quasi-periodic component (perhaps, amplitude-modulated), or noise. This is followed by a reconstruction of the original series. Each corresponding stage involves two primary steps, for decomposition; embedding and singular value decomposition (SVD) and for reconstruction; grouping and reconstruction. For a detailed description of each stage see Golyandina et al (2001). In short, decomposition breaks the time series down into its constituent components (in this instance repeating seasonal and long-term patterns in rainfall). Once isolated it is possible to identify stochastic noise within the leftover signal and remove it before reconstructing a new noise-free time series.

Here SSA is applied to the rainfall time series for both French Guiana and Akonolinga. Each seasonal component of both series' was first identified using periodograms, graphical representations of the distribution of power (or variance) among different frequencies. Independence of each seasonal component was also tested. The main concept in studying SSA component properties is 'separability', which characterizes how well different components can be separated from each other. SSA decomposition of the series Y_T can only be successful if the resulting additive components of the series are approximately separable. A natural measure of dependence between two time series $Y_T^{(1)}$and $Y_T^{(2)}$ is the weighted correlation or "ω-correlation"(Golyandina et al. 2001). To identify correlations between all the components within the time series a ω-correlation matrix was created. This shows the ω-correlation for the components in a 20-grade grey scale from white to black corresponding to the absolute values of correlations from 0 to 1 (with 0 being no correlation and 1 being absolute correlation).

For forecasting the French Guiana data the series was further analyzed using sequential SSA, which refers to an analysis of the residuals (Hassani 2007). As a result of sequential SSA, it is possible to identify signal components which were incorrectly classified as noise through the earlier decomposition and then combine them together in order to build a signal which is shown as total rainfall residual. It is then possible to include the total signal extracted from the residuals in to the earlier reconstruction and use the post sequential SSA reconstruction for forecasting new data points.

4.3.3.2 Empirical Mode Decomposition

To explore relationships between seasonal components of the rainfall time series and any corresponding seasonal changes in BU, repeating intra and inter annual patterns in BU incidence were extracted using Empirical Mode Decomposition (EMD). EMD is a technique developed specifically to decompose non-stationary and non-linear time series and has been successfully applied to a number of climatic and epidemiological datasets (Chaves et al. 2012; Hurtado et al. 2014). The method is an iterative process which builds a number of oscillatory signals called intrinsic mode functions (IMF), that are repeatedly subtracted from the time series, each iteration results in an IMF with a longer periodic component until just a trend signal is present. For a detailed description of EMD see Huang et al 1998. A periodogram was created for each IMF to assess the frequency of the repeating pattern through time.

4.3.3.3 Trend coherence analysis

To identify the relationship between the rainfall trend obtained via SSA, SST, and the number of human BU cases in French Guiana, continuous wavelet transforms were used. Wavelets have been utilised previously in various branches of

ecological theory (Torrence and Compo 1998; Klvana et al. 2004; Maraun and Kurths 2004; Maraun et al. 2007; Cazelles et al. 2008) and to identify relationships between disease and long-term climatic patterns (Chowell et al. 2009). They have the benefit of being unaffected by non-stationary time series, which are often found in ecology (Torrence and Compo 1998). Cross wavelet coherence analysis was performed using the biwavelet R package (Gouhier 2013) with the Morlet mother wavelet and 2,000 Monte Carlo randomisations. In order to further characterise the association between the time series', phase analysis was also undertaken to identify the phasing difference between the two, for example whether one precedes the other (Cazelles and Stone 2003), this is indicated by arrows on the wavelet plots. To further assess the statistical significance of the patterns exhibited by the wavelet approach, null models were tested. To create time series for the null models bootstrapping was used to construct from the observed time series, control datasets, which share properties of the original series under the following null hypothesis: the variability of the observed time-series or the association between two time-series is no different to that expected from outbreaks independent of the rainfall trend.

4.3.3.4 Seasonal coherence analysis

To identify correlations between seasonal components of the rainfall time series extracted via SSA, SST and seasonal changes in BU cases represented by the extracted IMF signals, cross correlation functions were used (CCF) (Shumway and Stoffer 2011). These measure the similarity between two oscillating time series as a function of a time lag applied to one of them and can be used with stationary time series (i.e. series which statistical properties including mean, variance and autocorrelation are consistent over time).

4.4 Results

4.4.1 French Guiana

4.4.1.1 SSA

Periodograms of the raw rainfall time series (Figure 4.2A) identified seasonal components which oscillated yearly for approximately 4, 6 and 12 month periods (repeating patterns occurring, tri-annually, bi-annually and once per year, Figure 4.2B) the signals of each repeating oscillation were isolated along with the long-term trend of rainfall (Figure 4.2C-F). Remaining data was considered noise and the components were reconstructed into a noise free signal (Figure 4.2G), a second periodogram of the new reconstructed signal shows the repeating patterns are intact, whilst stochastic noise has been removed (Figure 4.2H). Figure 4.2I shows the resulting output of forecasting from sequential SSA, with rainfall beginning to decline as French Guiana enters a trough of dryer years after several years of high rainfall.

The 4 month component corresponds to two rainy seasons, one long rainy season and one short, and also to the main dry season from August to December (Figure 4.3A). The 6 month bi-annual component corresponds to the two rainy seasons and is a reflection of the strength of these two seasons (Figure 4.3B), whilst the 12 month component shows the overall rainfall level for the year (Figure 4.3C). Separability of these components was confirmed with a ω-correlation matrix showing that these seasonal components did not show high levels of correlation with other components (Figure 4.4).

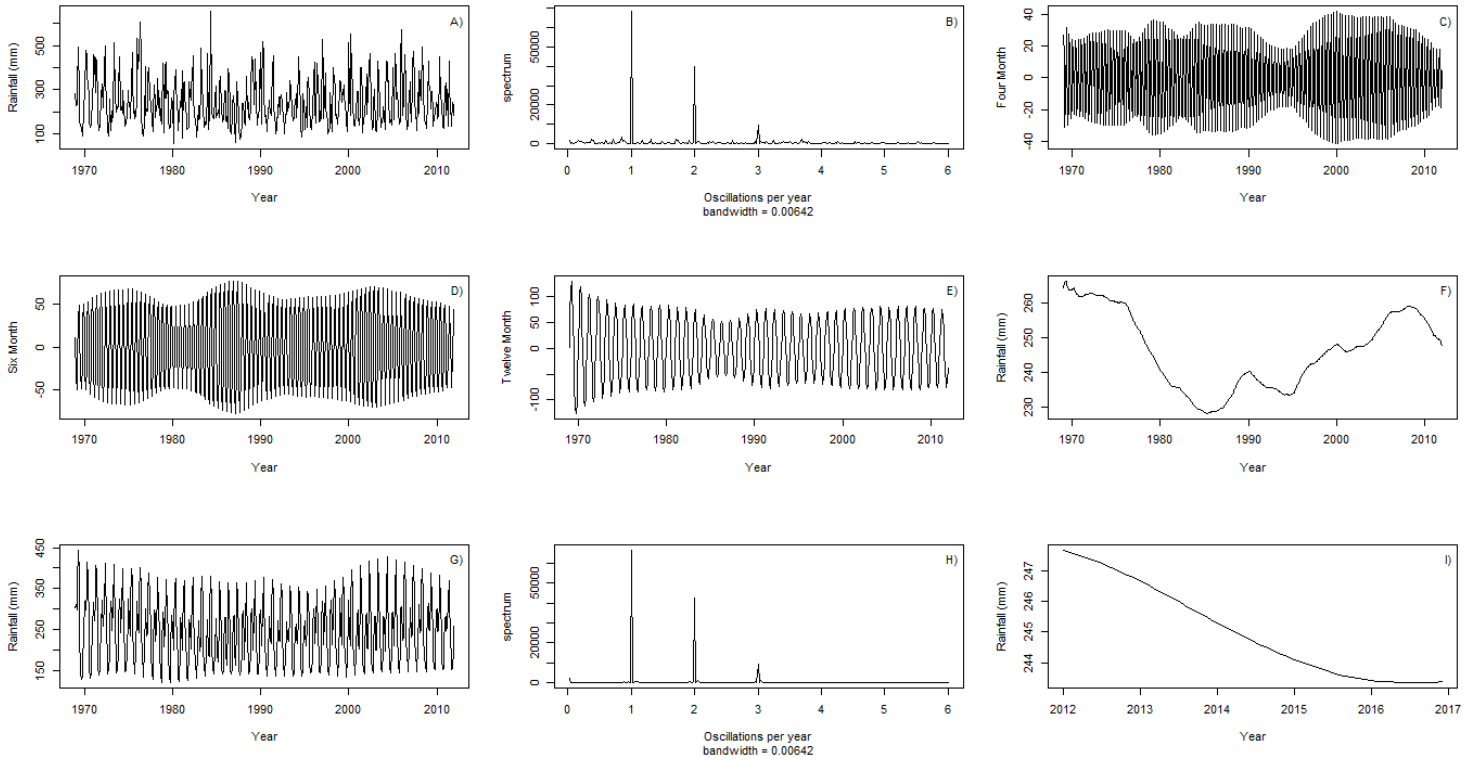


Figure 4.2 French Guiana, monthly time series data showing the decomposition, reconstruction and forecasting of datapoints using Singular Spectrum Analysis. A) Original rainfall time series average for 17 weather stations along the coast of French Guiana, from 1969 to 2012. B) Periodograms of the rainfall time series identifying significant repeating patterns once per year, twice per year and three times per year. SSA extracted component corresponding to periodogram spike of C) Three times per year, four month component, D) Twice per year, six month component, E) Once per year, twelve month component. F) The extracted rainfall trend. G) The reconstructed rainfall time series after the removal of stochastic noise. H) A second periodogram of the reconstructed rainfall series showing less stochastic noise around the 3 main repeating patterns. I) Forecasting of the rainfall trend to 2017 using sequential SSA.

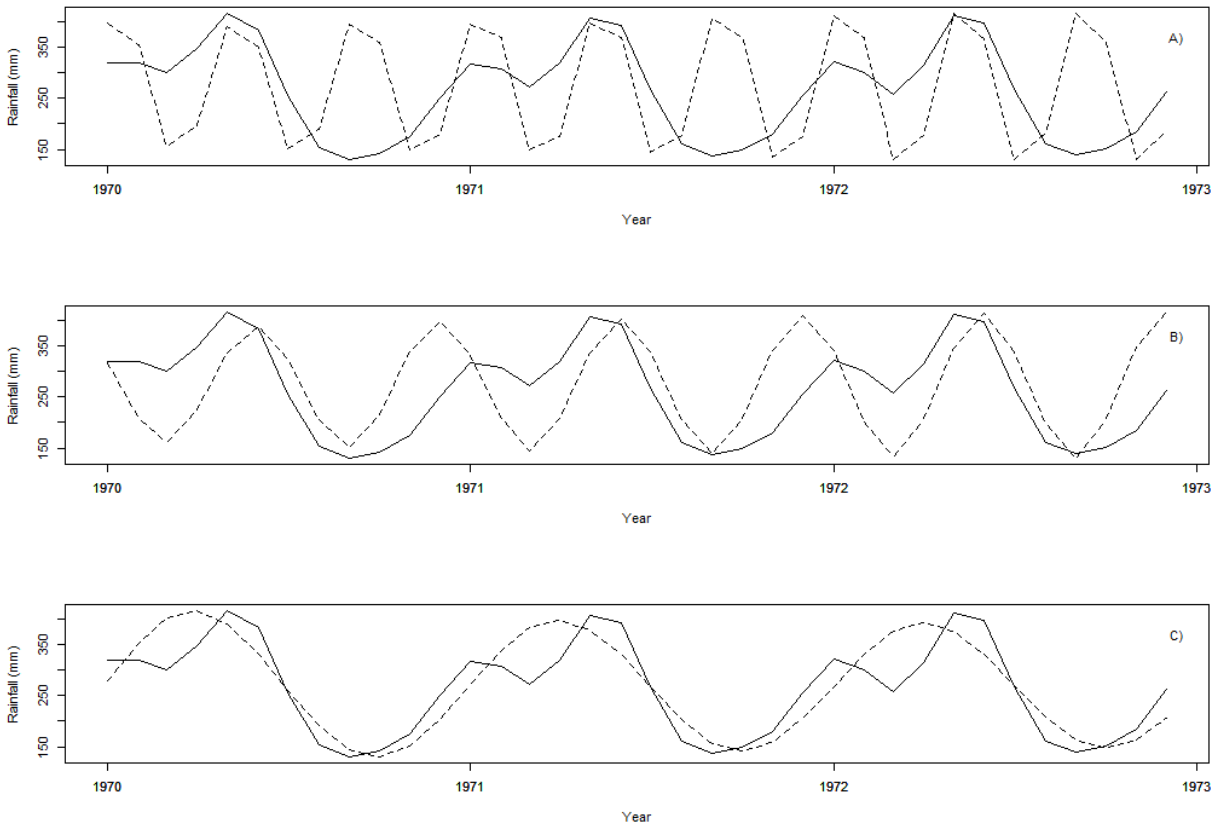


Figure 4.3 French Guiana, three year period of extracted components (dashed line) plotted against the same period of the reconstructed rainfall (solid line), showing which parts of the rainfall series the components represent. A) Four month component corresponds to both rainy seasons and the dry season. C) Six month component corresponds to the two rainy seasons. C) Twelve month component represents the rainfall for the whole year.

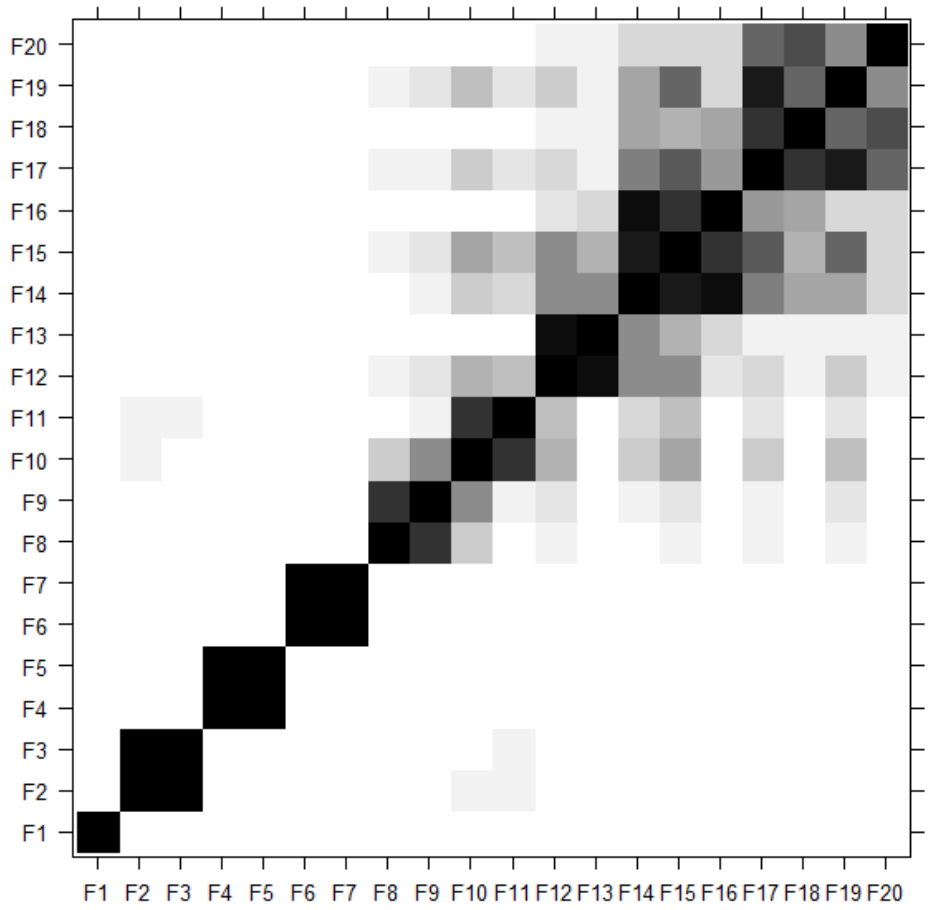


Figure 4.4 French Guiana, ω-correlation matrix, the F values represent oscillating components within a year (i.e. F6 is the 6 month bi-annual component). The level of correlation can be found by finding the component of interest along the x-axis and looking up the y-axis to see where it corresponds with other rainfall components. Large values of ω-correlation between reconstructed components indicate that they should possibly be gathered into one group and correspond to the same component in SSA decomposition. The matrix uses a 20-grade grey scale from white to black corresponding to the absolute values of correlations from 0 to 1 (with 0 being no correlation and 1 being absolute correlation).

4.4.1.2 Empirical mode decomposition (EMD)

Seven IMF series' were produced by EMD of repeating patterns from the BU case data (Figures 4.5A-H), periodograms of these show that the first IMF has high levels of variation across the spectrum and therefore is likely noise, the second is a bi-yearly repeating pattern and the third is a measure of cases per year, the fourth over two years and the fifth over 4 years. Subsequent IMF series' are long-term trends (Figures 4.5I-O).

4.4.1.3 Trend coherence analysis

During the period 1969-2012 the series was dominated by four inter-annual peaks in rainfall followed by three inter-annual periods of recessions, with 3 corresponding peaks and recessions of BU disease cases. The results of the wavelet coherence analyses showed a statistically significant correlation between the two time series for 1979-2000, with cohering peaks and troughs over periods of approximately 8 years (Figure 4.6A). Phase analysis, indicated by the arrows pointing downwards suggests a preceding relationship of rainfall change occurring before BU cases. In essence, after a peak in rainfall, during a dry period, the number of BU cases increases. Null models showed no significant relationships between rainfall and disease cases. The analysis between cases and SST showed less coherence; however there were two short periods during the mid 1970's and early 1990's which weakly corresponded (Figure 4.6B).

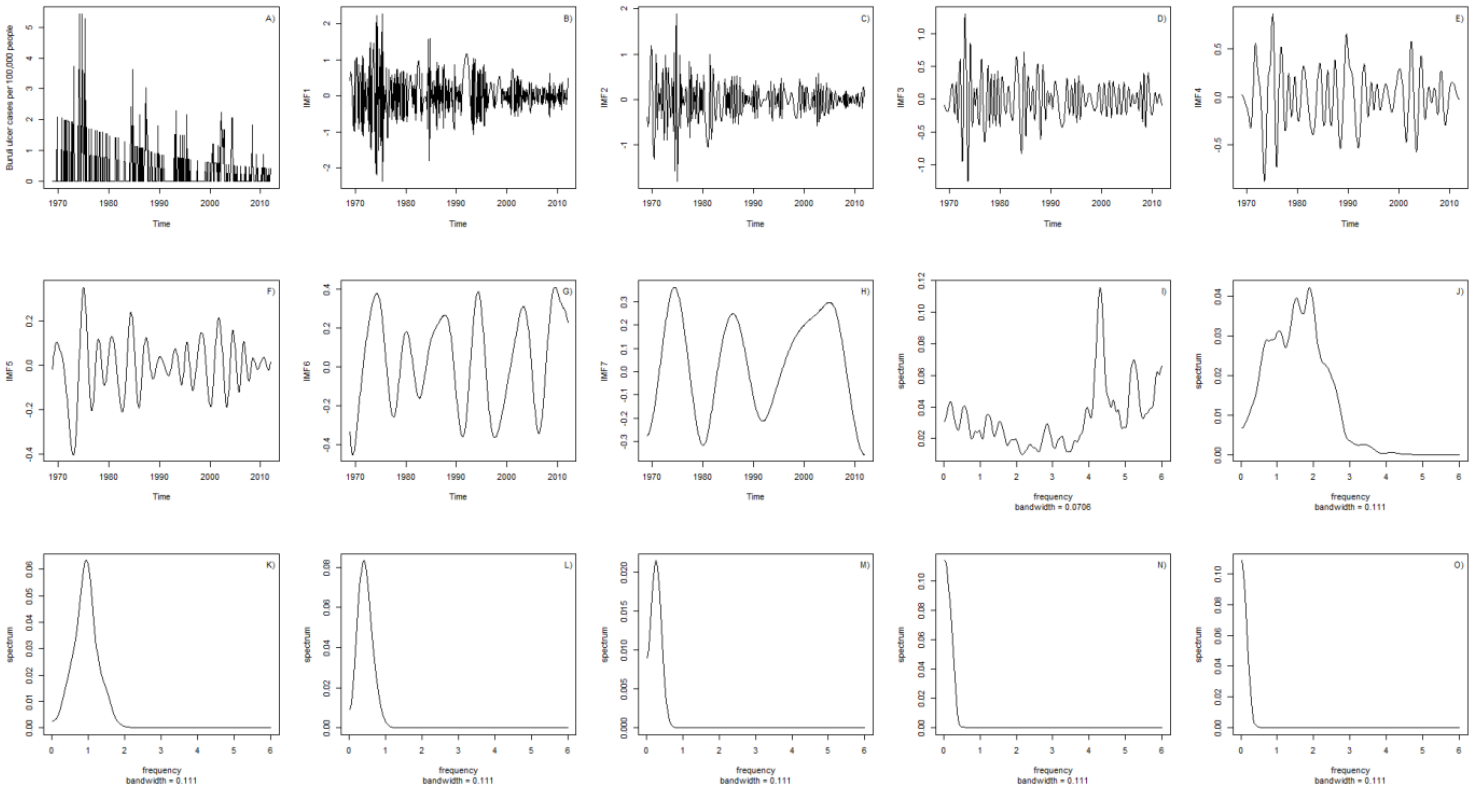


Figure 4.5 French Guiana, monthly time series and Empirical Mode Decomposition of BU cases per 100,000 people. A) Monthly time series of BU cases per 100,000 people from 1969 to 2012. B) First Intrinsic Mode Function (IMF). C) Second IMF. D) Third IMF. E) Fourth IMF. F) Fifth IMF. G) Sixth IMF. H) Seventh IMF. Periodograms for I) first IMF showing a high level of variation across the spectra and therefore should be considered white noise. J) Second IMF which has its highest power at two cycles per year. K) Third IMF with its highest power at one cycle per year. L) Fourth IMF with highest power approximately at one cycle every two years. M) Fifth IMF with highest power approximately at one cycle every 4 years. N) Sixth IMF with a low level of cycles representing very long-term trends. O) Seventh IMF also representing long-term trends.

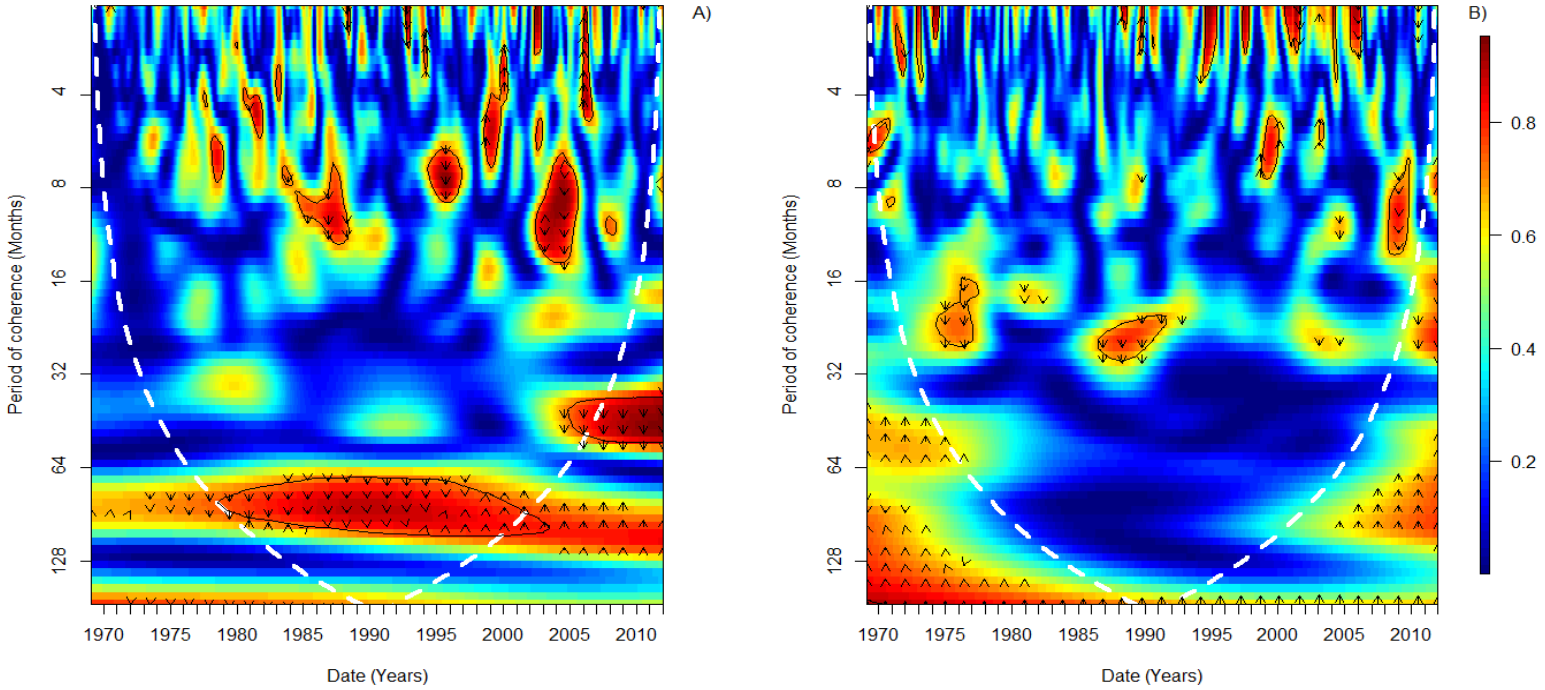


Figure 4.6 French Guiana, wavelet coherence between A) BUincidences per 100,000 people and the rainfall trend obtained from Singular Spectrum Analysis (SSA). B) BU incidences per 100,000 people and El Niño-Southern Oscillation (ENSO). The colours are coded from dark blue to dark red with dark blue representing low coherence through to high coherence with dark red. The solid black lines around areas of red show the α=5% significance levels computed based on 2000 Monte Carlo randomisations. The dotted white lines represent the cone of influence; outside this area coherence is not considered as it may be influenced by edge effects. The black arrows represent the phase analysis and adhere to the following pattern: Arrows pointing to the right mean that rainfall and cases are in phase, arrows pointing to the left mean that they are in anti-phase., arrows pointing up mean that cases lead rainfall and arrows pointing down mean that rainfall leads cases.

4.4.1.4 Seasonal coherence analysis

Cross correlation functions between the IMF signals representing intra and inter-annual patterns in BU cases (first, second, third, fourth and fifth IMFs) and both SST and the reconstructed rainfall pattern from SSA analysis showed a number of corresponding signatures.

SST did not correlate with the SSA derived rainfall series (Figure 4.7A) but did corresponded with rainfall before SSA was applied (Figure 4.7B) suggesting that SST spikes cause higher levels of unpredictable rainfall anomalies, which during SSA analysis were classified as "noise" or stochastic events. SST also corresponded with the first IMF of BU cases (Figure 4.7C), which was similarly classified as noise. This may mean that SST creates rainfall anomalies which cause high levels of BU but do not follow any set seasonal or long-term patterns (i.e. random one off events). SST fluctuations further matched with inter-annual variation of BU cases with long lag periods suggesting the total number of cases over these periods is increased by the influence of SST driven anomalies (Figure 4.7D-G). In particular the fourth IMF, where a below average SST (La Niña) value produces a higher than average level of BU cases over 2 years after a lag time of approximately 18 months (Figure 4.7F). Conversely a peak in SST (El Niño) creates a decline in the 4 year oscillation of BU cases (Figure 4.7G).

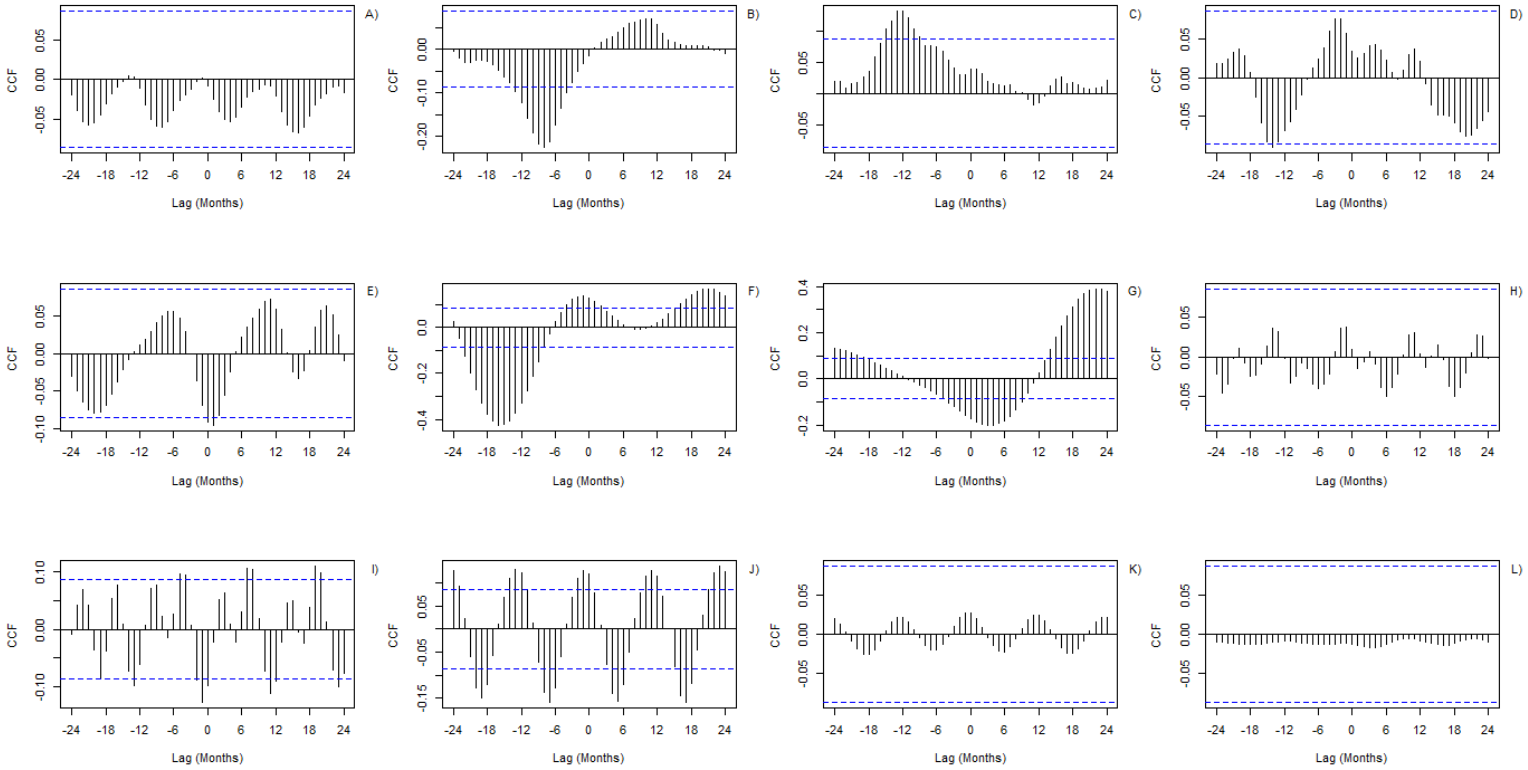


Figure 4.7 French Guiana, cross correlation analysis between A) El Niño-Southern Oscillation (ENSO) and reconstructed rainfall time series from Singular Spectrum Anlaysis. B) ENSO and original rainfall time series. ENSO and C) First Intrinsic Mode Function (IMF) from BU cases. D) Second IMF. E) Third IMF. F) Fourth IMF. G) Fifth IMF. Reconstructed rainfall series and H) First IMF. I) Second IMF. J) Third IMF. K) Fourth IMF. L) Fifth IMF. Dashed horizontal blue lines in all panels represent the 95% confidence limit; black vertical lines which go beyond the dashed line can be considered non-random cohering oscillations between the two time series being assessed, with the lag period between an above average oscillation in the first time series and a subsequent above average oscillation in the second shown on the x-axis.

The reconstructed rainfall series has a corresponding relationship with both the second and third IMF (Figures 4.7I and J) suggesting that an above average level of precipitation contributes to an above average spike in bi-annual BU cases, and also the overall yearly number of cases. By looking at the SSA derived components of rainfall against the IMF signatures of BU it is possible to identify the exact rainfall components which are driving these spikes. The SSA derived four month component correlated with no IMF signatures (Figures 4.8A-C). Figures 4.8D-F show the correlations of the first, second and third IMF signatures with the SSA derived 6 month component (i.e. the strength of both the two wet seasons), a significant correlation is identified with the bi-annual BU cases (Second IMF) with an approximate five month lag, therefore because of the average reported incubation periods of 3-5 months (Trubiano et al. 2013) and the lag time of several weeks before diagnosis, the two spikes in cases per year are most likely to be driven by the strength of the two spikes in rainfall per year rather than the two dry seasons. The 12 month component which is a measure of the total rainfall per year did not correlate with any seasonal IMF signatures (8G and 8H) but as expected correlated with the overall level of BU per year (Figure 4.8I).

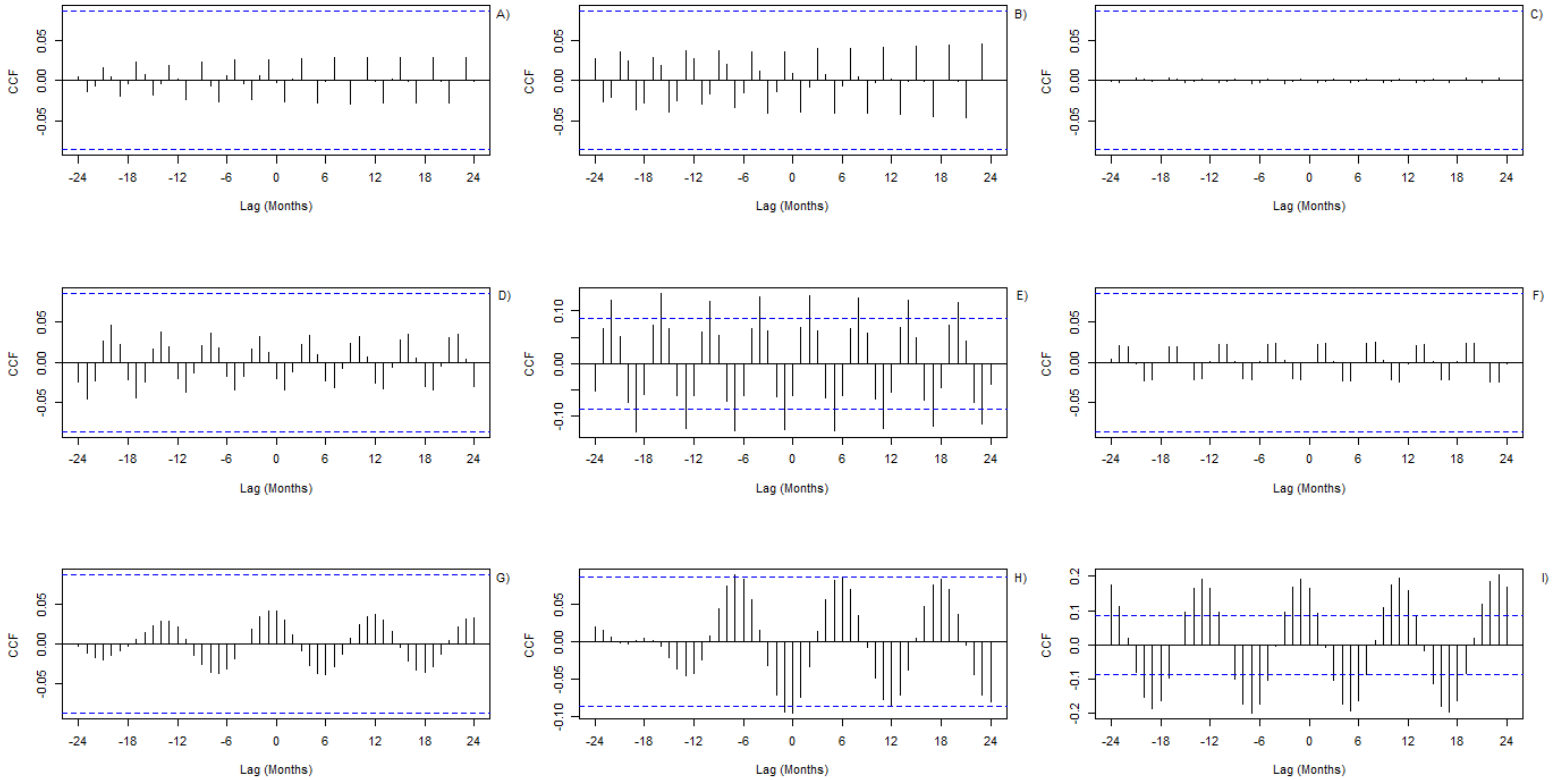


Figure 4.8 French Guiana, cross correlation analysis between rainfall seasonal components extracted by Singular Spectrum Analysis and seasonal Intrinsic Mode Function (IMF) series of BU. Four month component and A) First IMF. B) Second IMF. C) Third IMF. Six month component and D) First IMF. E) Second IMF. F) Third IMF. Twelve month component and G) First IMF. H) Second IMF.I) Third IMF. Dashed horizontal blue lines in all panels represent the 95% confidence limit; black vertical lines which go beyond the dashed line can be considered non-random cohering oscillations between the two time series being assessed, with the lag period between an above average oscillation in the first time series and a subsequent above average oscillation in the second shown on the x-axis.

4.4.2 Akonolinga

4.4.2.1 SSA

Periodograms of the raw rainfall time series identified a highly similar pattern to French Guiana (Figure 4.9A) with seasonal components which oscillated yearly for approximately 4, 6 and 12 month periods (repeating patterns occurring, tri-annually, bi-annually and once per year, Figure 4.9B) the signals of each repeating oscillation were isolated along with the long-term trend of rainfall (Figure 4.9C-F). Remaining data was considered noise and the components were reconstructed into a noise free signal (Figure 4.9G), a second periodogram of the new reconstructed signal shows the repeating patterns are intact, whilst stochastic noise has been removed (Figure 4.9H). Forecasting was not performed for the Akonolinga data as the dataset was too short for accurate long-term predictions. Separability of these components was confirmed with a ω-correlation matrix showing that these seasonal components did not show high levels of correlation with other components (Figure 4.10).

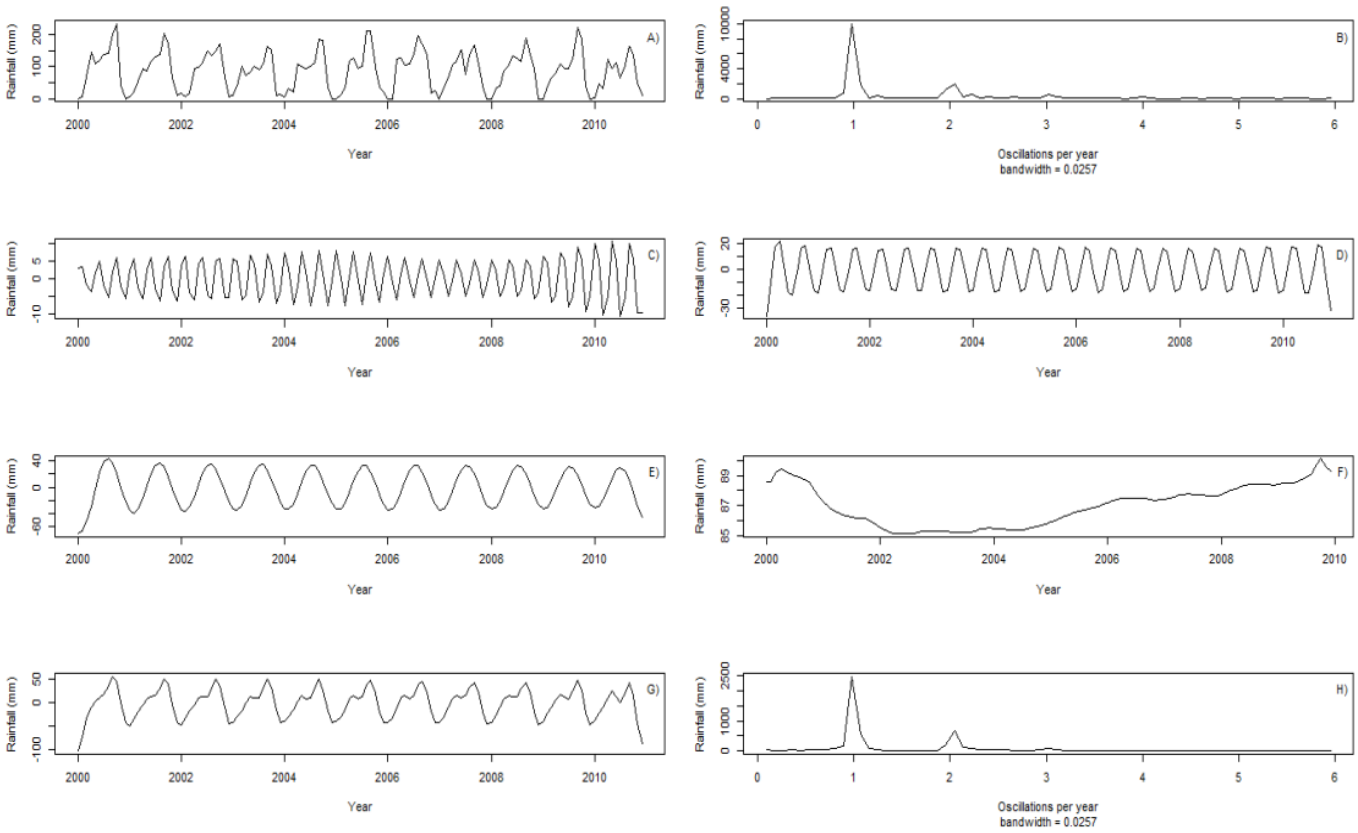


Figure 4.9 Akonolinga, monthly time series data showing the decomposition, reconstruction and forecasting of datapoints using Singular Spectrum Analysis. A) Original rainfall time series average for 17 weather stations along the coast of French Guiana, from 1969 to 2012. B) Periodograms of the rainfall time series identifying significant repeating patterns once per year, twice per year and three times per year. SSA extracted component corresponding to periodogram spike of C) Three times per year, four month component, D) Twice per year, six month component, E) Once per year, twelve month component. F) The extracted rainfall trend. G) The reconstructed rainfall time series after the removal of stochastic noise. H) A second periodogram of the reconstructed rainfall series showing less stochastic noise around the 3 main repeating patterns.

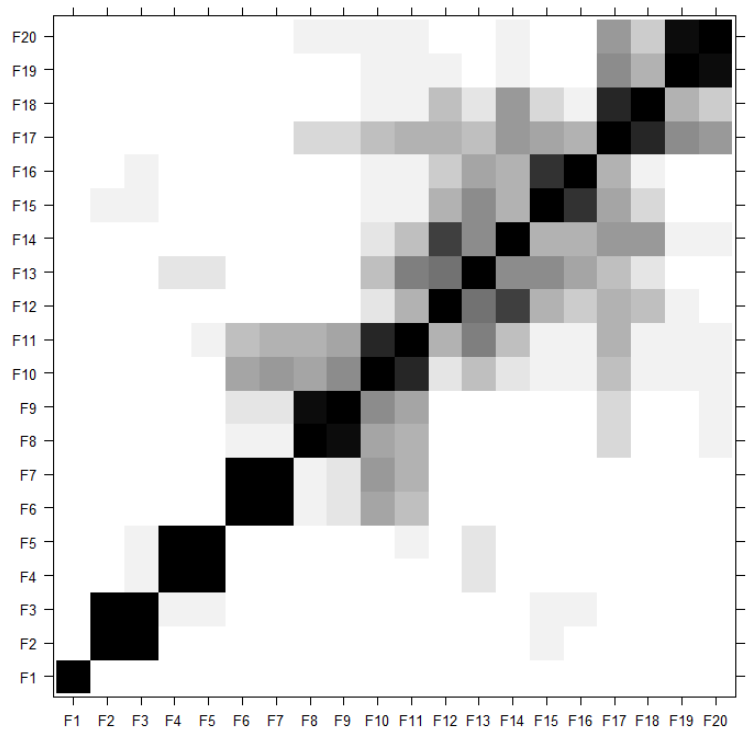


Figure 4.10 Akonolinga, ω-correlation matrix, the F values represent oscillating components within a year (i.e. F6 is the 6 month bi-annual component). The level of correlation can be found by finding the component of interest along the x-axis and looking up the y-axis to see where it corresponds with other rainfall components. Large values of ω-correlation between reconstructed components indicate that they should possibly be gathered into one group and correspond to the same component in SSA decomposition. The matrix uses a 20-grade grey scale from white to black corresponding to the absolute values of correlations from 0 to 1 (with 0 being no correlation and 1 being absolute correlation).

4.4.2.2 Empirical mode decomposition (EMD)

As with French Guiana Seven IMF series' were produced by EMD of repeating patterns in the BU case data (Figure 4.11A-H), these followed an almost identical model, with two main repeating patterns being identified; yearly and bi-yearly. The yearly pattern is to be expected in both data sets and shows the repeating yearly cycle of cases, the bi-yearly demonstrates that both data sets, French Guiana and Akonlinga have two seasonal peaks in BU.

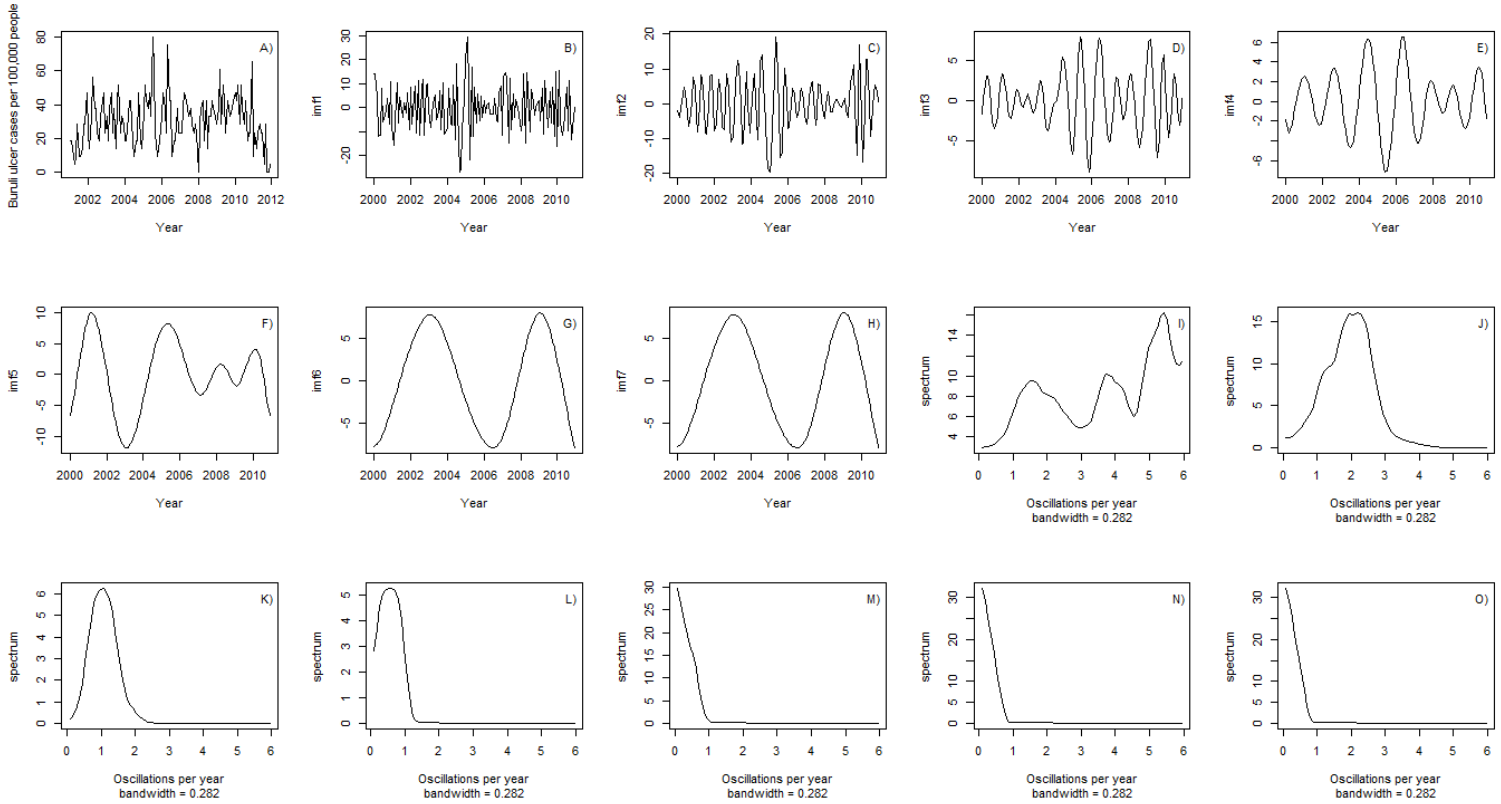


Figure 4.11 Akonolinga, monthly time series and Empirical Mode Decomposition of BU cases per 100,000 people. A) Monthly time series of BU cases per 100,000 people from 1969 to 2012. B) First Intrinsic Mode Function (IMF). C) Second IMF. D) Third IMF. E) Fourth IMF. F) Fifth IMF. G) Sixth IMF. H) Seventh IMF. Periodograms for I) first IMF showing a high level of variation across the spectra and therefore should be considered white noise. J) Second IMF which has its highest power at two cycles per year. K) Third IMF with its highest power at one cycle per year. L) Fourth IMF with highest power approximately at one cycle every two years. M) Fifth IMF with highest power approximately at one cycle every 4 years. N) Sixth IMF with a low level of cycles representing very long-term trends. O) Seventh IMF also representing long-term trends.

4.4.2.3 Seasonal coherence analysis

Cross correlation functions between the IMF signals representing intra and inter-annual patterns in BU cases (first, second, third, fourth and fifth IMFs) and both SST and the reconstructed rainfall pattern from SSA analysis showed a number of corresponding signatures.

SST did not correlate with the SSA derived rainfall series (Figure 4.12A) nor did it with rainfall before SSA was applied (Figure 4.12B) this suggests that unlike with French Guiana there is no significant influence of SST on the rainfall in Akonlinga. SST did however correlate with the third IMF signal of BU cases (Figure 4.12F), representing the bi-yearly seasonal peaks. Both the SSA reconstructed rainfall series and original rainfall series correlated with the second and third IMF (Figure 4.12I-J), representing seasonal BU cases and the overall yearly cases respectively.

The SSA derived four month component correlated with no IMF signatures (Figures 4.13A-C). Figures 4.13D-F show the correlations of the first, second and third IMF signatures with the SSA derived 6 month component (i.e. the strength of both the two wet seasons), again as in the French Guiana data a significant correlation is identified with the bi-annual BU cases (Second IMF) with an approximate five month lag. The 12 month component which is a measure of the total rainfall per year did not correlate with any seasonal IMF signatures (4.12G and 4.12H) but again as expected correlated with the overall level of BU per year (Figure 4.12I).

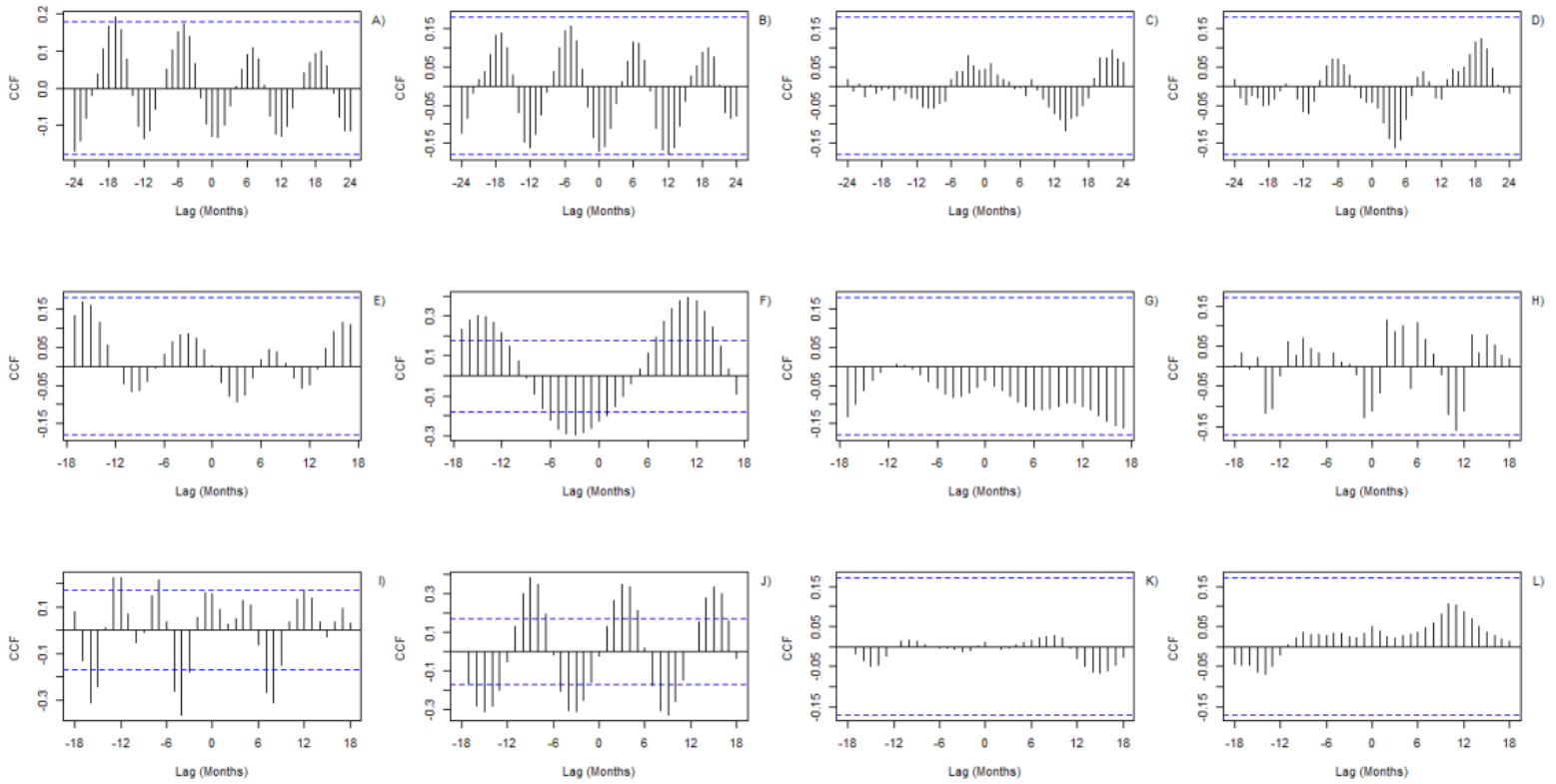


Figure 4.12 Akonlinga, cross correlation analysis between A) El Niño-Southern Oscillation (ENSO) and reconstructed rainfall time series from Singular Spectrum Anlaysis. B) ENSO and original rainfall time series. ENSO and C) First Intrinsic Mode Function (IMF) from BU cases. D) Second IMF. E) Third IMF. F) Fourth IMF. G) Fifth IMF. Reconstructed rainfall series and H) First IMF. I) Second IMF. J) Third IMF. K) Fourth IMF. L) Fifth IMF. Dashed horizontal blue lines in all panels represent the 95% confidence limit; black vertical lines which go beyond the dashed line can be considered non-random cohering oscillations between the two time series being assessed, with the lag period between an above average oscillation in the first time series and a subsequent above average oscillation in the second shown on the x-axis.

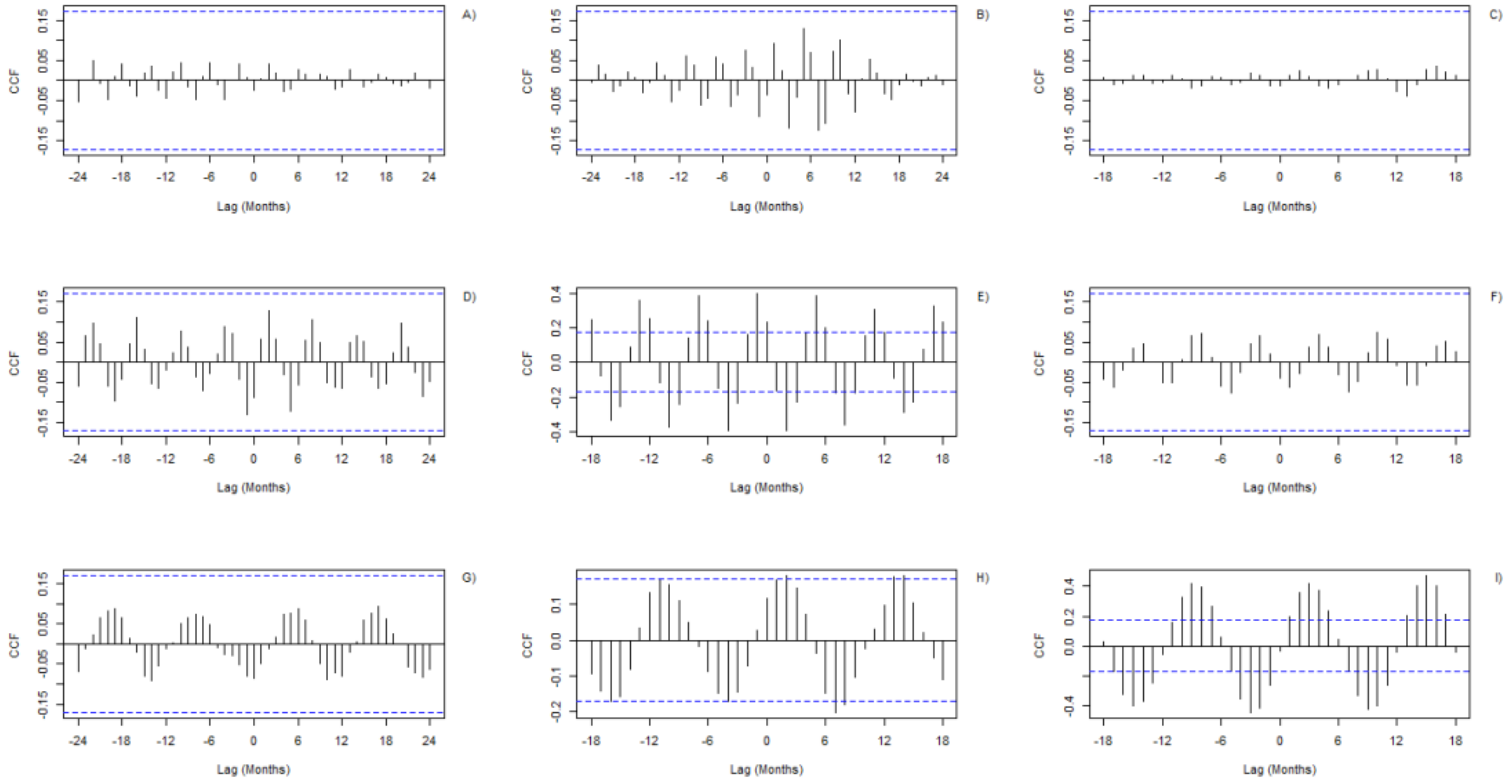


Figure 4.13 Akonolinga**,** Cross correlation analysis between rainfall seasonal components extracted by Singular Spectrum Analysis and seasonal Intrinsic Mode Function (IMF) series of BU. Four month component and A) First IMF. B) Second IMF. C) Third IMF. Six month component and D) First IMF. E) Second IMF. F) Third IMF. Twelve month component and G) First IMF. H) Second IMF.I) Third IMF. Dashed horizontal blue lines in all panels represent the 95% confidence limit; black vertical lines which go beyond the dashed line can be considered non-random cohering oscillations between the two time series being assessed, with the lag period between an above average oscillation in the first time series and a subsequent above average oscillation in the second shown on the x-axis.

4.5 Discussion

In French Guiana the identification of more than one temporal correlation between rainfall and BU disease, in addition to wet weather anomalies outside of usual seasonal peaks in rainfall being driven by SST, highlights the complexity of using environmental changes in predicting disease outbreaks. The stochastic SST driven incidences and the influence of long-term rainfall in addition to seasonal drivers is a first for BU in the world. Whilst a similar long and short-term climatic pattern has already been linked to cholera outbreaks (Pascual et al. 2002; Koelle et al. 2005), the results of this paper show that such patterns are likely to be more spread among aquatic infectious pathogens such as with MU and that SST relationships' may be driving stochastic cases. The same seasonal and yearly patterns being repeated in Akonolinga goes further to cement the relationship BU has with climatic fluctuations, although in this series there is no evidence of a correlation between stochastic rainfall events and BU cases outside of the main seasonal cycles. There is however a correlation between the two seasonal cycles and ENSO, and the correlation between the reconstructed rainfall series and ENSO is on the boundaries of significant with an 18 month lag period. This suggests that ENSO may be having an effect on the seasonal BU cases, potentially increasing their intensity, but it is not driving stochastic events. This could be expected as the influence of ENSO is limited in this area of Africa with the main changes in rainfall happening further north. There is also a much shorter data set and so only a very few peaks of ENSO are recorded over this period.

Whilst further study will be required to fully understand what niche MU occupies within the environment, these results show the identification of a relationship with rainfall and provide important testable insights into its disease ecology. It also exemplifies the importance of using long-term datasets when trying to establish

relationships between the environment and infectious disease, and the use of techniques such as wavelets, SSA and EMD to look deeper into time series.

By removing noise and decomposing the time series using SSA it was possible to look at the influence of each individual seasonal pattern on BU and to identify a cause of important rainfall anomalies which occur outside the regular seasonal patterns. By further extracting a long-term trend and relating this to cases, the effects of several components of rainfall become apparent, something which would perhaps be lost without decomposition.

During the analysis EMD had some advantages and disadvantages over SSA which makes it suitable for differing time series, dependent on the application. In this instance the ability to successfully identify noise and a seasonal component, in addition to a hierarchy of increasing long-term periodic components in disease cases was beneficial, particularly when linking disease patterns to SST. SSA however provided a more accurate separation of seasonal components within the rainfall data.

The increasing use of wavelets (Torrence and Compo 1998; Maraun and Kurths 2004; Maraun et al. 2007; Cazelles et al. 2008; Chowell et al. 2009) is an important development for time series analysis in epidemiology, previously in this field, non-stationarity presented a serious problem for relating ecological, epidemiological and climatic datasets (Hastings 2001; Benton et al. 2006; Cazelles and Hales 2006). Wavelets provide the advantage of being localized in both time and frequency whereas the standard Fourier transform, traditionally used in time series analysis is only localized in frequency (Torrence and Compo 1998) ,which although useful for identifying constant periodic components, is not able to account for changes in frequency over time.

The results of the relationship between SST and BU cases in French Guiana broadly agrees with similar observations in Australia where it was found that periods of wet weather approximately 16-17 months prior to an outbreak, followed by a period of dry weather for 5 months to be the most suitable for BU emergence (van Ravensway et al. 2012). The results presented here suggest that outbreaks of BU over a long period (at least 2 years, as represented by the fourth IMF), correspond with a decline in SST (i.e. a La Niña event) 17 months prior, whilst the opposite is true for El Niño events, with the effect causing a below average decrease in BU cases over the preceding 4 years. A La Niña event in the north of South America corresponds to a marked increase in wet weather anomalies, (as corroborated by the significant CCF between the pre-SSA rainfall and SST), whilst El-Niño signals an increase in dryer weather. As SST also correlates with high rainfall events and BU cases classified as stochastic, it is possible that SST driven rainfall anomalies create random outbreaks in BU, independent of the usual seasonal cycles over a two year period.

The long-term and seasonal relationships between rainfall and cases, i.e. the long-term peaks in BU driven by a recession in rainfall over several years, and the bi-annual peaks in cases driven by spikes in the two rainy seasons, could be explained in several ways. Whilst the limited knowledge of BU disease transmission makes it only possible to speculate, the analysis does present an opportunity to shed further light on the ecological niche occupied by MU and its environmental heterogeneity in space and time. Long periods of wet weather followed by a decrease in rainfall may increase the number of stagnant water bodies and swampland, flowing rivers may recede into a series of isolated pools (Cazelles et al. 2005), whilst newly formed channels and wetlands will be cut off. This could spark an outbreak of disease vectors, host carriers and other aquatic species that have been identified to have a relationship with BU and which thrive in these conditions (Portaels et al.

1999; Marsollier et al. 2002; Mosi et al. 2008; Williamson et al. 2008; Fyfe et al. 2010), or beneficially change the aquatic community structure. A second hypothesis, which assumes the disease agent does not necessarily require a true symbiotic relationship, would be that the initial rainfall washes the agent into new territory, and once established the dry weather will again increase the number of preferential water-bodies thereby increasing the number of cases. Both events have the potential to co-occur over a long and short period of time and could potentially be working at differing spatial scales, driven by either long or short-term rainfall patterns. Previous studies have shown how community assemblages and system dynamics can be significantly altered over varying periods of time at sites with differing size, hydrology and landscape parameters (Harris 1980; Kratz et al. 1987; Maio and Corkum 1995; Arthington et al. 2005; Ruetz et al. 2005; Garchitorena et al. 2014). The majority of the BU cases occur when both the long and short-term temporal changes coincide, suggesting weather conducive to providing an increase in infectious habitats or human contact with these habitats during such a period. Using the argument that the disease agent is most abundant in stagnant water, it is possible to hypothesise that large bodies of water created after several years of high rainfall may recede slowly, stagnate and undergo significant changes in biotic community over several dry years. Nested within this period high peaks in rainy seasons as suggested by an increase in the biannual component, or SST driven rainfall anomalies, will cause these stagnating areas to swell, potentially flooding or seeping onto pathways where human contact is increased. In addition periods of high rainfall during periods of long-term dryer climates could also be conducive to the occurrence of potential invertebrate vectors. Previous work on mosquito populations for example, which have been linked to BU (Fyfe et al. 2010), show that they are at their highest during periods of high short-term climatic variability, cited as the amount of short-term fluctuations around a mean climate state on a fine time scale (Loevinsohn 1994; Githeko and Ndegwa 2001; Zhou et al. 2004;

Paaijmans et al. 2010). This may add weight to the idea of mosquito-based vectors, although it is possible that several less studied aquatic species exhibit similar responses. It must also be considered that the relationship could be influenced more by human behaviour, for example dryer years will often induce an increase in recreational use of local water bodies, notably for fishing and hunting.

The results of the forecasting are of potential importance in predicting future disease emergence. The most recent highest rainfall in French Guiana has occurred in 2009 and it is being followed by a 5-year period of dry weather (Figure 4.2I). Based on the analysis and the identification of a relationship with rainfall it is possible to predict that this should also be followed by an increase in BU cases in the region.

Caution must be expressed when using time series data, particularly over long periods, as changes in diagnostic capabilities (for example the introduction of PCR, knowledge of the disease and improvements in equipment and recording) and the accuracy in weather data collection will cause significant temporal discrepancies. These factors are difficult to address and are inherent to all such studies. It is also important to note that unpredictable external factors such as socio-economical changes can also be having an unknown influence. The use of long-term weather data also presents the problem of how to include landscape parameters, as these also have an influence on a disease but are often not available or poorly recorded early on in time. In this instance however it seems unlikely that a cyclical pattern is related to a steady change in population and landscape. The robust analysis shows that French Guiana time series for BU cases reveals interesting non-random patterns which are vital for understanding the ecological niche of this aquatic microbial agent.

4.6 References

Arthington, A. H., Balcombe, S. R., Wilson, G. A., Thoms, M. C. and Marshall, J., 2005. Spatial and temporal variation in fish-assemblage structure in isolated waterholes during the 2001 dry season of an arid-zone floodplain river, Cooper Creek, Australia. *Marine and Freshwater Research*, 56 (1), 25-35.

Benton, T. G., Plaistow, S. J. and Coulson, T. N., 2006. Complex population dynamics and complex causation: devils, details and demography. *Proceedings of the Royal Society B: Biological Sciences*, 273 (1591), 1173-1181.

Broomhead, D. S. and King, G. P., 1986a. Extracting qualitative dynamics from experimental data. *Physica D: Nonlinear Phenomena*, 20 (2–3), 217-236.

Broomhead, D. S. and King, G. P., 1986b. On the qualitative analysis of experimental dynamical systems. *In:* Sarkar, S., ed. *Nonlinear Phenomena and Chaos.* Bristol: Adam Hilger, 113-144.

Cazelles, B., Chavez, M., Berteaux, D., Ménard, F., Vik, J., Jenouvrier, S. and Stenseth, N., 2008. Wavelet analysis of ecological time series. *Oecologia*, 156 (2), 287-304.

Cazelles, B., Chavez, M., McMichael, A. J. and Hales, S., 2005. Nonstationary Influence of El Niño on the Synchronous Dengue Epidemics in Thailand. *PLoS Medicine*, 2 (4), e106.

Cazelles, B. and Hales, S., 2006. Infectious diseases, climate influences, and nonstationarity. *PLoS Medicine*, 3 (8).

Cazelles, B. and Stone, L., 2003. Detection of imperfect population synchrony in an uncertain world. *Journal of Animal Ecology*, 72 (6), 953-968.

Chaves, L. F. and Pascual, M., 2006. Climate Cycles and Forecasts of Cutaneous Leishmaniasis, a Nonstationary Vector-Borne Disease. *PLoS Medicine*, 3 (8), e295.

Chaves, L. F., Satake, A., Hashizume, M. and Minakawa, N., 2012. Indian Ocean Dipole and Rainfall Drive a Moran Effect in East Africa Malaria Transmission. *Journal of Infectious Diseases*, 205 (12), 1885-1891.

Chowell, G., Munayco, C., Escalante, A. and McKenzie, F. E., 2009. The spatial and temporal patterns of falciparum and vivax malaria in Peru: 1994-2006. *Malaria Journal*, 8 (1), 142.

Fyfe, J. A. M., Lavender, C. J., Handasyde, K. A., Legione, A. R., O'Brien, C. R., Stinear, T. P., Pidot, S. J., Seemann, T., Benbow, M. E., Wallace, J. R., McCowan, C. and Johnson, P. D. R., 2010. A Major Role for Mammals in the Ecology of *Mycobacterium ulcerans*. *PLoS Neglected Tropical Diseases*, 4 (8), e791.

Gagnon, A., Smoyer-Tomic, K. and Bush, A., 2002. The El Niño Southern Oscillation and malaria epidemics in South America. *International Journal of Biometeorology*, 46 (2), 81-89.

Gagnon, A. S., Bush, A. B. G. and Smoyer-Tomic, K. E., 2001. Dengue epidemics and the El Niño Southern Oscillation. *Climate Research*, 19 (1), 35-43.

Garchitorena, A., Roche, B., Kamgang, R., Ossomba, J., Babonneau, J., Landier, J., Fontanet, A., Flahault, A., Eyangoh, S. and Guégan, J.-F., 2014. *Mycobacterium ulcerans* Ecological Dynamics and Its Association with Freshwater Ecosystems and Aquatic Communities: Results from a 12-Month Environmental Survey in Cameroon. *PLoS Neglected Tropical Diseases*, 8 (5), e2879.

Githeko, A. and Ndegwa, W., 2001. Predicting Malaria Epidemics in the Kenyan Highlands Using Climate Data: A Tool for Decision Makers. *Global Change and Human Health*, 2 (1), 54-63.

Golyandina, N., Nekrutkin, V. and Zhigljavsky, A., 2001. *Analysis of Time Series Structure: SSA and related techniques*. Chapman and Hall/CRC.

Gouhier, T. C., 2013. *biwavelet: Conduct univariate and bivariate wavelet analyses*. 0.14. http://biwavelet.r-forge.r-project.org. [Accessed 01/06/2013].

Hanf, M., Adenis, A., Nacher, M. and Carme, B., 2011. The role of El Nino southern oscillation (ENSO) on variations of monthly Plasmodium falciparum malaria cases at the cayenne general hospital, 1996-2009, French Guiana. *Malaria Journal*, 10 (1), 100.

Harris, G. P., 1980. Temporal and Spatial Scales in Phytoplankton Ecology. Mechanisms, Methods, Models, and Management. *Canadian Journal of Fisheries and Aquatic Sciences*, 37 (5), 877-900.

Hassani, H., 2007. *Singular Spectrum Analysis: Methodology and Comparison*. University Library of Munich, Germany.

Hassani, H., Heravi, S. and Zhigljavsky, A., 2009. Forecasting European industrial production with singular spectrum analysis. *International Journal of Forecasting*, 25 (1), 103-118.

Hassani, H. and Thomakos, D., 2010. A Review on Singular Spectrum Analysis for Economic and Financial Time Series. *Statistics and its Interface*, 3, 377-397.

Hastings, A., 2001. Transient dynamics and persistence of ecological systems. *Ecology Letters*, 4 (3), 215-220.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N.-C., Tung, C. C. and Liu, H. H., 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454 (1971), 903-995.

Hurtado, L. A., Cáceres, L., Chaves, L. F. and Calzada, J. E., 2014. When climate change couples social neglect: malaria dynamics in Panamá. *Emerging Microbes and Infections*, 3 (4), e27.

Klvana, I., Berteaux, D. and Cazelles, B., 2004. Porcupine Feeding Scars and Climatic Data Show Ecosystem Effects of the Solar Cycle. *The American Naturalist*, 164 (3), 283-297.

Koelle, K., Rodo, X., Pascual, M., Yunus, M. and Mostafa, G., 2005. Refractory periods and climate forcing in cholera dynamics. *Nature*, 436 (7051), 696-700.

Kratz, T. K., Frost, T. M. and Magnuson, J. J., 1987. Inferences from Spatial and Temporal Variability in Ecosystems: Long-Term Zooplankton Data from Lakes. *The American Naturalist*, 129 (6), 830-846.

Landier, J., Gaudart, J., Carolan, K., Seen, D. L., Guégan, J.-F., Eyangoh, S., Fontanet, A. and Texier, G., 2014. Spatio-temporal Patterns and Landscape-Associated Risk of Buruli Ulcer in Akonolinga, Cameroon. *PLoS Neglected Tropical Diseases*, 8 (9), e3123.

Lipp, E. K., Huq, A. and Colwell, R. R., 2002. Effects of Global Climate on Infectious Disease: the Cholera Model. *Clinical Microbiology Reviews*, 15 (4), 757-770.

Loevinsohn, M. E., 1994. Climatic warming and increased malaria incidence in Rwanda. *The Lancet*, 343 (8899), 714-718.

Mahamat, A., Dussart, P., Bouix, A., Carvalho, L., Eltges, F., Matheus, S., Miller, M. A., Quenel, P. and Viboud, C., 2013. Climatic drivers of seasonal influenza epidemics in French Guiana, 2006–2010. *Journal of Infection*, 67 (2), 141-147.

Maio, J. D. and Corkum, L. D., 1995. Relationship between the spatial distribution of freshwater mussels (Bivalvia: Unionidae) and the hydrological variability of rivers. *Canadian Journal of Zoology*, 73 (4), 663-671.

Maraun, D. and Kurths, J., 2004. Cross wavelet analysis: significance testing and pitfalls. *Nonlinear Processes in Geophysics*, 11 (4), 505-514.

Maraun, D., Kurths, J. and Holschneider, M., 2007. Nonstationary Gaussian processes in wavelet domain: Synthesis, estimation, and significance testing. *Physical Review E*, 75 (1), 016707.

Marsollier, L., Robert, R., Aubry, J., Saint Andre, J.-P., Kouakou, H., Legras, P., Manceau, A.-L., Mahaza, C. and Carbonnelle, B., 2002. Aquatic Insects as a Vector for *Mycobacterium ulcerans*. *Applied and Environmental. Microbiology.*, 68 (9), 4623-4628.

Marsollier, L., Stinear, T., Aubry, J., Saint André, J. P., Robert, R., Legras, P., Manceau, A.-L., Audrain, C., Bourdon, S., Kouakou, H. and Carbonnelle, B., 2004a. Aquatic Plants Stimulate the Growth of and Biofilm Formation by *Mycobacterium ulcerans* in Axenic Culture and Harbor These Bacteria in the Environment. *Applied and Environmental Microbiology*, 70 (2), 1097-1103.

Marsollier, L., Sévérin, T., Aubry, J., Merritt, R. W., Saint André, J.-P., Legras, P., Manceau, A.-L., Chauty, A., Carbonnelle, B. and Cole, S. T., 2004b. Aquatic Snails, Passive Hosts of *Mycobacterium ulcerans*. *Applied and Environmental Microbiology*, 70 (10), 6296-6298.

Merritt, R. W., Benbow, M. E. and Small, P. L. C., 2005. Unraveling an Emerging Disease Associated with Disturbed Aquatic Environments: The Case of Buruli Ulcer. *Frontiers in Ecology and the Environment*, 3 (6), 323-331.

Merritt, R. W., Walker, E. D., Small, P. L. C., Wallace, J. R., Johnson, P. D. R., Benbow, M. E. and Boakye, D. A., 2010. Ecology and Transmission of Buruli Ulcer Disease: A Systematic Review. *PLoS Neglected Tropical Diseases*, 4 (12), e911.

Morris, A., Gozlan, R., Marion, E., Marsollier, L., Andreou, D., Sanhueza, D., Ruffine, R., Couppié, P. and Guégan, J.-F., 2014. First Detection of *Mycobacterium ulcerans* DNA in Environmental Samples from South America. *PLOS Neglected Tropical Diseases*, 8 (1), e2660.

Mosi, L., Williamson, H., Wallace, J. R., Merritt, R. W. and Small, P. L. C., 2008. Persistent Association of *Mycobacterium ulcerans* with West African Predaceous Insects of the Family Belostomatidae. *Applied and Environmental Microbiology*, 7 4 (22), 7036-7042.

Paaijmans, K. P., Blanford, S., Bell, A. S., Blanford, J. I., Read, A. F. and Thomas, M. B., 2010. Influence of climate on malaria transmission depends on daily temperature variation. *Proceedings of the National Academy of Sciences*, 107 (34), 15135-15139.

Pascual, M., Bouma, M. J. and Dobson, A. P., 2002. Cholera and climate: revisiting the quantitative evidence. *Microbes and Infection*, 4 (2), 237-245.

Pascual, M., Cazelles, B., Bouma, M. J., Chaves, L. F. and Koelle, K., 2008. Shifting patterns: malaria dynamics and rainfall variability in an African highland. *Proceedings of the Royal Society B: Biological Sciences*, 275 (1631), 123-132.

Pascual, M., Rodó, X., Ellner, S. P., Colwell, R. and Bouma, M. J., 2000. Cholera Dynamics and El Niño-Southern Oscillation. *Science*, 289 (5485), 1766-1769.

Patz, J. A., Campbell-Lendrum, D., Holloway, T. and Foley, J. A., 2005. Impact of regional climate change on human health. *Nature*, 438 (7066), 310-317.

Portaels, F., Elsen, P., Guimaraes-Peres, A., Fonteyne, P.-A. and Meyers, W. M., 1999. Insects in the transmission of *Mycobacterium ulcerans* infection. *The Lancet*, 353 (9157), 986.

Rodó, X., Pascual, M., Fuchs, G. and Faruque, A. S. G., 2002. ENSO and cholera: A nonstationary link related to climate change? *Proceedings of the National Academy of Sciences*, 99 (20), 12901-12906.

Ruetz, C. R., Trexler, J. C., Jordan, F., Loftus, W. F. and Perry, S. A., 2005. Population dynamics of wetland fishes: spatio-temporal patterns

synchronized by hydrological disturbance? *Journal of Animal Ecology*, 74 (2), 322-332.

Shumway, R. and Stoffer, D., 2011. Time Series Regression and Exploratory Data Analysis. *In: Time Series Analysis and Its Applications.* Springer New York, 47-82.

Torrence, C. and Compo, G. P., 1998. A Practical Guide to Wavelet Analysis. *Bulletin of the American Meteorological Society*, 79 (1), 61-78.

Trubiano, J. A., Lavender, C. J., Fyfe, J. A. M., Bittmann, S. and Johnson, P. D. R., 2013. The incubation period of Buruli ulcer (*Mycobacterium ulcerans* infection). *PLoS Neglected Tropical Diseases*, 7 (10), e2463.

van Ravensway, J., Benbow, M. E., Tsonis, A. A., Pierce, S. J., Campbell, L. P., Fyfe, J. A. M., Hayman, J. A., Johnson, P. D. R., Wallace, J. R. and Qi, J., 2012. Climate and Landscape Factors Associated with Buruli Ulcer Incidence in Victoria, Australia. *PLoS ONE*, 7 (12), e51074.

Williamson, H. R., Benbow, M. E., Nguyen, K. D., Beachboard, D. C., Kimbirauskas, R. K., McIntosh, M. D., Quaye, C., Ampadu, E. O., Boakye, D., Merritt, R. W. and Small, P. L. C., 2008. Distribution of *Mycobacterium ulcerans* in Buruli Ulcer Endemic and Non-Endemic Aquatic Sites in Ghana. *PLoS Neglected Tropical Diseases*, 2 (3), e205.

World Health Organisation, 2001a. *Buruli Ulcer. Diagnosis of Mycobacterium ulcerans disease. A manual for health care providers* [online]. Geneva: Available from: http://www.who.int/iris/handle/10665/67000#sthash.cioj71jV.dpuf.

World Health Organisation, 2001b. *Buruli Ulcer Management of Mycobacterium ulcerans disease* [online]. Available from: http://whqlibdoc.who.int/hq/2001/WHO_CDS_CPE_GBUI_2001.3.pdf.

Zhou, G., Minakawa, N., Githeko, A. K. and Yan, G., 2004. Association between climate variability and malaria epidemics in the East African highlands.

Proceedings of the National Academy of Sciences of the United States of America, 101 (8), 2375-2380.

Chapter 5 Functional diversity provides a new perspective for understanding the ecology of emerging infectious diseases.

5.1 Summary

EID outbreaks are increasingly suspected to be a consequence of human pressures exerted on natural ecosystems and their biodiversity, and can be described as biological tags strongly tied with ecosystem and habitat perturbations (see chapter 1.1.2). Previously, taxonomic communities have been used as indicators of disease presence, and the loss of their diversity has been implicated as a driver of disease prevalence. The mechanistic details in how such pathogen-host systems function however cannot always be explained by phylogenetic variation or loss. For example, a change in taxonomic communities is unlikely to correlate with an emerging generalist pathogen that can associate itself with numerous functionally similar, but phylogenetically distinct species. Here functional diversity indices are utilised in addition to methods based on Gower's dissimilarity to quantify invertebrate functional groups and their diversity at sites endemic for MU. Changes in these metrics allowed the rapid categorization of the ecological niche of the bacteria and the ability to relate specific traits to its survival. This technique enables for the first time the testing of the dilution theory on a generalist pathogen and the assessment of whether diversity of species with traits conducive to the bacteria's survival, or functional diversity and evenness of occupied trait space across a whole site acts as a barrier for the disease. It was found that functional diversity loss and evenness had no bearing on the bacteria's presence, and whilst there was a strong correlation between certain functional traits, diversity of host carrier species with these traits is unlikely to be a compounding factor to its survival. This suggests the dilution effect of host diversity is perhaps less applicable to some emerging diseases and postulate that high diversity at certain

trophic levels, caused by mid-level disturbance, may drive the evolution of pathogenic traits or allow such diseases to proliferate.

5.2 Introduction

One of the consequences of a rapid increase in global human population is a similarly rapid decline in biological diversity both in terrestrial and aquatic ecosystems (Rahel 2002; McKinney 2006, 2008). Current estimates suggest that from about 10,000 to 20,000 freshwater species have either become extinct or are under severe threat (Harvell et al. 2002; Gleick 2003; Meybeck 2003; Vorosmarty et al. 2010), with direct consequences for freshwater ecosystem functioning and therefore the emergence of associated infectious diseases. In effect, based on the complex relationships between hosts and pathogens, it is expected that along with the current changes of freshwater community structures, rapid changes will also arise in the known epidemiological pattern of water-borne infectious diseases and the potential emergence of new pathogens (Chapin et al. 2000; Ostfeld and Keesing 2000a, 2000b; Schmidt and Ostfeld 2001; LoGiudice et al. 2003; Ezenwa et al. 2006; Swaddle and Calos 2008; Keesing et al. 2009; Ogden and Tsao 2009; Ostfeld 2009; Johnson and Thieltges 2010; Keesing et al. 2010; Rigaud et al. 2010; Sehgal 2010; Searle et al. 2011), (see chapter 1.1). Until now however, understanding of the consequences of these dramatic changes of host communities on the emergence of infectious pathogens has remained limited (Jones et al. 2008; Keesing et al. 2010; Roche et al. 2013; Roche and Guégan 2011;Randolph and Dobson 2012). This is an important component in comprehending how human influence and pressures exerted by population expansion can lead to the emergence of disease and can also be seen as a proxy for environmental surveillance of habitat degradation.

A central theme to the Eco-Health paradigm is determining the role of biodiversity as a source of infectious pathogens, e.g. an amplifier of their prevalence or a limitation to their transmission due to dilution and buffer effects or similar (Ostfeld and Keesing 2000b; Roche and Guégan 2011; Roche et al. 2012 ; Randolph and Dobson 2012). The loss of some host species (i.e. loss of biodiversity) however may be difficult to relate to generalist multi-host pathogen transmission, as these organisms are likely to inhabit niche space provided or facilitated by species that share specific traits, rather than genetic relatedness. For example, pathogens that are not dependent on vector hosts often have ecological niches that include associations with abiotic and biotic substrates favourable to their growth (Carolan et al. 2014), these can be found on or within many organisms. Such multi-host pathogens equate to a large proportion of infectious diseases, with the majority of emerging and re-emerging human pathogens (circum 50 to 60% having more than 3 different host species) often having the broadest range of hosts (Woolhouse and Gowtage-Sequeria 2005). Many bacteria that have a soil or water-borne origin typically fall into this category of pathogens that travel within the environment without a clear route of propagation (Sachs et al. 2011). These pathogens are also more likely to be impacted by changes in groups of organisms that inhabit certain functional categories, which in turn are more directly linked with changes in ecosystem function; e.g by soil erosion, habitat destruction and land-use.

The need for greater mechanistic details in how such pathogen-host systems function cannot therefore be provided by taxonomic variation in biotic communities, particularly with many phylogenetically close species (Petchey and Gaston 2002, 2007, 2009). It is thus important to characterise the specific role of functional groups as potential drivers of pathogen emergence and determine macro-ecological and evolutionary patterns responsible for infectious diseases (Dizney

and Ruedas 2009; Elwell et al. 2009; Garrett et al. 2009; Rigaud et al. 2010; Roche and Guégan 2011; Randolph and Dobson 2012; Roche et al. 2013).

Despite MU being strongly associated with aquatic environments and associated with numerous potential aquatic invertebrate reservoir species, (see chapter 1.3.4). Previous attempts to link the bacteria to traditional taxonomic indicators have had varying success and the relevance of each taxon in the lifecycle of the bacteria remains disputed (Marsollier et al. 2002; Marsollier et al. 2004a; Marsollier et al. 2004b; Benbow et al. 2008; Fyfe et al. 2010; Benbow et al. 2013). Relationships between disease cases and climate suggest outbreaks are most common during long-term dryer periods over several years (van Ravensway et al. 2012) but are triggered by higher than average levels of seasonal rainfall, or by extreme stochastic rainfall events as detailed in chapter 4 (Morris et al. 2014). Similar associations are seen in other bacteria, notably the spore-forming *Bacillus anthracis*. In a comparable manner MU may enter a dormant phase during dry periods in the soil (Lamont et al. 2012), the negative charge of the outer lipid layer (Deshayes et al. 2013) would cause it to be repelled from soil particles and therefore susceptible to being washed into water-bodies during heavy rainfall. After time with the right conditions, these water-bodies can become hotspots for the bacteria, waiting for heavy rainfall again to increase human contact. For example, during dry periods, isolated water-bodies engage in a process of acidification through stagnation, thus increasing the number of positively charged soil particles with which the bacillus's negatively charged outer lipid layer can readily bind to and concentrate. Unlike *B. anthracis* the negative charge on MU is not lost at a lower *pH* (Deshayes et al. 2013) and so the bacteria is less likely to perish in these acidic environments. Organisms in the soil and sediment are potentially key to the movement of the bacteria into suitable hosts, initially by causing changes in the ionic structure of particles, breaking down organic matter

and releasing the bacteria from their bonds. Furthermore, inferences from the MU genome indicate the necessity for a relationship with other organisms when active, with reductive evolution suggesting a requirement for a very stable microhabitat and the provision of complex sugars provided by a host (Stinear et al. 2008) (see also chapter 1.3.3). The physical movement of bacteria from below the sediment surface to the water by epibenthic burrowers, in a similar mechanism to *B. anthracis* with earthworms, as well as their predation by macro-invertebrates could also be important in moving the disease up into the water column and through the food chain (Willson et al. 2013). This may increase the chances of other disease reservoirs becoming infected or if necessary for transmission to humans, infection of vectors capable of inoculating a human host. The possibility of deceased animals contaminated with the bacteria being a component in the life-cycle (Fyfe et al. 2010), or of the bacteria being introduced as detritus from plants, aquatic or riparian which it may have a symbiotic relationship with should also be considered.

Here for the first time functional diversity and changes in functional groups, rather than taxonomic changes will be related to the presence of MU in the environment, thereby helping categorise its ecological niche within the environment. This will also identify whether groups such as soil dwelling taxa are important in the bacteria's life cycle. The results will then be used to enable the testing of the dilution theory using related functional groups of organism, i.e if the bacteria is dependent on a specific functional group, does taxonomic diversity within this group affect its survival. This type of approach to disease ecology can be applied to a wide range of environment-borne infectious diseases such as cholera, anthrax, tuberculosis, legionellosis, histoplasmosis or acinetobacter-disease. The aim is for a better characterisation of the ecological niche of non-vectored pathogens and to identify the effects of changes in functional communities, enabling the prediction of current and future hot spots of emergence.

5.3 Methods

5.3.1 Data collection

This chapter uses data from a previous study conducted in Ghana, Africa. Invertebrates, a *priori* considered to be non-host or host-carriers for MU, were collected from 94 sites in areas endemic and non-endemic for BU during the months of July, August and September over a three-year period (2005-2007) using standardized sampling methods published in detail previously (Benbow et al. 2008; Williamson et al. 2008). The sites were also classified into lotic (flowing waters, often very slow) and lentic (non-flowing) habitats for determining habitat-specific relationships of MU and invertebrate functional traits and groups. Habitat descriptions can be found in (Benbow et al. 2013).

5.3.2 PCR analysis

Environmental samples were comprised of water filtrate and dominant aquatic plant biofilms collected during the same time period as previously described (Benbow et al. 2013). These samples were evaluated for the presence of MU DNA using PCR targeting the insertion sequence IS*2404* and the the enoyl reductase (ER) domain from the polyketide synthase (PKSs) genes coding for mycolactone synthesis (Williamson et al. 2008). Only samples positive for both primer pairs were used in the present study.

5.3.3 Functional group identification

Aquatic taxa were in most cases determined at the family level. A functional trait matrix was built based on current literature for each family of invertebrates, including traits considered to have an impact on how the bacteria may persist in the environment and based on the average trait type for species within a family in a specific habitat (Cummins et al. 2005; Petchey and Gaston 2006) (see Apendix II

Table S5.1 for a list of traits used and their definitions, and all families with their associated traits). In addition, traits were weighted by importance in the following order: Feeding group and food type, body type and mobility, respiration type and main body constituents (see Appendix II Table S5.1). Using Gower dissimilarity (Gower and Legendre 1986) a dendrogram was constructed from the functional traits matrix (Figure 1), K-Means partitioning using Simple Structure Index (SSI) (Dolnicar et al. 1999) was applied to characterise the effective number of functional groups based on the dendrogram (Figure SM1). Principal coordinates analysis (PCoA) was then used to compute PCoA axis as functional traits and calculate a number of functional diversity indices (Laliberté and Shipley 2011): Functional Dispersion (Laliberté and Legendre 2010), Functional Evenness, (Villéger et al. 2008) and Rao's Quadratic Entropy (Botta-Dukát 2005). It was not possible to calculate Functional Divergence because all data were categorical. Taxa with less than a total of 10 organisms at all sites were excluded.

5.3.4 Analysis

Linear mixed models (LMM) were performed using the lme4 R package (Bates et al. 2012). Firstly, functional indices (functional dispersion, functional evenness and Rao's quadratic entropy) were compared with the proportion of MU positive environmental samples at a site (considered a proxy for the level of MU present). Secondly, proportions of each functional group were compared with the proportion of MU positive environmental samples. Thirdly, individual taxa were tested against the proportion of MU positive environmental samples. Spatial autocorrelation was tested by evaluating residuals of the models using both Moran's *I* test (95% confidence interval) and correlogs smoothed with a spline function (95% pointwise bootstrap confidence intervals) (Bjørnstad and Falck 2001). Spatial autocorrelation was accounted for in the LMM by factoring geographical location into the model as a random effect, the previous tests were then repeated to confirm the removal of

this correlation. P-values were estimated from the model using the pval.fnc function of the language R package with 10,000 Markov chain Monte-Carlo samples (Baayen 2011). The sites were then divided into lentic, i.e. stagnant waters such as swamps and ponds, and lotic, i.e. flowing waters such as rivers and creeks, ecosystem categories and the same analysis was performed again. All models were validated by checking residuals for normality and homogeneity, and where needed data was appropriately transformed. Temporal independence was also evaluated using a regression of residuals against time.

AIC values of functional groups against MU environmental levels were compared to individual models of each taxon within a group against MU levels, in addition to analysis of variance between these models. This was to identify whether the functional group as a whole was a better fit model for predicting levels of MU within and across aquatic localities than individual species within that group. Individual species within a group being a better fit would work against the necessity for grouping by function.

To test the dilution effect of an increase in diversity of host carriers having a detrimental effect on pathogen presence, groups found to have a positive relationship with the level of MU DNA were further tested for a cohering relationship between taxonomic diversity and MU DNA. This change was measured by calculating two diversity indices, Simpson and Shannon for the group, a LMM accounting for spatial autocorrelation was then used identify whether a relationship between within-group diversity and bacterial DNA exists.

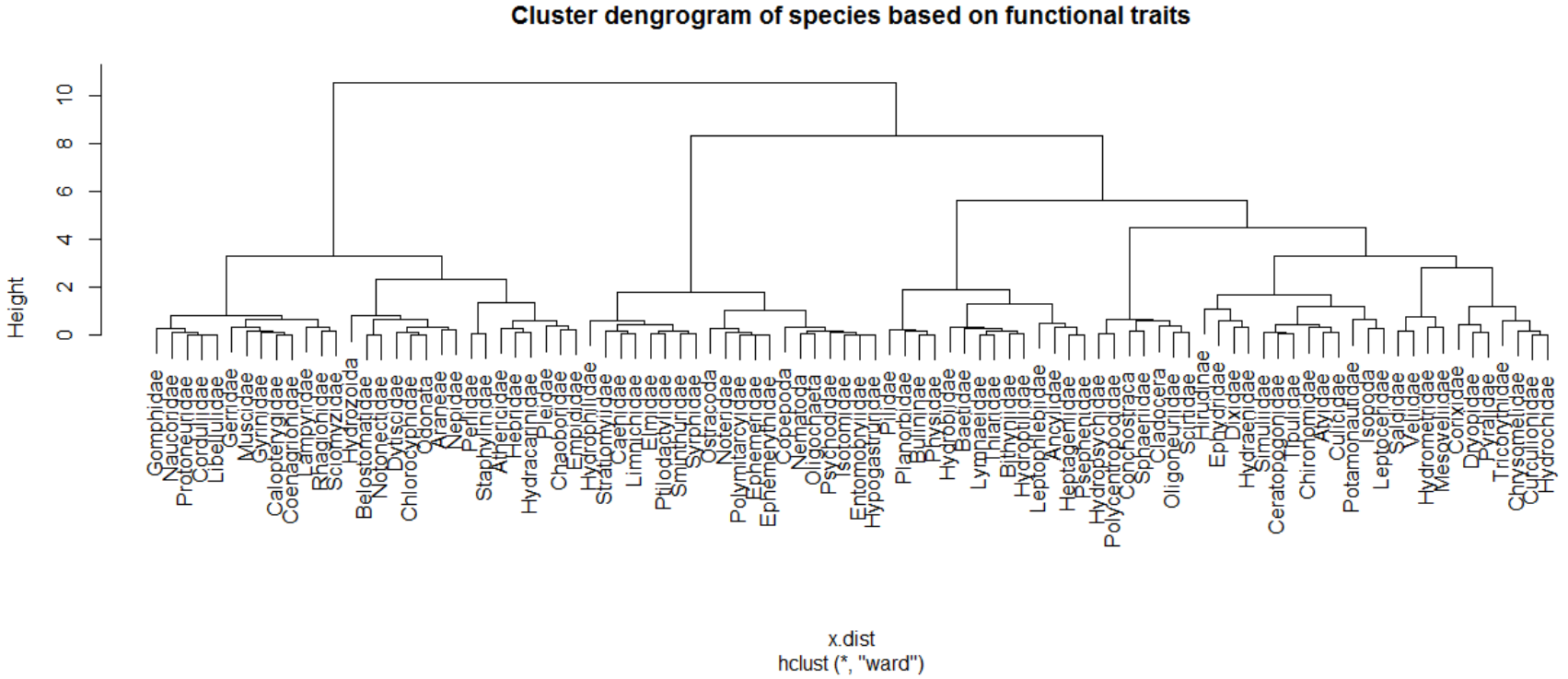


Figure 5.1 Cluster dendrogram of species based on their functional traits, dendrogram created using Gower dissimilarity; taxon groups on branches which are closer together have a higher number of functional traits in common. See Appendix II Table S5.4 for list of species, traits and their corresponding grouping

5.4 Results

5.4.1 Indices of functional diversity

Indices of functional diversity, functional dispersion, functional evenness and Rao's quadratic entropy showed no correlation with the proportion of positive MU samples at a site, this was true for all sites and also when lentic and lotic sites were tested separately (Table 5.1).

Table 5.1 Mixed effects model of functional indices at sites against the proportion of samples positive for MU DNA. Values of $p<0.05$ highlighted in bold. Models show how functional evenness and two measures of diversity do not vary between sites as the presence of MU DNA increases. Coefficient is the slope coefficient of regression. See Appendix II Table S5.1 for list of species, traits and their corresponding grouping. F.D. = Functional dispersion, F.E. = Functional evenness, R.Q. = Rao's quadratic entropy.

	All sites		Lotic sites		Lentic sites	
Indices	P value	Coefficient	P value	Coefficient	P value	Coefficient
F.D.	0.815	-0.074	0.746	0.174	0.494	-0.260
F.E.	0.837	0.055	0.700	0.162	0.955	-0.019
R.Q.	0.780	-0.128	0.669	0.326	0.382	-0.496

5.4.2 Functional group analysis

Two functional groups showed correlations with the proportion of positive environmental samples, with one group at lentic sites and another at lotic sites (Table 5.2). Group 9 was correlated with an increase in MU positive samples at lotic sites, from this group individual taxa which met the criteria of more than 10 organisms comprised primarily of small filtering collectors including: Cladocera, Sphaeriidae, Conchostraca, Ephemeroptera, Trichoptera (Hydropsychidae and Polycentropodidae) and larval Scirtidae (Figure 2). These species are often found in areas where there is a high abundance of organic matter and runoff in the water. Group 10 was correlated with an increase in positive samples at lentic sites and

included the springtails Entomobryidae and Isotomidae, Copepoda, Oligochaeta, Ostracoda, Ephemeroptera (Polymitarcyidae) and Diptera (Psychodidae). Again most species in this group feed on detritus, algae and other micro-organisms and predominantly reside in the soil or sediment of lentic ecosystems (Figure 2).

Table 5.2 Mixed effects model of functional groups at sites against the proportion of samples positive for MU. Values of $p<0.05$ highlighted in bold. Models show how functional groups 9 and 10 increase as the presence of MU DNA increases at sites in lentic and lotic habitats. See Appendix II Table S5.4 for list of species, traits and their corresponding grouping.

	All sites		Lotic sites		Lentic sites	
Group	p.value	Coefficient	p.value	Coefficient	p.value	Coefficient
1	0.590	-0.070	0.920	-0.019	0.256	-0.229
2	0.086	-0.643	0.949	-0.061	0.059	-0.727
3	0.895	-0.057	0.112	-1.344	0.321	0.469
4	0.161	-0.163	0.728	-0.067	0.075	-0.245
5	0.766	0.088	0.593	0.420	0.942	0.022
6	0.964	0.007	0.778	-0.060	0.292	0.785
7	0.968	-0.018	0.231	0.976	0.314	-0.518
8	0.445	1.570	0.953	0.166	0.129	4.881
9	0.083	0.598	**0.003**	**2.779**	0.610	0.171
10	**0.033**	**0.293**	0.736	0.129	**0.004**	**0.392**
11	0.196	-1.934	0.369	-2.354	0.393	-1.561

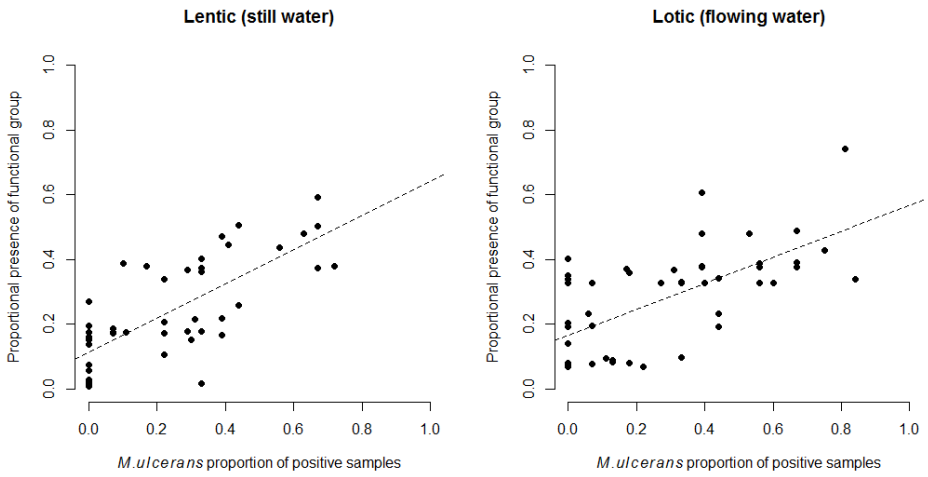


Figure 5.2 Model fitted values for the proportion of environmental samples, which were positive for MU DNA against the proportional presence of functional groups related to the breakdown and consumption of organic material in lentic habitats (group 10) and the filtering of organic material and sediment in lotic habitats (group 9). Dashed line represents line of best fit through the fitted points. See Appendix II Table S5.4 for list of species, traits and their corresponding grouping.

In all cases the AIC values were lower and the model was a significantly better fit when tested as a functional group rather than individual taxa. Group 9 showed the most improvement of model fit through grouping of organisms, with a significantly lower AIC value than any individual taxa tested. Group 10 showed a distinct improvement as a functional group over all the taxa in that group taken individually (see Appendix II Table S5.1 for all model results).

Taxonomic diversity within groups which had a positive relationship with MU DNA did not exhibit any change between sites of low DNA and high DNA (Table 5.3). As the group itself exhibits a higher correlation and better model fit between DNA than any individual taxa within it, it can be concluded that whilst the overall proportional presence of the group rose in relation to MU DNA and the number of different taxa stayed even, the presence of different taxa varied, providing the same niche space but from different organisms.

Table 5.3 Mixed effects model of biotic diversity within functional groups at sites against the proportion of samples positive for MU. Values of $p<0.05$ highlighted in bold. Models show how the diversity within functional groups 9 and 10 stays the same as the presence of MU DNA increases at all sites. Model 9 was tested at lotic sites, as this is where it showed a positive relationship with MU, whilst similarly group 10 was tested at lentic sites.

Functional group	Shannon		Simpson	
	P value	Coefficient	P value	Coefficient
10	0.510	-0.073	0.694	0.134
9	0.574	-0.037	0.561	-0.070

5.5 Discussion

For the first time, it is possible to directly link the presence of an infectious bacterium from a suspected aquatic origin, to functional groups of larger organisms, giving us a new insight into the functional role of communities in the potential distribution of a pathogen within the ecosystem. Furthermore these findings provide strong evidence that generalist environmentally persistent bacteria (in this case, a bacteria found in multiple taxa) such as MU can be associated with specific functional traits rather than taxonomic groups of organisms, increasing understanding of the environmental drivers of emerging diseases and subsequent transfer into humans.

The natural or human driven change from lentic to lotic environments brings with it a change in species morphology, yet a similarity in functional traits. In lentic environments, the numbers of environmental samples positive for the bacteria were correlated with a group of gathering collectors, Entomobryidae and Isotomidae, Oligochaeta, Copepoda, Ostracoda, Ephemeroptera (Polymitarcyidae) and Diptera (Psychodidae), which consume small microorganisms including microphytes, bacteria, biofilms and micro-invertebrates, and break down plant material and detritus. A similar relationship was seen in lotic environments with a group that feeds on small organisms, microphytes and detritus, albeit predominantly through

filter feeding, Cladocera, Sphaeriidae, Conchostraca, Ephemeroptera, Trichoptera (Hydropsychidae and Polycentropodidae) and larval Scirtidae. Within the lotic sites, body type and life history traits of organisms were also more suited to flowing water, with all of the species having adaptations more requisite to this environment, or of a semi-aquatic lifestyle at the water's edge (See Appendix II Tables S5.1-4). This builds on previous work on BU looking at feeding groups (Benbow et al. 2013), however by analysing a multitude of functional traits, specifically those relating to habitat preference, greater precision of grouping when testing between lotic and lentic sites has been possible, helping discover previously missed relationships.

The association with aquatic organisms that feed on decaying material, microorganisms and detritus, predominantly in or around the benthic layer, suggests that MU is potentially soil-borne or resides within organic material close to the bottom of a water-body. Previous theoretical studies have also identified sediment dwelling species as having a potential role in the presence of the bacteria (Roche et al. 2013), in particular Oligochaete worms which form part of group 10, although bear no significant relationship as an individual taxon in this study, highlighting the importance of testing functional groupings. The identification of such species suggests the niche of the bacteria is within sites of high organic content, most likely in the sediment. It is also possible that through decomposition and consumption of this material it is being brought into the wider animal community. The identification of a relationship with filter feeders in lotic sites also suggests these organisms filter what is leeched from suitable stagnant areas after rainfall; such an influx of material, including microorganisms would allow them to flourish, potentially concentrating the bacterium in a new area.

The lack of a link between changes in functional diversity or functional evenness across sites and varying levels of positive samples suggests that from an overall

functional perspective the bacteria can infect a range of habitats. This agrees with recent studies which have suggested the bacterium is relatively ubiquitous in endemic countries (Williamson et al. 2008).

As MU DNA increased in lentic and lotic habitats, the unchanging level of diversity yet the variation in which taxa were present within group 10 and group 9 respectively, suggests that from a population ecology perspective the dilution effect (Schmidt and Ostfeld 2001; Keesing et al. 2010) may need a certain level of host specificity before becoming apparent. Whilst organisms sharing a number of traits are conducive to the bacteria's propagation, which with current understanding of the bacillus's biology suggests a likely symbiotic relationship (Stinear et al. 2007; Stinear et al. 2008; Roche et al. 2013), the homogenous diversity and variation in taxa within the group between sites, suggests the bacteria are able to utilise the niche provided by any of these species without significant detriment to survival. This ability to utilise a number of hosts with varying immune responses and life histories could be a key trait of an emerging pathogenic species. The unusually slow growing nature of the bacterium (Portaels et al. 2008) suggests that rapid evolution to colonise different hosts in this instance is unlikely. It is possible therefore that counter to the dilution theory, the pressure of occupying a niche with a wide range of taxonomically distinct hosts has pushed it to evolve traits that allow the undermining of numerous defences (George et al. 1999; George et al. 2000; Stinear et al. 2004; Stinear et al. 2007), thereby creating a species capable of jumping from an unrelated source into humans. Whilst this may appear to counter the notion that human disturbance is a cause of disease, potential diversity loss in the groups presented here are not necessarily a factor of human impacts. At this trophic level, intermediate disturbance can allow a variety of species to inhabit a similar niche space, thereby increasing diversity (Hobbs and Huenneke 1992; Reynolds et al. 1993; Kimberling et al. 2001; Molino and Sabatier 2001;

McKinney 2002; Roxburgh et al. 2004). Caution is required therefore when discussing theories such as the dilution effect, whilst disturbance might be a pre-cursor to disease emergence; it is not necessarily a generic loss in biodiversity which is pressuring microbes to become pathogenic, particularly emerging ones.

Whilst the outcome of this study is of significant interest in MU research, the central result shows that functional changes, and in particular the use of functional groups, rather than taxonomic indicators and indices of biodiversity, can in some instances represent a better framework for characterising macro-ecological patterns responsible for infectious generalist pathogen presence in nature. Ecosystem changes, biodiversity alteration and functional diversity modification could impact on these categories of disease agents, and as a consequence developing research programmes taking into consideration local and regional changes to functional communities in both time and space is strongly recommended. This type of approach to ecological medicine opens up a whole new way to better understand the environment as a set of functional pathways constituted of interchangeable host carrier species that underpin the spread of some pathogens across the various levels of organisation of an ecosystem.

5.6 References

Baayen, R. H., 2011. *languageR: Data sets and functions with "Analyzing Linguistic Data: A practical introduction to statistics R package version 1.2*. [Accessed 01/02/2013].

Bates, D., Maechler, M. and Bolker, B., 2012. *Linear mixed-effects models using S4 classes.R package version 0.999999-0*. [Accessed 01/02/2013].

Benbow, M. E., Kimbirauskas, R., McIntosh, M. D., Williamson, H., Quaye, C., Boakye, D., Small, P. L. C. and Merritt, R. W., 2013. Aquatic Macroinvertebrate Assemblages of Ghana, West Africa: Understanding the Ecology of a Neglected Tropical Disease. *EcoHealth*, 1-16.

Benbow, M. E., Williamson, H., Kimbirauskas, R., McIntosh, M. D., Kolar, R., Quaye, C., Akpabey, F., Boakye, D., Small, P. and Merritt, R. W., 2008. Aquatic Invertebrates as Unlikely Vectors of Buruli Ulcer Disease. *Emerging Infectious Diseases*, 14 (8), 1247-1254.

BjØrnstad, O. N. and Falck, W., 2001. Nonparametric spatial covariance functions: Estimation and testing. *Environmental and Ecological Statistics*, 8 (1), 53-70.

Botta-Dukát, Z., 2005. Rao's quadratic entropy as a measure of functional diversity based on multiple traits. *Journal of Vegetation Science*, 16 (5), 533-540.

Carolan, K., Garchitorena, A., Garcia-Pena, G., Morris, A., Landier, J., Fontanet, A., Le Gall, P., Texier, G., Marsollier, L., Gozlan, R.E., Eyangoh, S., Lo Seen, D., and Guégan, J-F., (2014) Topography and land cover of the watershed predicts the distribution of the pathogen *Mycobacterium ulcerans* in aquatic insects. *PLoS Neglected Tropical Diseases* (in Press).

Chapin, F. S., Zavaleta, E. S., Eviner, V. T., Naylor, R. L., Vitousek, P. M., Reynolds, H. L., Hooper, D. U., Lavorel, S., Sala, O. E., Hobbie, S. E.,

Mack, M. C. and Diaz, S., 2000. Consequences of changing biodiversity. *Nature*, 405 (6783), 234-242.

Cummins, K. W., Merritt, R. W. and Andrade, P. C. N., 2005. The use of invertebrate functional groups to characterize ecosystem attributes in selected streams and rivers in south Brazil. *Studies on Neotropical Fauna and Environment*, 40 (1), 69-89.

Debacker, M., Aguiar, J., Steunou, C., Zinsou, C., Meyers, W. M., Guédénon, A., Scott, J. T., Dramaix, M. and Portaels, F., 2004. *Mycobacterium ulcerans* Disease (Buruli ulcer) in Rural Hospital, Southern Benin, 1997–2001. *Emerging Infectious Diseases*, 10 (8).

Deshayes, C., Angala, S. K., Marion, E., Brandli, I., Babonneau, J., Preisser, L., Eyangoh, S., Delneste, Y., Legras, P., De Chastellier, C., Stinear, T. P., Jackson, M. and Marsollier, L., 2013. Regulation of Mycolactone, the *Mycobacterium ulcerans* Toxin, Depends on Nutrient Source. *PLoS Neglected Tropical Diseases*, 7 (11), e2502.

Dizney, L. J. and Ruedas, L. A., 2009. Increased host species diversity and decreased prevalence of Sin Nombre virus. *Emerging Infectious Diseases*, 15 (7), 1012.

Dolnicar, S., Grabler, K., Mazanec, J. A., Woodside, A. G., Crouch, G. I., Oppermann, M. and Sakai, M. Y., 1999. *A tale of three cities: perceptual charting for analysing destination images*. CABI Publishing.

Elwell, L. C. S., Kerans, B. L. and Zickovich, J., 2009. Host–parasite interactions and competition between tubificid species in a benthic community. *Freshwater Biology*, 54 (8), 1616-1628.

Ezenwa, V. O., Godsey, M. S., King, R. J. and Guptill, S. C., 2006. Avian diversity and West Nile virus: testing associations between biodiversity and infectious disease risk. *Proceedings of the Royal Society B: Biological Sciences*, 273 (1582), 109-117.

Fyfe, J. A. M., Lavender, C. J., Handasyde, K. A., Legione, A. R., O'Brien, C. R., Stinear, T. P., Pidot, S. J., Seemann, T., Benbow, M. E., Wallace, J. R., McCowan, C. and Johnson, P. D. R., 2010. A Major Role for Mammals in the Ecology of *Mycobacterium ulcerans*. *PLoS Neglected Tropical Diseases*, 4 (8), e791.

Garrett, K. A., Zúñiga, L. N., Roncal, E., Forbes, G. A., Mundt, C. C., Su, Z. and Nelson, R. J., 2009. Intraspecific functional diversity in hosts and its effect on disease risk across a climatic gradient. *Ecological Applications*, 19 (7), 1868-1883.

George, K. M., Chatterjee, D., Gunawardana, G., Welty, D., Hayman, J., Lee, R. and Small, P. L. C., 1999. Mycolactone: A Polyketide Toxin from *Mycobacterium ulcerans* Required for Virulence. *Science*, 283 (5403), 854.

George, K. M., Pascopella, L., Welty, D. M. and Small, P. L. C., 2000. A *Mycobacterium ulcerans* Toxin, Mycolactone, Causes Apoptosis in Guinea Pig Ulcers and Tissue Culture Cells. *Infection and Immunity.*, 68 (2), 877-883.

Gleick, P. H., 2003. Global Freshwater Resources: Soft-Path Solutions for the 21st Century. *Science*, 302 (5650), 1524-1528.

Gower, J. C. and Legendre, P., 1986. Metric and Euclidean properties of dissimilarity coefficients. *Journal of Classification*, 3 (1), 5-48.

Harvell, C. D., Mitchell, C. E., Ward, J. R., Altizer, S., Dobson, A. P., Ostfeld, R. S. and Samuel, M. D., 2002. Climate Warming and Disease Risks for Terrestrial and Marine Biota. *Science*, 296 (5576), 2158-2162.

Hobbs, R. J. and Huenneke, L. F., 1992. Disturbance, diversity, and invasion: implications for conservation. *Conservation Biology*, 6 (3), 324-337.

Johnson, P. T. J. and Thieltges, D. W., 2010. Diversity, decoys and the dilution effect: how ecological communities affect disease risk. *The Journal of Experimental Biology*, 213 (6), 961-970.

Jones, K. E., Patel, N. G., Levy, M. A., Storeygard, A., Balk, D., Gittleman, J. L. and Daszak, P., 2008. Global trends in emerging infectious diseases. *Nature*, 451 (7181), 990-993.

Keesing, F., Belden, L. K., Daszak, P., Dobson, A., Harvell, C. D., Holt, R. D., Hudson, P., Jolles, A., Jones, K. E., Mitchell, C. E., Myers, S. S., Bogich, T. and Ostfeld, R. S., 2010. Impacts of biodiversity on the emergence and transmission of infectious diseases. *Nature*, 468 (7324), 647-652.

Keesing, F., Brunner, J., Duerr, S., Killilea, M., LoGiudice, K., Schmidt, K., Vuong, H. and Ostfeld, R. S., 2009. Hosts as ecological traps for the vector of Lyme disease. *Proceedings of the Royal Society B: Biological Sciences*, 276 (1675), 3911-3919.

Kimberling, D. N., Karr, J. R. and Fore, L. S., 2001. Measuring human disturbance using terrestrial invertebrates in the shrub–steppe of eastern Washington (USA). *Ecological Indicators*, 1 (2), 63-81.

Laliberté, E. and Legendre, P., 2010. A distance-based framework for measuring functional diversity from multiple traits. *Ecology*, 91 (1), 299-305.

Laliberté, E. and Shipley, B., 2011. *FD: measuring functional diversity from multiple traits, and other tools for functional ecology. R package version 1.0-11*. [Accessed 01/02/2013].

Lamont, E. A., Bannantine, J. P., Armién, A., Ariyakumar, D. S. and Sreevatsan, S., 2012. Identification and characterization of a spore-like morphotype in chronically starved Mycobacterium avium subsp. paratuberculosis cultures. *PloS one*, 7 (1), e30648.

LoGiudice, K., Ostfeld, R. S., Schmidt, K. A. and Keesing, F., 2003. The ecology of infectious disease: effects of host diversity and community composition on Lyme disease risk. *Proceedings of the National Academy of Sciences*, 100 (2), 567-571.

Marsollier, L., Robert, R., Aubry, J., Saint Andre, J.-P., Kouakou, H., Legras, P., Manceau, A.-L., Mahaza, C. and Carbonnelle, B., 2002. Aquatic Insects as a Vector for *Mycobacterium ulcerans*. *Applied and Environmental Microbiology*, 68 (9), 4623-4628.

Marsollier, L., Stinear, T., Aubry, J., Saint André, J. P., Robert, R., Legras, P., Manceau, A.-L., Audrain, C., Bourdon, S., Kouakou, H. and Carbonnelle, B., 2004a. Aquatic Plants Stimulate the Growth of and Biofilm Formation by *Mycobacterium ulcerans* in Axenic Culture and Harbor These Bacteria in the Environment. *Applied and Environmental Microbiology*, 70 (2), 1097-1103.

Marsollier, L., Sévérin, T., Aubry, J., Merritt, R. W., Saint André, J.-P., Legras, P., Manceau, A.-L., Chauty, A., Carbonnelle, B. and Cole, S. T., 2004b. Aquatic Snails, Passive Hosts of *Mycobacterium ulcerans*. *Applied and Environmental Microbiology*, 70 (10), 6296-6298.

Marston, B. J., Diallo, M. O., Horsburgh, C. R., Diomande, I., Saki, M. Z., Kanga, J. M., Patrice, G., Lipman, H. B., Ostroff, S. M. and Good, R. C., 1995. Emergence of Buruli ulcer disease in the Daloa region of Cote d'Ivoire. *The American Journal of Tropical Medicine and Hygiene*, 52 (3), 219-224.

McKinney, M., 2008. Effects of urbanization on species richness: A review of plants and animals. *Urban Ecosystems*, 11 (2), 161-176.

McKinney, M. L., 2002. Do human activities raise species richness? Contrasting patterns in United States plants and fishes. *Global Ecology and Biogeography*, 11 (4), 343-348.

McKinney, M. L., 2006. Urbanization as a major cause of biotic homogenization. *Biological Conservation*, 127 (3), 247-260.

Merritt, R. W., Benbow, M. E. and Small, P. L. C., 2005. Unraveling an Emerging Disease Associated with Disturbed Aquatic Environments: The Case of Buruli Ulcer. *Frontiers in Ecology and the Environment*, 3 (6), 323-331.

Merritt, R. W., Walker, E. D., Small, P. L. C., Wallace, J. R., Johnson, P. D. R., Benbow, M. E. and Boakye, D. A., 2010. Ecology and Transmission of Buruli Ulcer Disease: A Systematic Review. *PLoS Neglected Tropical Diseases*, 4 (12), e911.

Meybeck, M., 2003. Global analysis of river systems: from Earth system controls to Anthropocene syndromes. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 358 (1440), 1935-1955.

Meyers, W. M., Tignokpa, N., Priuli, G. B. and Portaels, F., 1996. *Mycobacterium ulcerans* infection (Buruli ulcer): first reported patients in Togo. *British Journal of Dermatology*, 134 (6), 1116-1121.

Molino, J.-F. and Sabatier, D., 2001. Tree diversity in tropical rain forests: a validation of the intermediate disturbance hypothesis. *Science*, 294 (5547), 1702-1704.

Morris, A., Gozlan, R. E., Hassani, H., Andreou, D., Couppié, P. and Guégan, J.-F., 2014. Complex temporal climate signals drive the emergence of human water-borne disease. *Emerging Microbes and Infections*, 3 (8), e56.

Mosi, L., Williamson, H., Wallace, J. R., Merritt, R. W. and Small, P. L. C., 2008. Persistent Association of *Mycobacterium ulcerans* with West African Predaceous Insects of the Family Belostomatidae. *Applied and Environmental Microbiology*, 74 (22), 7036-7042.

Ogden, N. H. and Tsao, J. I., 2009. Biodiversity and Lyme disease: Dilution or amplification? *Epidemics*, 1 (3), 196-206.

Ostfeld, R. S., 2009. Biodiversity loss and the rise of zoonotic pathogens. *Clinical Microbiology and Infection*, 15 (s1), 40-43.

Ostfeld, R. S. and Keesing, F., 2000a. Biodiversity and disease risk: the case of Lyme disease. *Conservation Biology*, 14 (3), 722-728.

Ostfeld, R. S. and Keesing, F., 2000b. Biodiversity series: the function of biodiversity in the ecology of vector-borne zoonotic diseases. *Canadian Journal of Zoology*, 78 (12), 2061-2078.

Petchey, O. L. and Gaston, K. J., 2002. Functional diversity (FD), species richness and community composition. *Ecology Letters*, 5 (3), 402-411.

Petchey, O. L. and Gaston, K. J., 2006. Functional diversity: back to basics and looking forward. *Ecology letters*, 9 (6), 741-758.

Petchey, O. L. and Gaston, K. J., 2007. Dendrograms and measuring functional diversity. *Oikos*, 116 (8), 1422-1426.

Petchey, O. L. and Gaston, K. J., 2009. Dendrograms and measures of functional diversity: a second instalment. *Oikos*, 118 (7), 1118-1120.

Portaels, F., Meyers, W. M., Ablordey, A., Castro, A. G., Chemlal, K., de Rijk, P., Elsen, P., Fissette, K., Fraga, A. G., Lee, R., Mahrous, E., Small, P. L. C., Stragier, P., Torrado, E., Van Aerde, A., Silva, M. T. and Pedrosa, J., 2008. First Cultivation and Characterization of *Mycobacterium ulcerans* from the Environment. *PLoS Neglected Tropical Diseases*, 2 (3), e178.

Rahel, F. J., 2002. Homogenization of freshwater faunas. *Annual Review of Ecology and Systematics*, 291-315.

Randolph, S. E. and Dobson, A. D. M., 2012. Pangloss revisited: a critique of the dilution effect and the biodiversity-buffers-disease paradigm. *Parasitology*, 139 (07), 847-863.

Reynolds, C. S., Padisák, J. and Sommer, U., 1993. Intermediate disturbance in the ecology of phytoplankton and the maintenance of species diversity: a synthesis. *Hydrobiologia*, 249 (1-3), 183-188.

Rigaud, T., Perrot-Minnot, M.-J. and Brown, M. J. F., 2010. Parasite and host assemblages: embracing the reality will improve our knowledge of parasite transmission and virulence. *Proceedings of the Royal Society B: Biological Sciences*, 277 (1701), 3693-3702.

Roche, B., Benbow, M. E., Merritt, R., Kimbirauskas, R., McIntosh, M., Small, P. L. C., Williamson, H. and Guégan, J.-F., 2013. Identifying the Achilles heel of multi-host pathogens: the concept of keystone 'host'species illustrated by *Mycobacterium ulcerans* transmission. *Environmental Research Letters*, 8 (4), 045009.

Roche, B., Dobson, A. P., Guégan, J.-F. and Rohani, P., 2012. Linking community and disease ecology: the impact of biodiversity on pathogen transmission. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367 (1604), 2807-2813.

Roche, B. and Guégan, J.-F., 2011. Ecosystem dynamics, biological diversity and emerging infectious diseases. *Comptes Rendus Biologies*, 334 (5–6), 385-392.

Roche, B., Rohani, P., Dobson, A. P. and Guégan, J.-F., 2013. The impact of community organization on vector-borne pathogens. *The American Naturalist*, 181 (1), 1-11.

Roxburgh, S. H., Shea, K. and Wilson, J. B., 2004. The intermediate disturbance hypothesis: patch dynamics and mechanisms of species coexistence. *Ecology*, 85 (2), 359-371.

Sachs, J. L., Skophammer, R. G. and Regus, J. U., 2011. Evolutionary transitions in bacterial symbiosis. *Proceedings of the National Academy of Sciences*, 108 (Supplement 2), 10800-10807.

Schmidt, K. A. and Ostfeld, R. S., 2001. Biodiversity and the dilution effect in disease ecology. *Ecology*, 82 (3), 609-619.

Searle, C. L., Biga, L. M., Spatafora, J. W. and Blaustein, A. R., 2011. A dilution effect in the emerging amphibian pathogen Batrachochytrium dendrobatidis. *Proceedings of the National Academy of Sciences*, 108 (39), 16322-16326.

Sehgal, R. N. M., 2010. Deforestation and avian infectious diseases. *The Journal of Experimental Biology*, 213 (6), 955-960.

Stienstra, Y., van der Graaf, W. T. A., Asamoa, K. and van der Werf, T. S., 2002. Beliefs and attitudes toward Buruli ulcer in Ghana. *American Journal of Tropical Medicine and Hygiene*, 67 (2), 207-213.

Stinear, T. P., Mve-Obiang, A., Small, P. L. C., Frigui, W., Pryor, M. J., Brosch, R., Jenkin, G. A., Johnson, P. D. R., Davies, J. K., Lee, R. E., Adusumilli, S., Garnier, T., Haydock, S. F., Leadlay, P. F. and Cole, S. T., 2004. Giant plasmid-encoded polyketide synthases produce the macrolide toxin of *Mycobacterium ulcerans*. *Proceedings of the National Academy of Sciences of the United States of America*, 101 (5), 1345-1349.

Stinear, T. P., Seemann, T., Harrison, P. F., Jenkin, G. A., Davies, J. K., Johnson, P. D. R., Abdellah, Z., Arrowsmith, C., Chillingworth, T. and Churcher, C., 2008. Insights from the complete genome sequence of Mycobacterium marinum on the evolution of Mycobacterium tuberculosis. *Genome research*, 18 (5), 729-741.

Stinear, T. P., Seemann, T., Pidot, S., Frigui, W., Reysset, G., Garnier, T., Meurice, G., Simon, D., Bouchier, C., Ma, L., Tichit, M., Porter, J. L., Ryan, J., Johnson, P. D. R., Davies, J. K., Jenkin, G. A., Small, P. L. C., Jones, L. M., Tekaia, F., Laval, F., Daffé, M., Parkhill, J. and Cole, S. T., 2007. Reductive evolution and niche adaptation inferred from the genome of *Mycobacterium ulcerans*, the causative agent of Buruli ulcer. *Genome Research*, 17 (2), 192-200.

Swaddle, J. P. and Calos, S. E., 2008. Increased avian diversity is associated with lower incidence of human West Nile infection: observation of the dilution effect. *PloS ONE*, 3 (6), e2488.

van Ravensway, J., Benbow, M. E., Tsonis, A. A., Pierce, S. J., Campbell, L. P., Fyfe, J. A. M., Hayman, J. A., Johnson, P. D. R., Wallace, J. R. and Qi, J.,

2012. Climate and Landscape Factors Associated with Buruli Ulcer Incidence in Victoria, Australia. *PLoS ONE*, 7 (12), e51074.

Villéger, S., Mason, N. W. H. and Mouillot, D., 2008. New Multidimensional Functional Diversity Indices for a Multifaceted Framework in Functional Ecology. *Ecology*, 89 (8), 2290-2301.

Vorosmarty, C. J., McIntyre, P. B., Gessner, M. O., Dudgeon, D., Prusevich, A., Green, P., Glidden, S., Bunn, S. E., Sullivan, C. A., Liermann, C. R. and Davies, P. M., 2010. Global threats to human water security and river biodiversity. *Nature*, 467 (7315), 555-561.

Williamson, H. R., Benbow, M. E., Campbell, L. P., Johnson, C. R., Sopoh, G., Barogui, Y., Merritt, R. W. and Small, P. L. C., 2012. Detection of *Mycobacterium ulcerans* in the Environment Predicts Prevalence of Buruli Ulcer in Benin. *PLoS Neglected Tropical Diseases*, 6 (1), e1506.

Williamson, H. R., Benbow, M. E., Nguyen, K. D., Beachboard, D. C., Kimbirauskas, R. K., McIntosh, M. D., Quaye, C., Ampadu, E. O., Boakye, D., Merritt, R. W. and Small, P. L. C., 2008. Distribution of *Mycobacterium ulcerans* in Buruli Ulcer Endemic and Non-Endemic Aquatic Sites in Ghana. *PLoS Neglected Tropical Diseases*, 2 (3), e205.

Willson, S. J., Kaufman, M. G., Merritt, R. W., Williamson, H. R., Malakauskas, D. M. and Benbow, M. E., 2013. Fish and amphibians as potential reservoirs of *Mycobacterium ulcerans*, the causative agent of Buruli ulcer disease. *Infection Ecology and Epidemiology*, 3.

Woolhouse, M. E. and Gowtage-Sequeria, S., 2005. Host range and emerging and reemerging pathogens. *Emerging Infectious Diseases*, 11 (12), 1842-1847.

World Health Organisation, 2001a. *Buruli Ulcer. Diagnosis of Mycobacterium ulcerans disease. A manual for health care providers*. Geneva. Available from: http://www.who.int/iris/handle/10665/67000#sthash.cioj71jV.dpuf.

World Health Organisation, 2001b. *Buruli Ulcer Management of Mycobacterium ulcerans disease*. World Health Organisation. Available from: http://whqlibdoc.who.int/hq/2001/WHO_CDS_CPE_GBUI_2001.3.pdf.

Chapter 6 Land-use/deforestation-mediated food web collapse promotes infectious disease carrying species in tropical freshwater systems.

6.1 Summary

Many of the approximately 250 known human emerging infectious diseases are directly linked to regions of tropical rainforest, of these many originate from freshwater aquatic systems. As urbanisation, population encroachment and deforestation increase across the globe, their effect on tropical ecosystems can be dramatic, the loss of function and collapse in food webs can often lead to opportunistic, highly fecund species being able to take hold and flourish in a more restricted niche space. Such shifts in community organization invariably lead to a change in the dynamics of tropical water-borne diseases. Understanding these changes is central in predicting future outbreaks and identifying hotspots for many emerging and re-emerging pathogens. Here for the first time food web metrics, stable isotopic data and quantitative PCR have been used to link deforestation and land-use pressures in the tropics to the emergence of infectious diseases. The results show that these rapid terrestrial habitat changes lead to the collapse of freshwater food web structure. This collapse is accompanied by an increase in basal opportunistic and highly fecund generalist hosts of MU. These species live at the interfaces between soil and water or between soil and vegetation, a huge source of potentially millions of new microbial species which could make the leap from environmentally persistent to infectious. As urbanisation, agriculture and deforestation intensify similar trends are expected to become apparent for many other generalists emerging and re-emerging infections.

6.2 Introduction

Despite a large percentage of the approximately 250 EID's being directly linked to tropical forests (Despommier et al. 2006), the relationship between these biotic regions and EID's is not yet fully understood. Many EID's originate from aquatic systems within the forests, either directly or via intermediate hosts or vectors, with the water also often acting as an interface between the soil and the terrestrial environment. Within these areas, as a consequence of a rapid increase in global human population, settlements and encroachments, there is a similarly rapid decline in biological diversity, often as a result of change in land-use and deforestation (Rahel 2002; McKinney 2006; McKinney 2008). Current estimates suggest that about 10,000 to 20,000 freshwater species have either become extinct or are under severe threat (Harvell et al. 2002; Gleick 2003; Meybeck 2003; Vorosmarty et al. 2010), with direct consequences for freshwater ecosystem functioning and therefore the potential for dramatic shifts in associated infectious diseases dynamics (Chapin et al. 2000; Ostfeld and Keesing 2000; Ostfeld 2009; Keesing et al. 2010). Until now however, understanding of the consequences of such dramatic changes in host communities on the emergence of infectious pathogens has remained limited (Jones et al. 2008; Keesing et al. 2010; Roche and Guégan 2011; Randolph and Dobson 2012).

The effects of land-use and deforestation have primarily been concerned with established vectored diseases, notably malaria and leishmaniasis. These are a natural choice for study as large volumes of human cases can be directly compared to the surrounding environment and known vectors or hosts can be more readily surveyed and used as a proxy for potential infection rates. However, to understand the change in more generalist EID's, which often do not rely on a specific host or vector, it is necessary to assess in detail the shifts in communities that arise from land-use changes. Land use could have a strong impact on community

composition, with some species being more adapted to new habitat conditions than others, and in turn, it is key to characterise how each organism is related to the presence of a disease agent (i.e. host/non-host). Beyond this direct relationship, land use change can also result in a shift in the level of species interactions, through a modification of food-web structure. As this network is responsible for species regulation through direct (e.g. predation) and indirect effects (e.g. extinction cascade), it is crucial to understand this aspect of the community to assess the consequences of habitat change on generalist EID's.

Here based on a comprehensive investigation of tropical freshwater aquatic communities, across a gradient of land-use, the following techniques were employed to build detailed structural pictures of freshwater communities and assess their changes in relation to MU presence under varying levels of anthropogenic stress:

- Stable isotopic analysis to identify trophic positions of taxa within a community and to quantify the trophic spread of taxa across a site, in this instance calculated as niche width (Layman et al. 2007). This measurement can be considered to represent the total trophic niche width at a site (with a larger spread of isotopic data representing a larger spread of trophic niche availability) and a proxy for the total extent of trophic diversity of organisms.

- Food web models to assess community interactions. Food webs for each site were constructed using a combination of prior knowledge and the allometric diet breadth model (ADBM). A web consists of nodes and links, each node represents a single taxa and its links represent both consumers and those which it consumes. ADBM is based on foraging theory and uses both species size and diet breadth, which is derived from the potential

energy gained from other organisms, encounter rates and handling times (Beckerman et al. 2006; Petchey et al. 2008).

- Quantitative Polymerase Chain Reactions (qPCR) allows a quantitative measure of the bacterial load based on measures of DNA replicates within a sample, and can be used to identify changes in bacterial presence within individuals.

From this information it is possible to show for the first time how structural changes in hosts and non-host community networks influence the presence of MU. Such food web approaches to generalist EID agents, can be seen as a model for other similar environmentally-persistent pathogens, and notably bacteria in general.

6.3 Methods

6.3.1 Data collection

Seventeen sites in French Guiana were identified as endemic or having the potential to harbour MU from the previous data, (see chapter 2 and Morris et al. 2014), these sites represented gradients of land-use from urban, through agricultural to pristine rainforest, see chapter 2.3.1 for detailed descriptions of biotic surveys and 2.3.2 for abiotic measurments. For organism identification, preservation and measurement methodology see chapter 2.3.3, for stable isotope analysis see chapter 2.3.4 and 2.3.5 for DNA extraction and qPCR protocols.

6.3.2 Landscape data

Landscape data was extracted from two maps, firstly the 25m 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 (Figure 6.1). The second the Hansen's deforestation maps, which are a high resolution map of global forest change to 30m resolution with individual data for each year from 2001 to 2012 (Hansen et al. 2013). Each site location was plotted onto the land-cover maps and buffer zones of 500m, 1km and 2km were drawn around each. From within each buffer zone, land cover data in m^2 was extracted for the CORINE map and deforestation data for the years 2001 to 2012 for the Hansen's map. Data extraction was conducted in Quantum GIS using the LecoS landscape data analysis plugin (Jung 2013).

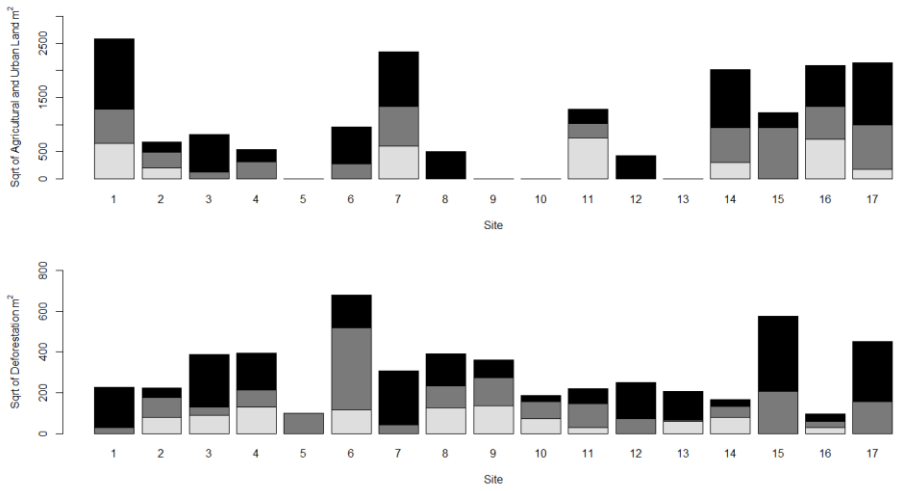


Figure 6.1 Bar plot showing the land cover data for each site, the top chart shows the m^2 cover of Agricultural and Urban land around each site within the 2km (black), 1km (dark grey) and 500m (light grey) buffer zones. The bottom chart shows the m^2 cover of deforestation within the 3 years prior to the surveys.

6.3.4 Trophic niche width and food web metric calculations of sites and aquatic invertebrates and vertebrates.

Trophic niche width (Layman et al. 2007) was calculated from the stable isotope data as the total convex hull area of the mean $\delta^{13}C$ and $\delta^{15}N$ isotopic readings in bi-plot space for the entire surveyed population of all organisms within a site (see Appendix III).

The ADBM was used to predict the links between nodes in each food web, these were subsequently checked for accuracy from prior knowledge and any biologically innacurate or missing links were updated accordingly.

Once a network was built a number of local metrics were calculated for each site. This included the connectance which is a measure of the number of node connections in a food web, or the mean proportional diet breadth of all species; the generality, a measure of the average number of differing organisms a taxa is able to feed on and the vulnerability, a measure of the average number of species an organism is preyed upon by.

6.3.7 Biodiversity measures

For each site taxa abundance was used to calculate both Shannon and Simpsons divesity indices. In addition species richness was estimated using jacknifing techniques based on the abundance of organisms found in each sample taken from a site. This gives a predictive estimate of the total species richness, based on the probability of species being missed during surveys (Heltshe and Forrester 1983).

6.3.8 Analysis

For each taxa identified during the surveys, the regional mean bacterial load was calculated, this encompassed all sites, survey dips and dates. In addition, each taxa's regional mean niche width, vulnerability, connectance and generality was also derived, weighted by their abundance at each site (i.e. if a species occurred predominantly at sites with low generality, its average generality metric would be lower, see supplementary materials in Appendix III).

To identify whether taxa with a higher mean bacterial load are most common at sites with a certain type of food web network (i.e high levels of connectance, generality or vulnerability). General additive models (GAM) with a tensor product

smoothing were performed between regional bacterial load and their mean regional food web metrics, non-significant metrics were removed until a final best fit model was found. Because many taxa had zero bacteria across all sites, a tweedie probability distribution was used, this distribution is well suited to zero-inflated models and provided the best fit to the data (Tweedie 1984)).

To identify how these metrics responded to a reduction in niche width of the population, GAM's of each metric; vulnerability, generality and connectance as a function of niche width were performed separately.

The same models were conducted for the abiotic metrics; *pH*, conductivity, temperature and dissolved oxygen, to identify whether the results are proxies for changes in the physical make-up of the site.

Relationships between land-cover data, biodiversity indices, species richness and niche width were identified at the local site level using GAMS' of the level of deforestation in m^2 within each buffer zone in addition to urban and agricultural cover (m^2). For deforestation, data from the preceding 3 years to the survey date was used; this was to allow 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 one year after 2008 when the last SPOT V satellite data for the agricultural and urban land-cover was taken, so it is possible to be confident that this land was present when the deforestation occurred.

In all models data was appropriately transformed to adhere to model assumptions, residuals of each model were tested for spatial autocorrelation using both Moran's *I* test (95% confidence interval) and correlogs smoothed with a spline function (95% pointwise bootstrap confidence intervals) (BjØrnstad and Falck 2001).

6.4 Results

For the 17 sites surveyed, over 3,600 invertebrates and fish were collected and catalogued. 90 different taxa were identified, of which 383 organisms representing 44 taxa tested positive for both IS2404 and KR (see supplementary materials in Appendix III) Bacterial concentrations ranged from 6 bacilli per mg of organism to 7,837 (mean = 266.5, standard deviation = 1,220) in those which were positive. The highest concentrations of bacteria were in species lower in the food chain, represented by a low $\delta^{15}N$ and a low $\delta^{13}C$, suggesting a diet high in aquatic algae, detritus, diatoms and similar food sources (Figure 6.2). Taxa in the upper quartile of bacterial load were predominantly invertebrates including; Corixidae, Caenidae, Coengrionidae, Baetidae, Veliidae, Leptophlebiidae and Simuliidae, with two fish species *Copella carsevenennsis* and *Rivulus lungi*. *C. carsevenennsis's* high bacterial load may however be a result of its feeding habits, as they are known to primarily consume Ephemeroptera larvae such as Caenidae and Baetidae, *R. lungi* may show similar behaviour although it is not as well documented. Despite this, it is not possible to discount their potential as hosts to the bacteria itself. Both these fish species are found in the bottom quartile of average niche width and are in the top 3 fish species for low niche width, suggesting they are most common at sites where there is a decline in trophic space (table S6.1 Appendix III); this may also pre-dispose them to feeding on bacterial hosts which thrive in these areas. All invertebrates in the upper quartile of bacterial load were functionally similar except for Veliidae, this was the only predatory taxa. Similar to *R. lungi*, it is quite likely however, that this taxa has been feeding on numerous basal species, it is one of the smallest predators surveyed and therefore unlikely to supplement its diet with other predatory organisms.

There was no relationship between organisms higher in bacterial load and any of the regionally calculated abiotic factors; *pH*, dissolved oxygen, temperature and conductivity.

The GAM between regional food web metrics and regional bacterial load identified two primary explanatory variables, vulnerability and generality (Adjusted R^2 = 0.295, deviance explained = 29.9%, p.values = 0.0003 and 0.002 for vulnerability and generality respectively) (Figure 6.3), connectance however, did not provide any significant explanatory power. Regional trophic niche width was closely correlated with generality and vulnerability, which both declined as niche width was reduced, (vulnerability; Adjusted R^2 = 0.321, deviance explained = 33%, p=<0.0001, generality; Adjusted R^2 = 0.465, deviance explained = 47.2%, p= <0.0001) (see Figure 6.4).

Local trophic niche width declined with increasing deforestation in the preceding 3 years and increasing levels of agricultural/urban land, the model, which fitted the data most accurately, occurred within the 1km buffer zone (R^2 = 0.58, p= 0.0008 on 14 d.f.), (Figure 6.5 and Table 6.2). This relationship was lost at 2km and 500m scales. This was potentially the result of there being too little or too much land-cover information at these spatial resolutions. A number of sites had no deforestation or agricultural/urban land-use within the 500m buffer zone and the data exhibited a highly skewed distribution, which could not be adequately normalised through transformation or using zero inflated models. At the 2km scale several sites showed more extreme levels of deforestation against the average, however much of this was over a kilometre from the sites, and less likely to have an effect on the conditions there. Biodiversity and species richness was unaffected by deforestation or agricultural practices (Table 6.3), suggesting the functional type of species found at a site changed rather than the number of taxonomically distinct organisms.

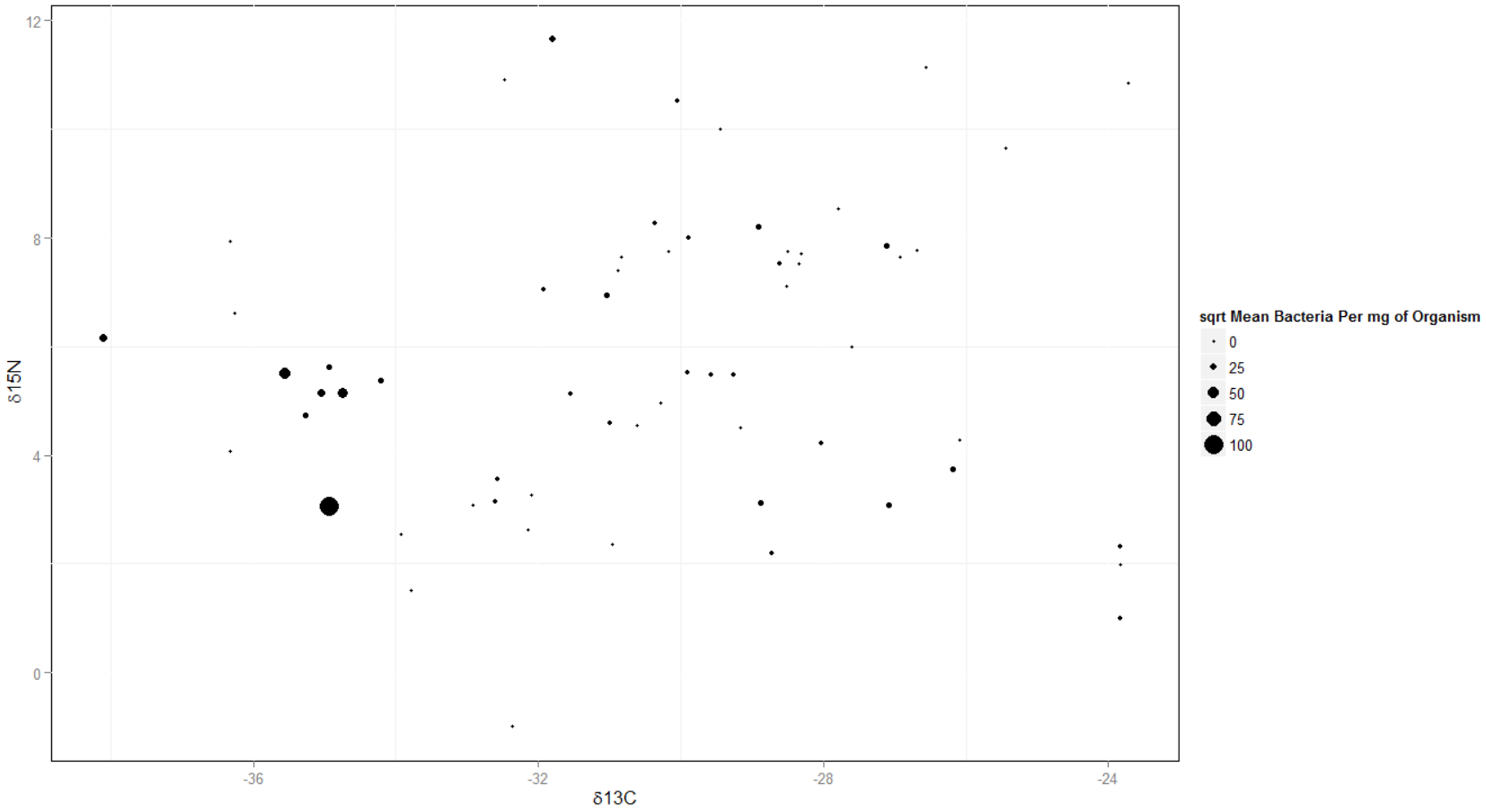


Figure 6.2 The average δ ^{15}N and a low δ^{13}C in bi-plot space for all recorded host and non-host organisms from the 17 sites. δ ^{15}N and δ^{13}C were derived for each taxa using the data from the 3 sites which were analysed for stable isotopic readings. For each point the square root transformed mean number of bacteria of that taxa is represented by the size of the circle.

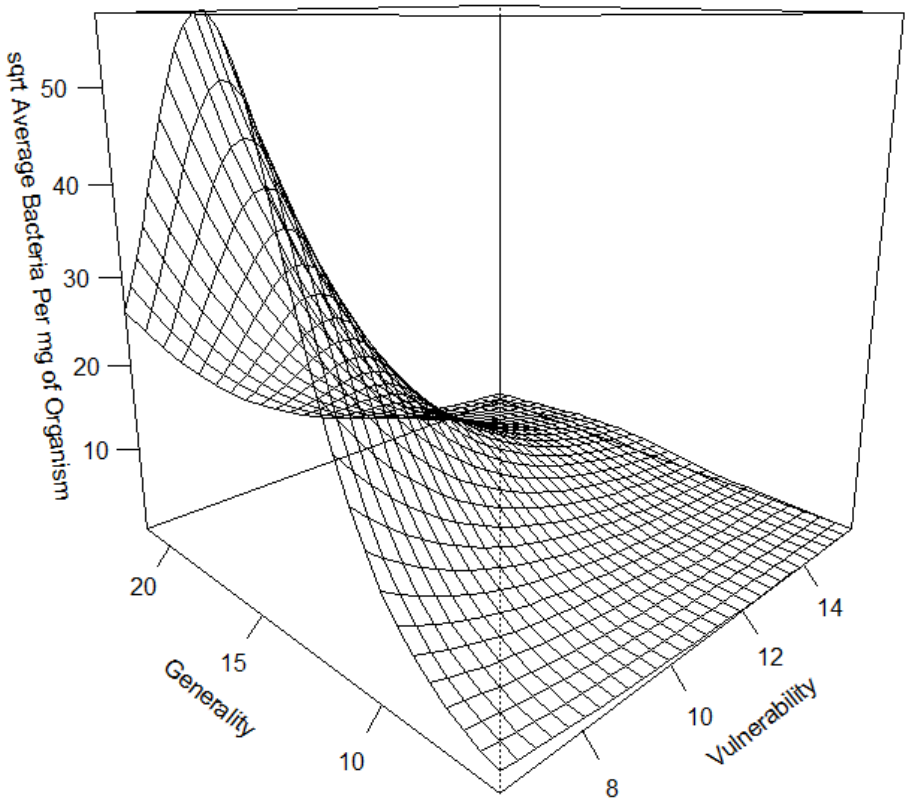


Figure 6.3 Plot showing the metrics of web networks that allow taxa, which on average carry a higher bacterial load to propagate. For all taxa, along the bottom axes are their mean regional food web metrics for vulnerability and generality (i.e. a measure of the food web metrics of sites at which they are most abundant), whilst the vertical axis shows their mean regional bacterial load of MU.

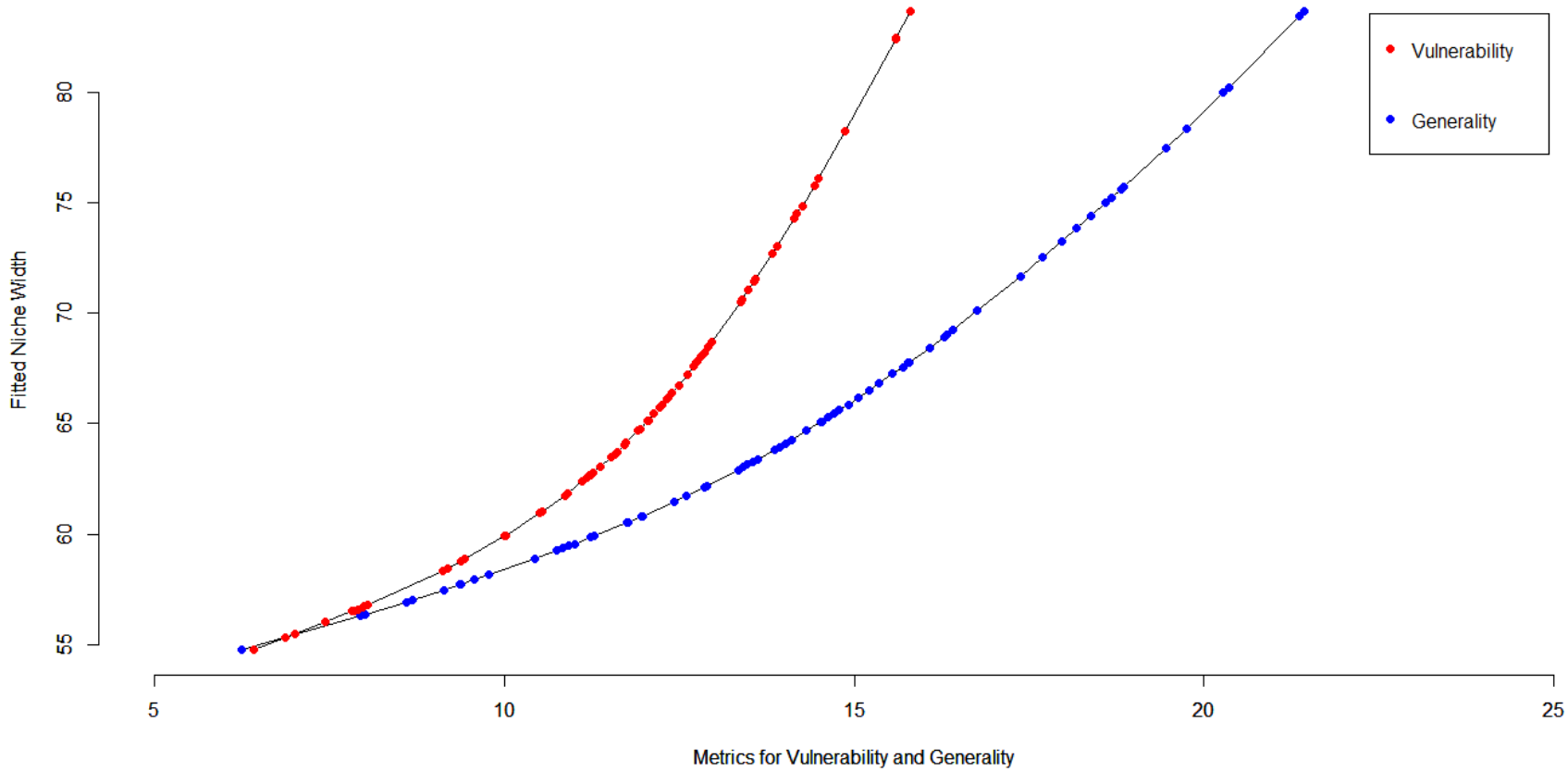


Figure 6.4 Plot showing the results of GAM's between the mean regional niche width of all organisms and their mean regional food web metrics; vulnerability and generality. The plot shows that as niche width decreases, vulnerability and generality of the food webs also decrease.

Table 6.2 Results of multiple linear regressions between calculated niche widths for the entire population of organisms surveyed at each site and the square meter cover of deforestation and agricultural/urban land within buffer zones of 500m, 1km and 2km. Coefficient is the slope coefficient, a negative value indicates a decline in explanatory variable as the dependant variable (niche width) increases.

Buffer zone	**Coefficient Agriculture and urban land**	**Coefficient Deforestation**	**R^2**	**Adjusted R^2**	**d.f.**	**p.value**
500m	-0.001	-0.017	0.060	-0.074	14	0.649
1km	**-0.020**	**-0.147**	**0.635**	**0.583**	**14**	**0.0008**
2km	-0.014	-0.061	0.179	0.062	14	0.250

Table 6.3 Results of multiple linear regressions between calculated biodiversity indices, actual and theoretical species richness calculated via jacknife for the entire population of organisms surveyed at each site and the square meter cover of deforestation and agricultural/urban land within buffer zones of 500m, 1km and 2km.

Indices	**Buffer zone**	**Coefficient Agriculture and urban land**	**Coefficient Deforestation**	**R^2**	**Adjusted R^2**	**d.f.**	**p.value**
Simpson	500m	-0.006	-0.007	0.250	0.143	14	0.132
	1km	-0.003	-0.013	0.111	-0.016	14	0.439
	2km	-0.003	0.026	0.135	0.012	14	0.360
Shannon	500m	-0.017	-0.013	0.080	-0.051	14	0.556
	1km	-0.067	-0.067	0.083	-0.048	14	0.544
	2km	-0.014	0.139	0.160	0.040	14	0.294
Species Richness	500m	0.011	0.046	0.063	-0.071	14	0.636
	1km	-0.035	-0.140	0.132	0.008	14	0.372
	2km	-0.035	0.097	0.074	-0.057	14	0.580
Jack-knifed Richness	500m	0.013	0.062	0.065	-0.068	14	0.623
	1km	-0.090	-0.322	0.135	0.012	14	0.361
	2km	-0.045	0.106	0.071	-0.060	14	0.593

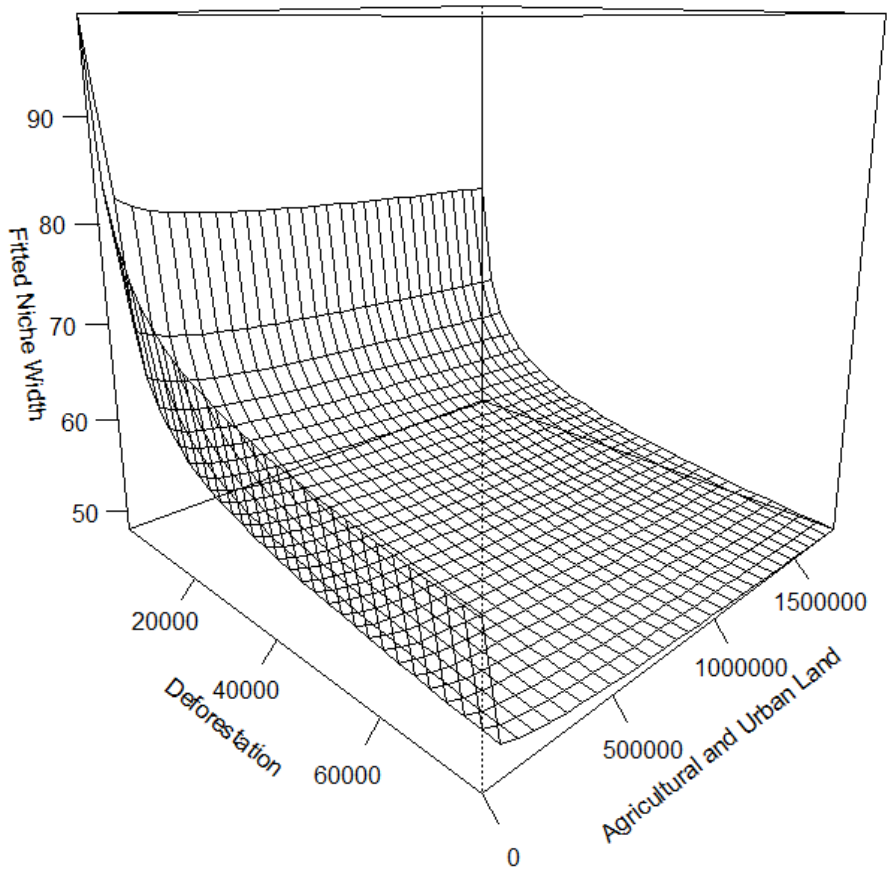


Figure 6.5 Plot showing the relationship between local site niche widths and the land cover in m^2 of agricultural and urban landscape and deforestation in a 1km buffer zone. The change to niche width caused by deforestation appears to be a steady decline, whilst the presence of agricultural land causes a sharp drop from where there is no agriculture, before exhibiting a similar steady decline in niche width as with deforestation, albeit less extreme.

6.5 Discussion

The decline in trophic niche width from deforestation and agricultural or urban intrusion has a powerful effect on aquatic food web networks. Most notably there is a decrease in the mean vulnerability of taxa, coupled with a decrease in generality. It is likely that these decreases are both being driven by a loss in the overall number of predators, allowing low-level trophic organisms to flourish through a knock-on process. This has an important effect on the potential for systems to harbour the bacterial pathogen MU, as taxa, which on average carry a high level of bacilli, are most abundant at sites where there is a very low level of vulnerability and a mid-level of generality (Figure 6.3). This response could be explained by hosts of the bacteria being low trophic-level, generalist organisms. For example, at sites with low vulnerability these species will flourish due to a lack of predation. However, if the predators at a site have a high generality the basal species will be more diverse leading to a wider, functionally more distinct niche space. This is supported by the relationship between high niche width and high generality. With a decrease in abundance of generalist predators as site disturbance increases, the overall average generality of the trophic network at a site declines, causing high levels of bacterial hosts to flourish. However, when this disturbance reaches its peak, a site becomes suitable for only a very few taxa, which have a decreased level of contact with jungle dwelling species and the more microbial rich soils found in and on the edge of the forest. As such the number of bacterial hosts begins again to decline. These results would explain why BU is most common in areas on the forest edge where there is likely a low level of vulnerability from top predator loss, and a mid-level of generality caused by an increase in niche space from intermediate disturbance (Roxburgh et al. 2004).

For the first time food web metrics and stable isotopic data have been used to link deforestation and land-use to the emergence of infectious diseases. The relationship

between disease, community structure and land cover is clearly established and represents a significant breakthrough for our understanding of other emergent infections. As urbanisation, agriculture and deforestation intensifies similar trends for other EIDs are expected to become apparent. Basal opportunistic and highly fecund species (i.e. *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 these basal taxa live at the interfaces between soil and water or between soil and vegetation, their high numbers and potential reduced investment in immunity (Nunn et al. 2003; Martin Ii et al. 2006; Martin et al. 2007; Lee et al. 2008) allow numerous opportunities for emerging microbes to take hold, or opportunistic microbes to make the leap from commensalism or symbiosis to infectious.

In terrestrial environments, recent increases in plague and leptospirosis have been linked to urbanisation and climatic shifts towards extreme weather events (Lau et al. 2010; Andrianaivoarimanana et al. 2013). As recently seen, there is a strong link between with BU and extreme rainfall (Morris et al. 2014b) along with the current evidence that fecund opportunists species also have a significant relationship with the disease agent (Roche et al. 2013). As more people come into contact with such combinations of land-use traits and extreme weather through global urbanisation, human settlement/encroachment and climate change, it is not unreasonable to predict that emergence and re-emergence of such infectious diseases will become more common.

These basal taxa could have been proxies for environmental conditions suitable for the bacteria. However the lack of a direct relationship between water chemistry and food web structure is puzzling. The decline of top predators in these tropical aquatic systems is most likely an indirect effect of deforestation and change in land-use. In effect, the loss of tree cover will significantly reduce the leaf litter and

carbon input into the water as well as increase bird, amphibian and reptile predation, with a direct knock-on effect on the fish communities. Also as a site loses its surrounding forest cover it becomes far more susceptible to extreme weather, previous publications have linked BU cases to high levels of rainfall (van Ravensway et al. 2012; Garchitorena et al. 2014; Morris et al. 2014b) and high levels of silting and flooding at an exposed site could cause a significant change in local and regional community structure.

Despite our increasing understanding of MU ecological niche, the process of infection to humans remains uncertain. The two main competing theories of mosquito (Fyfe et al. 2007; Fyfe et al. 2010) or water-bug (Marsollier et al. 2002; Marsollier et al. 2004b; Mosi et al. 2008; Marion et al. 2010) vector transmissions are still partially supported by our data, with both these taxa testing positive (Belostomatidae 4/34 pooled samples, Culicidae 1/13 pooled samples). With mosquitoes it may be a direct transfer from the environment to the larvae, although there is less evidence to support the potential for the bacilli to be maintained through to adulthood. Water-bugs are likely to become infected, or temporarily carry the bacteria after feeding on positive organisms and species such as Belostomatidae are highly mobile and can fly from site to site and easily pick up the bacteria from feeding where there are high numbers of positive taxa. Similarly mosquitoes are ubiquitous in such areas and will be more likely to flourish where there are fewer predators. Regardless of the theory, the potential for disease transmission is likely to increase if either a) there are more prey items carrying the bacteria or b) the presence of the bacteria in the environment is increased through the presence of suitable bacterial hosts.

In chapter 5 specific functional groups of taxa which were trophically low level basal organisms, were found to correlate with high levels of environmental bacilli in Ghana, Africa. These were similar to those found to contain high levels of the

bacilli in French Guiana. Whilst a direct comparison of taxa is difficult as only two family groups from lotic sites (lentic sites were not surveyed in this study), were found in both geographic regions. These two families were found to be positive for bacilli in French Guiana, Oligochaeta and Ostracoda. This is a significant breakthrough in understanding the ecological niche of the bacteria, previously Oligochaeta have been highlighted as having a potentially key relationship with MU (Roche et al. 2013), these results provide the first empirical evidence for this, they also go further and support the results of chapter 5, in that groups of functionally similar taxa are important in harbouring the bacilli.

The characterisation of a complex link between terrestrial, aquatic and infectious processes shows the importance of community structure on disease risks mapping and monitoring. Such an approach can be applied to the emergence of other generalist pathogens and could be central in predicting how human induced environmental change could drive the occurrence of new infections in the world.

6.6 References

Andrianaivoarimanana, V., Kreppel, K., Elissa, N., Duplantier, J-M., Carniel, E., Rajerison, M. and Jambou, R., 2013. Understanding the Persistence of Plague Foci in Madagascar. *PLoS Neglected Tropical Diseases*, 7 (11), e2382.

Beckerman, A. P., Petchey, O. L. and Warren, P. H., 2006. Foraging biology predicts food web complexity. *Proceedings of the National Academy of Sciences*, 103 (37), 13745-13749.

Benke, A. C., Huryn, A. D., Smock, L. A. and Wallace, J. B., 1999. Length-mass relationships for freshwater macroinvertebrates in North America with particular reference to the southeastern United States. *Journal of the North American Benthological Society*, 308-343.

BjØrnstad, O. N. and Falck, W., 2001. Nonparametric spatial covariance functions: Estimation and testing. *Environmental and Ecological Statistics*, 8 (1), 53-70.

Brou, T., Broutin, H., Elguero, E., Asse, H. and Guegan, J-F., 2008. Landscape Diversity Related to Buruli Ulcer Disease in Côte d'Ivoire. *PLoS Neglected Tropical Diseases*, 2 (7), e271.

Chapin, F. S., Zavaleta, E. S., Eviner, V. T., Naylor, R. L., Vitousek, P. M., Reynolds, H. L., Hooper, D. U., Lavorel, S., Sala, O. E., Hobbie, S. E., Mack, M. C. and Diaz, S., 2000. Consequences of changing biodiversity. *Nature*, 405 (6783), 234-242.

Despommier, D., Ellis, B. R. and Wilcox, B. A., 2006. The role of ecotones in emerging infectious diseases. *EcoHealth*, 3 (4), 281-289.

Domínguez, E., 2006. *Ephemeroptera de América del Sur*. Vol. 2. Pensoft Publishers.

Froese, R., Thorson, J. T. and Reyes, R. B., 2014. A Bayesian approach for estimating length-weight relationships in fishes. *Journal of Applied Ichthyology*, 30 (1), 78-85.

Fyfe, J. A. M., Lavender, C. J., Handasyde, K. A., Legione, A. R., O'Brien, C. R., Stinear, T. P., Pidot, S. J., Seemann, T., Benbow, M. E., Wallace, J. R., McCowan, C. and Johnson, P. D. R., 2010. A Major Role for Mammals in the Ecology of *Mycobacterium ulcerans*. *PLoS Neglected Tropical Diseases*, 4 (8), e791.

Fyfe, J. A. M., Lavender, C. J., Johnson, P. D. R., Globan, M., Sievers, A., Azuolas, J. and Stinear, T. P., 2007. Development and Application of Two Multiplex Real-Time PCR Assays for the Detection of *Mycobacterium ulcerans* in Clinical and Environmental Samples. *Applied and Environmental Microbiology*, 73 (15), 4733-4740.

Garchitorena, A., Roche, B., Kamgang, R., Ossomba, J., Babonneau, J., Landier, J., Fontanet, A., Flahault, A., Eyangoh, S. and Guégan, J.-F., 2014. *Mycobacterium ulcerans* Ecological Dynamics and Its Association with Freshwater Ecosystems and Aquatic Communities: Results from a 12-Month Environmental Survey in Cameroon. *PLoS Neglected Tropical Diseases*, 8 (5), e2879.

Gleick, P. H., 2003. Global Freshwater Resources: Soft-Path Solutions for the 21st Century. *Science*, 302 (5650), 1524-1528.

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O. and Townshend, J. R. G., 2013. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science*, 342 (6160), 850-853.

Harvell, C. D., Mitchell, C. E., Ward, J. R., Altizer, S., Dobson, A. P., Ostfeld, R. S. and Samuel, M. D., 2002. Climate Warming and Disease Risks for Terrestrial and Marine Biota. *Science*, 296 (5576), 2158-2162.

Heckman, C. W., 2006. *Encyclopedia of South American Aquatic Insects: Odonata-Anisoptera: Illustrated Keys to Known Families, Genera, and Species in South America*. Springer.

Heckman, C. W., 2008. *Encyclopedia of South American Aquatic Insects: Odonata-Zygoptera: Illustrated Keys to Known Families, Genera, and Species in South America*. Springer.

Heckman, C. W., 2011. *Encyclopedia of South American Aquatic Insects: Hemiptera-Heteroptera: Illustrated Keys to Known Families, Genera, and Species in South America*. Springer.

Heltshe, J. F. and Forrester, N. E., 1983. Estimating species richness using the jackknife procedure. *Biometrics*, 1-11.

Jones, K. E., Patel, N. G., Levy, M. A., Storeygard, A., Balk, D., Gittleman, J. L. and Daszak, P., 2008. Global trends in emerging infectious diseases. *Nature*, 451 (7181), 990-993.

Jung, M., 2013. *LecoS-A QGIS plugin for automated landscape ecology analysis*. PeerJ PrePrints. 2167-9843 [Accessed 01/05/2014].

Keesing, F., Belden, L. K., Daszak, P., Dobson, A., Harvell, C. D., Holt, R. D., Hudson, P., Jolles, A., Jones, K. E., Mitchell, C. E., Myers, S. S., Bogich, T. and Ostfeld, R. S., 2010. Impacts of biodiversity on the emergence and transmission of infectious diseases. *Nature*, 468 (7324), 647-652.

Keith, P., Le Bail, P.-Y. and Planquette, P., 2000a. *Atlas des poissons d'eau douce de Guyane*. Publications scientifiques du MNHN.

Keith, P., Le Bail, P. Y. and Planquette, P., 2000b. Atlas des poissons d'eau douce de Guyane(Tome 2- Fascicule I, Batrachoidiformes, Mugiliformes,

Beloniformes, Cyprinodontiformes, Synbranchiformes, Perciformes, Pleuronectiformes, Tetraodontiformes). *Collection patrimoines naturels*.

Landier, J., Gaudart, J., Carolan, K., Seen, D. L., Guégan, J.-F., Eyangoh, S., Fontanet, A. and Texier, G., 2014. Spatio-temporal Patterns and Landscape-Associated Risk of Buruli Ulcer in Akonolinga, Cameroon. *PLoS Neglected Tropical Diseases*, 8 (9), e3123.

Lau, C. L., Smythe, L. D., Craig, S. B. and Weinstein, P., 2010. Climate change, flooding, urbanisation and leptospirosis: fuelling the fire? *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 104 (10), 631-638.

Layman, C. A., Arrington, D. A., Montaña, C. G. and Post, D. M., 2007. Can stable isotope ratios provide for community-wide measures of trophic structure? *Ecology*, 88 (1), 42-48.

Lee, K. A., Wikelski, M., Robinson, W. D., Robinson, T. R. and Klasing, K. C., 2008. Constitutive immune defences correlate with life-history variables in tropical birds. *Journal of Animal Ecology*, 77 (2), 356-363.

Marion, E., Eyangoh, S., Yeramian, E., Doannio, J., Landier, J., Aubry, J., Fontanet, A., Rogier, C., Cassisa, V., Cottin, J., Marot, A., Eveillard, M., Kamdem, Y., Legras, P., Deshayes, C., Saint-André, J.-P. and Marsollier, L., 2010. Seasonal and Regional Dynamics of *M. ulcerans Tr*ansmission in Environmental Context: Deciphering the Role of Water Bugs as Hosts and Vectors. *PLoS Neglected Tropical Diseases*, 4 (7), e731.

Marsollier, L., Robert, R., Aubry, J., Saint Andre, J.-P., Kouakou, H., Legras, P., Manceau, A.-L., Mahaza, C. and Carbonnelle, B., 2002. Aquatic Insects as a Vector for *Mycobacterium ulcerans*. *Applied and Environmental Microbiology*, 68 (9), 4623-4628.

Marsollier, L., Stinear, T., Aubry, J., Saint André, J. P., Robert, R., Legras, P., Manceau, A.-L., Audrain, C., Bourdon, S., Kouakou, H. and Carbonnelle, B., 2004a. Aquatic Plants Stimulate the Growth of and Biofilm Formation

by *Mycobacterium ulcerans* in Axenic Culture and Harbor These Bacteria in the Environment. *Applied and Environmental Microbiology*, 70 (2), 1097-1103.

Marsollier, L., Sévérin, T., Aubry, J., Merritt, R. W., Saint André, J.-P., Legras, P., Manceau, A.-L., Chauty, A., Carbonnelle, B. and Cole, S. T., 2004b. Aquatic Snails, Passive Hosts of *Mycobacterium ulcerans*. *Applied and Environmental Microbiology*, 70 (10), 6296-6298.

Martin Ii, L., Hasselquist, D. and Wikelski, M., 2006. Investment in immune defense is linked to pace of life in house sparrows. *Oecologia*, 147 (4), 565-575.

Martin, L. B., Weil, Z. M. and Nelson, R. J., 2007. Immune defense and reproductive pace of life in peromyscus mice. *Ecology*, 88 (10), 2516-2528.

McKinney, M., 2008. Effects of urbanization on species richness: A review of plants and animals. *Urban Ecosystems*, 11 (2), 161-176.

McKinney, M. L., 2006. Urbanization as a major cause of biotic homogenization. *Biological Conservation*, 127 (3), 247-260.

Merritt, R. W., Benbow, M. E. and Small, P. L. C., 2005. Unraveling an Emerging Disease Associated with Disturbed Aquatic Environments: The Case of Buruli Ulcer. *Frontiers in Ecology and the Environment*, 3 (6), 323-331.

Merritt, R. W. and Cummins, K. W., 1996. *An introduction to the aquatic insects of North America*. Kendall Hunt.

Merritt, R. W., Walker, E. D., Small, P. L. C., Wallace, J. R., Johnson, P. D. R., Benbow, M. E. and Boakye, D. A., 2010. Ecology and Transmission of Buruli Ulcer Disease: A Systematic Review. *PLoS Neglected Tropical Diseases*, 4 (12), e911.

Meybeck, M., 2003. Global analysis of river systems: from Earth system controls to Anthropocene syndromes. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 358 (1440), 1935-1955.

Miserendino, M. L., 2001. Length-mass relationships for macroinvertebrates in freshwater environments of Patagonia (Argentina). *Ecología Austral*, 11, 3-8.

Morris, A., Gozlan, R., Marion, E., Marsollier, L., Andreou, D., Sanhueza, D., Ruffine, R., Couppié, P. and Guégan, J.-F., 2014a. First Detection of *Mycobacterium ulcerans* DNA in Environmental Samples from South America. *PLOS Neglected Tropical Diseases*, 8 (1), e2660.

Morris, A., Gozlan, R. E., Hassani, H., Andreou, D., Couppié, P. and Guégan, J.-F., 2014b. Complex temporal climate signals drive the emergence of human water-borne disease. *Emerging Microbes and Infections*, 3 (8), e56.

Mosi, L., Williamson, H., Wallace, J. R., Merritt, R. W. and Small, P. L. C., 2008. Persistent Association of *Mycobacterium ulcerans* with West African Predaceous Insects of the Family Belostomatidae. *Applied and Environmental Microbiology*, 74 (22), 7036-7042.

Nunn, C. L., Gittleman, J. L. and Antonovics, J., 2003. A comparative study of white blood cell counts and disease risk in carnivores. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270 (1513), 347-356.

Ostfeld, R. S., 2009. Biodiversity loss and the rise of zoonotic pathogens. *Clinical Microbiology and Infection*, 15 (s1), 40-43.

Ostfeld, R. S. and Keesing, F., 2000. Biodiversity series: the function of biodiversity in the ecology of vector-borne zoonotic diseases. *Canadian Journal of Zoology*, 78 (12), 2061-2078.

Petchey, O. L., Beckerman, A. P., Riede, J. O. and Warren, P. H., 2008. Size, foraging, and food web structure. *Proceedings of the National Academy of Sciences*, 105 (11), 4191-4196.

Rahel, F. J., 2002. Homogenization of freshwater faunas. *Annual Review of Ecology and Systematics*, 291-315.

Randolph, S. E. and Dobson, A. D. M., 2012. Pangloss revisited: a critique of the dilution effect and the biodiversity-buffers-disease paradigm. *Parasitology*, 139 (07), 847-863.

Roche, B., Benbow, M. E., Merritt, R., Kimbirauskas, R., McIntosh, M., Small, P. L. C., Williamson, H. and Guégan, J.-F., 2013. Identifying the Achilles heel of multi-host pathogens: the concept of keystone 'host'species illustrated by *Mycobacterium ulcerans* transmission. *Environmental Research Letters*, 8 (4), 045009.

Roche, B. and Guégan, J.-F., 2011. Ecosystem dynamics, biological diversity and emerging infectious diseases. *Comptes Rendus Biologies*, 334 (5–6), 385-392.

Roxburgh, S. H., Shea, K. and Wilson, J. B., 2004. The intermediate disturbance hypothesis: patch dynamics and mechanisms of species coexistence. *Ecology*, 85 (2), 359-371.

Smock, L. A., 1980. Relationships between body size and biomass of aquatic insects. *Freshwater biology*, 10 (4), 375-383.

Stinear, T., Ross, B. C., Davies, J. K., Marino, L., Robins-Browne, R. M., Oppedisano, F., Sievers, A. and Johnson, P. D. R., 1999. Identification and Characterization of IS2404 and IS2606: Two Distinct Repeated Sequences for Detection of *Mycobacterium ulcerans* by PCR. *Journal of Clinical Microbiology*, 37 (4), 1018-1023.

Tweedie, M. C. K., 1984, 1984. *An index which distinguishes between some important exponential families*.

van Ravensway, J., Benbow, M. E., Tsonis, A. A., Pierce, S. J., Campbell, L. P., Fyfe, J. A. M., Hayman, J. A., Johnson, P. D. R., Wallace, J. R. and Qi, J., 2012. Climate and Landscape Factors Associated with Buruli Ulcer Incidence in Victoria, Australia. *PLoS ONE*, 7 (12), e51074.

Vorosmarty, C. J., McIntyre, P. B., Gessner, M. O., Dudgeon, D., Prusevich, A., Green, P., Glidden, S., Bunn, S. E., Sullivan, C. A., Liermann, C. R. and Davies, P. M., 2010. Global threats to human water security and river biodiversity. *Nature*, 467 (7315), 555-561.

Wagner, T., Benbow, M. E., Brenden, T. O., Qi, J. and Johnson, R. C., 2008. Buruli ulcer disease prevalence in Benin, West Africa: associations with land use/cover and the identification of disease clusters. *Int J Health Geogr*, 7, 25.

World Health Organisation, 2001a. *Buruli Ulcer. Diagnosis of Mycobacterium ulcerans disease. A manual for health care providers* [online]. Geneva: Available from: http://www.who.int/iris/handle/10665/67000#sthash.cioj71jV.dpuf.

World Health Organisation 2001b. *Buruli Ulcer Management of Mycobacterium ulcerans disease* [online]. Available from: http://whqlibdoc.who.int/hq/2001/WHO_CDS_CPE_GBUI_2001.3.pdf

Chapter 7 Thesis discussion

7.1 Overview

EID's are difficult to study in the wild. Often their environmental sources are unknown or poorly understood and their transmission pathways hard to quantify. The necessity to characterise these above points however is central in predicting and preventing future outbreaks of serious consequence (Plowright et al. 2008). During this PhD, significant further light has been shed on several aspects of the EID Buruli ulcer, the results of which can be applied to other microbes which inhabit and propagate within similar environments.

Despite historical records of the disease in South America (Guerra et al. 2008; Merritt et al. 2010), for the first time, environmental sources of the bacterium have been identified. Both chapter 3 and chapter 6 provide further evidence of the aquatic nature of MU, with high numbers of the bacilli in or on numerous aquatic organisms and the presence of the disease in the water column and in biofilms. Environmental samples however yielded a much lower level of MU abundance. Thus, it is likely that the bacteria more commonly reside within or on a host organism. This would fit with the current understanding of MU and the reductive evolution it has recently undertaken (Stinear et al. 2007; Demangel et al. 2009). Such a difference between the number of positive environmental samples and those from aquatic macro-organisms has implications for our understanding of transmission and adds further evidence to the idea of key animal associations being involved, either as hosts or vectors. A recent paper has further suggested the necessity for an inoculation with the bacterium to occur before MU becomes infectious (Williamson et al. 2014). Whilst our results do not single out any one taxon capable of vectoring MU, numerous biting invertebrates were found to be positive. As previously postulated, it may be that as a generalist the bacteria are

capable of forming colonies within or on numerous biting taxa with no single route of transmission. This however, does not explain the lack of insect bites reported before development of BU. It may be argued that the long incubation period before symptoms occur (Trubiano et al. 2013) may be an explanation, or it may be that the vectors are not leaving serious and highly noticeable bites marks, thus eliminating the majority of large water-bugs as prime candidates as they leave lasting painful lesions. If a vector exists, it is perhaps more likely that less memorable bite's such as those from biting midges, commonly found to be positive for MU in chapter 6 are a significant source of inoculation.

The decline in cases around the Sinnamary area also coincides with the building of the Petite-Saut dam, which controls the main flow of water downstream. This may have a number of effects on the spread of the pathogen both from upstream and in areas it is already established by causing a general decrease in water levels, leading to a much drier environment where the bacteria is less able to proliferate, or by limiting downstream flooding. This relationship with water-flow links well with chapter 4, were it was found that BU cases are affected by complex temporal climatic changes. This association is a significant advance in the knowledge surrounding BU, when this PhD started such relationships were mostly speculative and how rainfall and water-levels related to cases was uncertain. The complexity of the patterns identified, with seasonal effects having a higher impact during certain 'drought' periods of long-term trends allowing for some stochastic rainfall events, gives a much clearer indication of the ecological niche of the bacilli. In addition it highlights the importance of using long-term datasets when linking disease outbreaks to climatic patterns and looking at different signals within a time series.

The association MU has with functionally similar basal groups of highly fecund invertebrates in aquatic systems is demonstrated in both chapters 5 and 6. How this relates to chapters 3 and 4 allows for the first time a much more detailed picture of

the ecological life history of the bacteria. For example, what may be causing the proliferation of MU and thus potentially BU, could start with anthropogenic environmental disturbance that would lead to a significant increase in basal species. Along this increase in the number of MU preferred hosts, the bacteria become more easily transmissible, either via a range of vectors or through an increase in overall abundance of bacteria in the ecosystem. In turn, disturbed aquatic habitats are more susceptible to climatic fluctuations. By being more exposed to the elements through agricultural and urban land cover and/or deforestation they will stagnate faster and flood more readily. This is compounded by the fact that many predators which would feed on MU hosts will also be more vulnerable. Such managed or impacted sites are also more likely to be in areas where people are more often accessing these aquatic environments (e.g. swimming, hunting and fishing). This combination of factors creates a hotspot for BU and is a significant advance in our knowledge of how the bacteria are transmitted. As previously surmised by Plowright et al. (2008) it also shows the importance of a multi-faceted approach to disease emergence.

7.2 Relating to other EID's

Whilst no two pathogens are identical and within emergent diseases this often becomes more apparent, certain characteristics, predominantly a lack of host specificity and variable transmission routes are common. As such the results of this study which takes a holistic view of ecological and environmental drivers of disease presence can be applied to other EID's. In particular the pattern of land-use change and human disturbance leading to a change in biotic communities is one which is likely to be applicable to other aetiological agents. In chapter 6 there is empirical evidence that anthropogenic disturbance shifts community composition towards an aquatic community more suited to the proliferation of the infectious agent, or where a disease is able to proliferate more readily. The pattern is a

quadratic response, with a decline in disease carrying taxa at very high levels of disturbance, suggesting the biggest risk to be on the boundaries of highly bio-diverse areas such as forests or in areas of intermediate disturbance. This fits well with many emerging diseases, which often are first reported in boundary towns and in areas where urbanisation has encroached into the forest.

Similarly patterns in climate have been applied to numerous diseases in the past, but often EID's are omitted from these studies. This oversight may occur because it is harder to understand the underlying causation of disease increases when the effects of climatic variability on the pathways of emergence are not clearly established. As shown here however this information can be vital in helping to build a more accurate picture of the ecological niche of a pathogen. In a more immediate sense, if climatic relationships are found it can also help target resources such as increased surveillance based on predictable climatic events such as el-niño.

7.3 Potential for future disease in French Guiana

The results of the thesis identify key sources of the disease agent within French Guiana, and suggest the bacteria is relatively common, it was found in the majority of sites where biotic surveys were conducted. The difficulty is in relating this information to human risk, there is no clear evidence of how much bacteria, and how many infected hosts need to be present before risk to humans increase. As our knowledge remains limited it is possible that a large presence of the bacteria within the tropics is the norm, and it is other factors which lead to either inoculation or a significant jump in bacterial concentration, such as with cholera and anthrax (Blackburn et al. 2007; Colwell and Spira 1992). With these diseases their presence in soil and water-bodies is fairly ubiquitous, and they can be found in relatively low levels under the right conditions across the globe. However, there are certain environmental and ecological changes which dramatically increase their numbers

within a given area; this coupled with human contact is what leads to outbreaks. Such a relationship may also be the case with MU, its presence may be fairly uniform across tropical areas, as shown in this and other studies (Benbow et al. 2013; Garchitorena et al. 2014; Williamson et al. 2008) but it is a combination of factors as previously discussed, which cause disease hotspots to appear.

French Guiana is predicted by the analysis in chapter 4 to be entering a period of drier years, during which BU is more likely to occur. Currently ENSO is also in an el-niño year which was implicated in a decarease in extreme weather events that trigger further stochastic outbreaks. However, the el-niño oscillation is so far fairly mild and has been fluctuating between both la-niña and el-niño, therefore it may not be causing any significant change in weather patterns. Despite this, the main long-term weather driver which is predicted to become more suitable for BU, suggests a significant increase above average of cases over the coming few years.

The extent of ecological change in this region is harder to quantify for the future, recently gold-mining has caused a lot of deforestation, both regulated and illegal (Hammond et al. 2007). Whilst agricultural land is not expanding at any significant rate, and large areas have been abandoned, such as the rice fields outside Mana due to coastal erosion, human population in French Guiana continues to increase rapidly (figure 1.2) and with it pressures on the environment, particularly on the forest boundaries. With the results shown in chapter 6 this suggests BU is likely to remain a persistent threat. Eradication of the bacteria is near impossible, owing to its ubiquitous nature, and where there is ecological pressure the environments shown in this thesis to be hotspots for the bacilli and its hosts are almost unavoidable, without significant resources and control over developments.

7.4 Future work

The results here still do not answer the specific question of how the disease is transmitted from the environment into human hosts. Unless this pathway is through contaminated water, further study of individual taxa highlighted in chapters 5 and 6 needs to be undertaken. For example, numerous biting midge larvae (Ceratopogonidae) were found to carry high levels of the bacilli and have been so far understudied as a potential vector, despite similar transmission routes for other diseases. Several studies still maintain the necessity for inoculation with the bacteria, rather than contact for it to become infective (Marion et al. 2014; Williamson et al. 2014), so all potential biting vectors should be explored. This group is highly abundant and analysis of adults within would quickly identify whether the bacteria are able to survive through metamorphosis.

Further profiling of MU ecovars would help to understand variation within sites, whether certain taxa carry differing species and which of these is most likely to carry human infecting forms. Whilst there is some debate as to how different the ecovars are (Pidot et al. 2010), and therefore their ability to manifest as disease, more work needs to be done on understanding the environmental pressures which give rise to them. Currently there is no direct evidence that they either live side by side or inhabit significantly different niches. There is some evidence for the ecovar *Mycobacterium ulcerans liflandii* to be predominantly an infection of amphibians (Tobias et al. 2013), there is no confirmation of whether it is still capable of infecting humans, and what specific traits have manifested from the genetic changes it has undergone. Direct ecological and environmental profiling of sites where these ecovars are found would go some way to explaining what causes their proliferation, along with genetic profiling of MU DNA from other hosts. For example, do these ecovars inhabit the same range of invertebrate and vertebrate hosts but only manifest as disease in different hosts? Or is there cross-over in their

persistence within sites and taxa. Whilst genetic profiling has enabled some separation of these closely related Mycobacteria, it is difficult to treat them separately in terms of risk to humans until comparisons of their life history can be made. This is important as there may be discrepancies between how each are transmitted, are some more virulent in those with different immune profiles or more likely to be passed on through differing human behaviour.

One key area of research into BU, which has largely been neglected, is the social economic and behavioural aspects of those infected. Whilst there is numerous anecdotal evidence one of the first large surveys of patients was only conducted in 2013 (Giles-Vernick et al. 2014). It suggests key moments within the past few decades where disease increased, including; changes in agricultural practices, ecological degradation leading to nutritional deficiencies and environmental events such as floods. Further work needs to be done in this area, particularly in behaviour, perhaps on a shorter local scale looking at events in individuals before contracting the disease. The difficulties in such studies are however obvious, with the long incubation, period people need to preferentially be logging events before contracting the infection. Such a study whilst not impossible is highly impractical and would require a large amount of time and resources.

7.5 Conclusions

During this PhD significant advances have been made in understanding the ecology of BU (Table 7.1). When the work was started, the ecological niche and its relationship with the environment and climate were poorly understood. The goal was to identify environmental sources of MU and predictors of its emergence, by combing the results from each chapter this has been achieved. This thesis provides empirical evidence for all the above points and goes someway in helping to fully understand the life history of the bacilli and the conditions that enable it to

proliferate. It is hoped that the results of this study can also be applied to other similar generalist emerging bacterial diseases, helping to predict and prevent future outbreaks. In particular it shows mechanisms by which infectious hosts can be encouraged through anthropogenic disturbance, which simultaneously increases human contact. Whilst there is still a lot of work to be done to fully understand this devastating infection, this thesis adds significant pieces to the puzzle and helps pave the way for future analysis.

Table 7.1 How knowledge of MU ecology has been furthered by the results of this PhD.

MU ecology	Before PhD	Results from PhD	PhD Reference
Environmental sources within French Guiana/ Southern America	Unknown, with no direct evidence of any environmental source for South American strain of MU	Key environmental sources of MU identified in French Guiana	Chapters 3 and 6
Relationship between climate and disease	One paper specifically linking recent outbreaks to climatic factors on a short temporal scale.	In depth, long term analysis of rainfall and BU cases over more than 50 years identifying specific patterns of outbreaks which are consistent in two continents.	Chapter 4
Identifying biotic hosts of MU	Previous papers have found the bacteria in numerous organisms but with no empirical relationship between them, other than being found within the same freshwater systems	Identification of specific functional groupings of taxa which are found in areas of disease presence and which have been found to carry high concentrations of the bacteria.	Chapters 5 and 6
Identifying how land use can affect the disease presence	No direct evidence	Identification of a relationship between land use and community structure and pinpointing how this can lead to an increase in disease carrying species.	Chapter 6

7.6 References

Benbow, M. E., Kimbirauskas, R., McIntosh, M. D., Williamson, H., Quaye, C., Boakye, D., Small, P. L. C. and Merritt, R. W., 2013. Aquatic Macroinvertebrate Assemblages of Ghana, West Africa: Understanding the Ecology of a Neglected Tropical Disease. *EcoHealth*, 1-16.

Blackburn, J. K., McNyset, K. M., Curtis, A. and Hugh-Jones, M. E., 2007. Modeling the geographic distribution of Bacillus anthracis, the causative agent of anthrax disease, for the contiguous United States using predictive ecologic niche modeling. *The American Journal of Tropical Medicine and Hygiene*, 77 (6), 1103-1110.

Colwell, R. R. and Spira, W. M., 1992. The ecology of Vibrio cholerae. *In: Cholera.* Springer, 107-127.

Demangel, C., Stinear, T. P. and Cole, S. T., 2009. Buruli ulcer: reductive evolution enhances pathogenicity of *Mycobacterium ulcerans*. *Nature Reviews Microbiology*, 7 (1), 50-60.

Garchitorena, A., Roche, B., Kamgang, R., Ossomba, J., Babonneau, J., Landier, J., Fontanet, A., Flahault, A., Eyangoh, S. and Guégan, J.-F., 2014. *Mycobacterium ulcerans* Ecological Dynamics and Its Association with Freshwater Ecosystems and Aquatic Communities: Results from a 12-Month Environmental Survey in Cameroon. *PLoS Neglected Tropical Diseases*, 8 (5), e2879.

Giles-Vernick, T., Owona-Ntsama, J., Landier, J. and Eyangoh, S., 2014. The puzzle of Buruli ulcer transmission, ethno-ecological history and the end of "love" in the Akonolinga district, Cameroon. *Social Science & Medicine*.

Guerra, H., Palomino, J. C., Falconí, E., Bravo, F., Donaires, N., Van Marck, E. and Portaels, F., 2008. *Mycobacterium ulcerans* disease, Peru. *Emerging Infectious Diseases*, 14 (3), 373.

Hammond, D. S., Gond, V., Thoisy, B. d., Forget, P.M. and DeDijn, B. P. E., 2007. Causes and consequences of a tropical forest gold rush in the Guiana Shield, South America. *AMBIO: A Journal of the Human Environment*, 36 (8), 661-670.

Marion, E., Chauty, A., Yeramian, E., Babonneau, J., Kempf, M. and Marsollier, L., 2014. A case of guilt by association: Water bug bite incriminated in *M. ulcerans* infection. *International Journal of Mycobacteriology*.

Merritt, R. W., Walker, E. D., Small, P. L. C., Wallace, J. R., Johnson, P. D. R., Benbow, M. E. and Boakye, D. A., 2010. Ecology and Transmission of Buruli Ulcer Disease: A Systematic Review. *PLoS Neglected Tropical Diseases*, 4 (12), e911.

Pidot, S. J., Asiedu, K., Käser, M., Fyfe, J. A. M. and Stinear, T. P., 2010. *Mycobacterium ulcerans* and other mycolactone-producing mycobacteria should be considered a single species. *PLoS Neglected Tropical Diseases*, 4 (7), e663.

Plowright, R. K., Sokolow, S. H., Gorman, M. E., Daszak, P. and Foley, J. E., 2008. Causal inference in disease ecology: investigating ecological drivers of disease emergence. *Frontiers in Ecology and the Environment*, 6 (8), 420-429.

Stinear, T. P., Seemann, T., Pidot, S., Frigui, W., Reysset, G., Garnier, T., Meurice, G., Simon, D., Bouchier, C., Ma, L., Tichit, M., Porter, J. L., Ryan, J., Johnson, P. D. R., Davies, J. K., Jenkin, G. A., Small, P. L. C., Jones, L. M., Tekaia, F., Laval, F., Daffé, M., Parkhill, J. and Cole, S. T., 2007. Reductive evolution and niche adaptation inferred from the genome of *Mycobacterium ulcerans*, the causative agent of Buruli ulcer. *Genome Research*, 17 (2), 192-200.

Tobias, N. J., Doig, K. D., Medema, M. H., Chen, H., Haring, V., Moore, R., Seemann, T. and Stinear, T. P., 2013. Complete genome sequence of the

frog pathogen *Mycobacterium ulcerans* ecovar Liflandii. *Journal of Bacteriology*, 195 (3), 556-564.

Trubiano, J. A., Lavender, C. J., Fyfe, J. A. M., Bittmann, S. and Johnson, P. D. R., 2013. The incubation period of Buruli ulcer (*Mycobacterium ulcerans* infection). *PLoS Neglected Tropical Diseases*, 7 (10), e2463.

Williamson, H. R., Mosi, L., Donnell, R., Aqqad, M., Merritt, R. W. and Small, P. L. C., 2014. *Mycobacterium ulcerans* Fails to Infect through Skin Abrasions in a Guinea Pig Infection Model: Implications for Transmission. *PLoS Neglected Tropical Diseases*, 8 (4), e2770.

Williamson, H. R., Benbow, M. E., Nguyen, K. D., Beachboard, D. C., Kimbirauskas, R. K., McIntosh, M. D., Quaye, C., Ampadu, E. O., Boakye, D., Merritt, R. W. and Small, P. L. C., 2008. Distribution of *Mycobacterium ulcerans* in Buruli Ulcer Endemic and Non-Endemic Aquatic Sites in Ghana. *PLoS Neglected Tropical Diseases*, 2 (3), e205.

Williamson, H. R., Mosi, L., Donnell, R., Aqqad, M., Merritt, R. W. and Small, P. L. C., 2014. *Mycobacterium ulcerans* Fails to Infect through Skin Abrasions in a Guinea Pig Infection Model: Implications for Transmission. *PLoS Neglected Tropical Diseases*, 8 (4), e2770.

Appendices

Appendix I

Analysis Site Photos

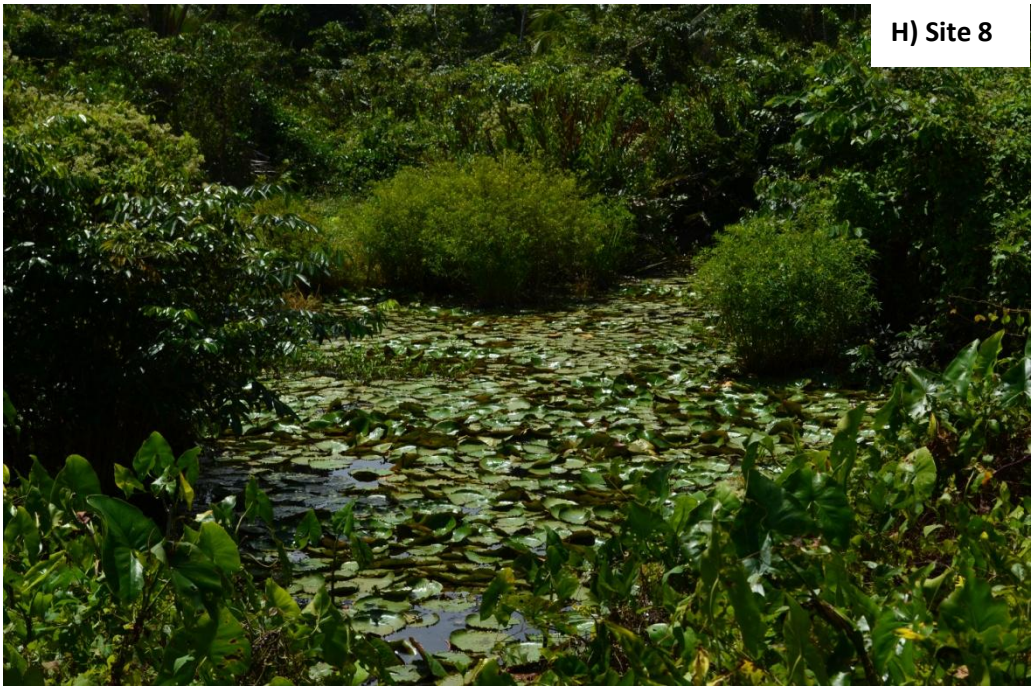
C) Site 3

D) Site 4

E) Site 5

F) Site 6

G) Site 7



H) Site 8

I) Site 9

J) Site 10

K) Site 11

L) Site 12

M) Site 13

N) Site 14

O) Site 15

P) Site 16

Q) Site 17

Appendix II

Figure S5.1: K-Means partitions comparison using Simple Structure Index (SSI) [1] shows a likely optimal number of functional groups to be 11. The values of the SSI criterion (right) are based on the maximum difference of each variable between the clusters of taxons on the dendrogram (Figure 5.1), the sizes of the most contrasting clusters and the deviation of a variable in the cluster centres compared to its overall mean.

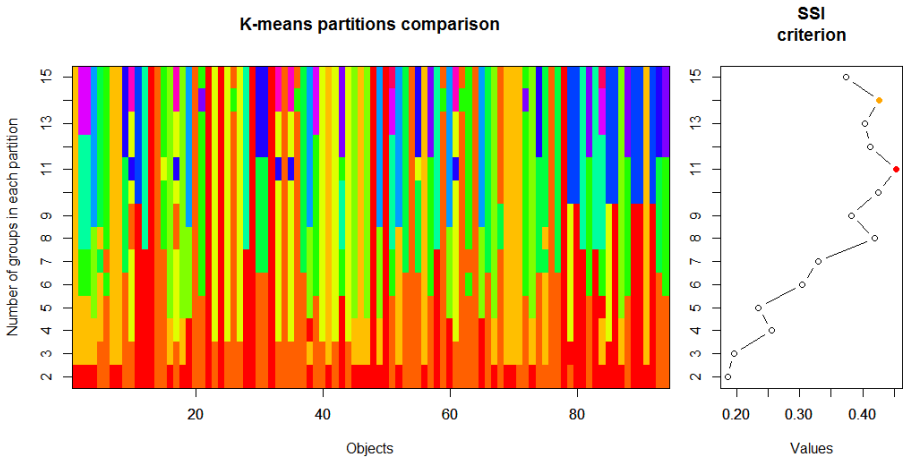


Table S5.1 Key for motility trait in table S5.2:

Abbreviation	Trait
EpB	Epibenthic burrower
SA	Semi-aquatic
SwS	Surface water swimmer
FwS	Free water swimmer
CL	Crawler
AT	Predominantly attached to a surface

Table S5.2 Definitions of feeding groups

Abbreviation	Trait
Scraper	Consumes biofilms and algae directly from a surface substrate
Predator	Hunts and feeds upon other invertebrates and/or vertebrates
Shredder	Chews litter, live or dead vascular plant tissue, or gauge wood
Gathering Collector	Passively gather loose particles of detritus and decaying tissue in addition to algae, bacteria and plankton
Filtering Collector	Passively gather loose particles of detritus and decaying tissue in addition to algae, bacteria and plankton
Predators/Scavengers	Can both hunt and feed on living organisms or will feed upon carcasses of dead organisms
Haematophagous	Sucks blood from other animals

Table S5.3 Definitions of food types

Abbreviation	Trait
Phytophagous	The tissue or fluids of living macro plants
Microphytes	Algae and plankton
Detritus/Microphytes	Decaying tissue of macroplants in addition to microphytes such as algae and plankton
Microinvertebrates	Small invertebrates, generally <1mm
Macroinvertabrates	Larger invertebrates, generally >1mm. Can also include small macrovertebrates, e.g juvenile fish or amphibians
Omnivorous	Used for taxons which do not feed on an easily defined food type and will consume most foods available, including through predation. Covers several taxons which have a wide range of feeding habits within the taxonomic class used
Microorganisms	Any microscopic organism, including: algae, bacteria, microinvertebrates and plankton
Parasitic	Feeds on the tissue or fluids of a living host
Decaying tissue	Decaying tissue of other animals

Table S5.4 Table of species listed with their corresponding functional group and functional traits.

Taxon	Group	Feeding group	Food type	Respiration	Body shape	Motility	Main body constituents
Ancylidae	1	Scraper	Microphytes	Lung	Spherical	AT	Calcium carbonate
Baetidae	1	Scraper	Microphytes	Gill	Flattened	CL	Chitin/sclerotin
Bithyniidae	1	Scraper	Microphytes	Gill	Cylindrical	CL	Calcium carbonate
Heptageniidae	1	Scraper	Microphytes	Gill	Flattened	AT	Chitin/sclerotin
Hydrobiidae	1	Scraper	Microphytes	Gill	Cylindrical	EpB	Calcium carbonate
Hydroptilidae	1	Scraper	Microphytes	Gill	Cylindrical	CL	Constructed Casing
Leptophlebiidae	1	Scraper	Detritus/Microphytes	Gill	Flattened	AT	Chitin/sclerotin
Lymnaeidae	1	Scraper	Microphytes	Lung	Cylindrical	CL	Calcium carbonate
Psephenidae	1	Scraper	Microphytes	Gill	Flattened	AT	Chitin/sclerotin
Thiaridae	1	Scraper	Microphytes	Lung	Cylindrical	CL	Calcium carbonate
Araneae	2	Predators	Macroorganisms	Lung	Spherical	CL	Chitin
Hydrozoida	2	Predator	Macroorganisms	Tegument	NA	NA	NA
Belostomatidae	2	Predators	Macroorganisms	Plastron	Flattened	FwS	Chitin/sclerotin
Chlorocyphidae	2	Predators	Macroorganisms	Gill	Flattened	CL	Chitin/sclerotin
Dytiscidae	2	Predators	Macroorganisms	Plastron	Flattened	CL	Chitin/sclerotin
Nepidae	2	Predators	Macroorganisms	Stigmata	Cylindrical	CL	Chitin/sclerotin
Notonectidae	2	Predators	Macroorganisms	Plastron	Flattened	FwS	Chitin/sclerotin
Odonata	2	Predators	Macroorganisms	Gill	Flattened	CL	Chitin/sclerotin
Athericidae	3	Predators	Microinvertebrates	Gill	Spherical	CL	Chitin
Chaoboridae	3	Predators	Microinvertebrates	Tegument	Cylindrical	FwS	Chitin
Empididae	3	Predators	Microinvertebrates	Gill	Cylindrical	EpB	Chitin
Hebridae	3	Predators	Microinvertebrates	Stigmata	Flattened	CL	Chitin/sclerotin
Hydracarinidae	3	Predators	Microinvertebrates	Tegument	Flattened	CL	Membranous
Noteridae	3	Predators	Microinvertebrates	Stigmata	Flattened	CL	Chitin/sclerotin

Perlidae	3	Predators	macroinvertebrates	Gill	Flattened	EpB	Chitin/sclerotin
Pleidae	3	Predators	Microinvertebrates	Plastron	Spherical	FwS	Chitin/sclerotin
Staphylinidae	3	Predators	macroinvertebrates	Stigmata	Flattened	EpB	Chitin/sclerotin
Atyidae	4	Gathering collectors	Omnivourus	Gill	Cylindrical	FwS	Calcium carbonate
Ceratopogonidae	4	Gathering collectors	Omnivourus	Gill	Cylindrical	CL	Chitin
Chironomidae	4	Gathering collectors	Omnivourus	Tegument	Cylindrical	EpB	Chitin
Culicidae	4	Gathering collectors	Omnivourus	Stigmata	Cylindrical	FwS	Chitin
Dixidae	4	Gathering collectors	Microphytes	Stigmata	Cylindrical	Sws	Chitin
Hydraenidae	4	Gathering collectors	Microphytes	Plastron	Cylindrical	CL	Chitin/sclerotin
Isopoda	4	Various	Omnivourus	Gill	NA	FwS	Chitin
Leptoceridae	4	Various	Omnivourus	Gill	Cylindrical	CL	Constructed Casing
Potamonautidae	4	Shredders	Omnivourus	Gill	Flattened	CL	Calcium carbonate
Simuliidae	4	Gathering collectors	Omnivourus	Stigmata	Cylindrical	CL	Chitin/sclerotin
Tipulidae	4	Gathering collectors	Omnivourus	Stigmata	Cylindrical	CL	Chitin
Bulininae	5	Scraper	Phytophagous	Lung	Cylindrical	CL	Calcium carbonate
Physidae	5	Scraper	Phytophagous	Lung	Cylindrical	CL	Calcium carbonate
Pilidae	5	Scraper	Phytophagous	Gill	Spherical	CL	Calcium carbonate
Planorbidae	5	Scraper	Phytophagous	Lung	Flattened	CL	Calcium carbonate
Caenidae	6	Gathering collectors	Detritus/Microphytes	Gill	Flattened	CL	Chitin/sclerotin
Elmidae	6	Gathering collectors	Detritus/Microphytes	Plastron	Cylindrical	CL	Chitin/sclerotin
Hydrophilidae	6	Predators/Scavengers	Detritus/Microphytes	Plastron	Cylindrical	CL	Chitin/sclerotin
Limnichidae	6	Gathering collectors	Detritus/Microphytes	Plastron	Flattened	CL	Chitin/sclerotin
Ptilodactylidae	6	Gathering collectors	Detritus/Microphytes	Stigmata	Cylindrical	CL	Chitin/sclerotin
Sminthuridae	6	Gathering collectors	Detritus/Microphytes	tegument	Cylindrical	CL	Chitin
Stratiomyiidae	6	Gathering collectors	Detritus/Microphytes	Stigmata	Flattened	CL	Chitin
Syrphidae	6	Gathering collectors	Detritus/Microphytes	Stigmata	Cylindrical	CL	Chitin
Calopterygidae	7	Predators	Macroinvertebrates	Gill	Cylindrical	CL	Chitin/sclerotin

Coenagrionidae	7	Predators	Macroinvertebrates	Gill	Cylindrical	CL	Chitin/sclerotin
Corduliidae	7	Predators	Macroinvertebrates	Gill	Flattened	CL	Chitin/sclerotin
Gerridae	7	Predators	Macroinvertebrates	Plastron	Cylindrical	SwS	Chitin/sclerotin
Gomphidae	7	Predators	Macroinvertebrates	Gill	Flattened	EpB	Chitin/sclerotin
Gyrinidae	7	Predators	Macroinvertebrates	Plastron	Cylindrical	CL	Chitin/sclerotin
Lampyridae	7	Predators	Macroinvertebrates	Stigmata	Flattened	SA	Chitin
Libellulidae	7	Predators	Macroinvertebrates	Gill	Flattened	CL	Chitin/sclerotin
Muscidae	7	Predators	Macroinvertebrates	Stigmata	Cylindrical	CL	Chitin
Naucoridae	7	Predators	Macroinvertebrates	Plastron	Flattened	CL	Chitin/sclerotin
Protoneuridae	7	Predators	Macroinvertebrates	Gill	Flattened	CL	Chitin/sclerotin
Rhagionidae	7	Predators	Macroinvertebrates	Stigmata	Cylindrical	EpB	Chitin
Sciomyzidae	7	Predators	Macroinvertebrates	Stigmata	Cylindrical	FwS	Chitin
Chrysomelidae	8	Shredders	Phytophagous	Plastron	Flattened	CL	Chitin/sclerotin
Corixidae	8	Shredders	Phytophagous	Plastron	Flattened	SwS	Chitin/sclerotin
Curculionidae	8	Shredders	Phytophagous	Plastron	Cylindrical	CL	Chitin/sclerotin
Dryopidae	8	Shredders	Phytophagous	Plastron	Cylindrical	CL	Chitin/sclerotin
Ephydridae	8	Shredders	Microphytes	Stigmata	Cylindrical	EpB	Chitin
Hirudinae	8	hematophagous	Parasitic	Tegument	Cylindrical	FwS	Membranous
Hydrochidae	8	Shredders	Phytophagous	Plastron	Cylindrical	CL	Chitin/sclerotin
Pyralidae	8	Shredders	Phytophagous	Stigmata	Cylindrical	CL	Chitin
Tricorythidae	8	Gathering collectors	Phytophagous	Gill	Flattened	CL	Chitin/sclerotin
Cladocera	9	Filtering collectors	Microorganisms	Tegument	Flattened	FwS	Membranous
Conchostraca	9	Filtering collectors	Microorganisms	Gill	Spherical	FwS	Calcium carbonate
Hydropsychidae	9	Filtering collectors	Microorganisms	Gill	Cylindrical	AT	Constructed Casing
Oligoneuriidae	9	Filtering collectors	Microorganisms	Gill	Flattened	CL	Chitin/sclerotin
Polycentropodidae	9	Filtering collectors	Microorganisms	Gill	Cylindrical	AT	Chitin
Scirtidae	9	Filtering collectors	Microorganisms	Stigmata	Flattened	CL	Chitin

Sphaeriidae	9	Filtering collectors	Microorganisms	Gill	Spherical	EpB	Calcium carbonate
Copepoda	10	Gathering collectors	Detritus/Microphytes	Tegument	Cylindrical	FwS	Chitin
Entomobryiidae	10	Gathering collectors	Detritus/Microphytes	Tegument	Cylindrical	EpB	Chitin
Ephemeridae	10	Gathering collectors	Detritus/Microphytes	Gill	Flattened	EpB	Chitin/sclerotin
Ephemerythidae	10	Gathering collectors	Detritus/Microphytes	Gill	Flattened	EpB	Chitin/sclerotin
Hypogastruridae	10	Gathering collectors	Detritus/Microphytes	Tegument	Cylindrical	EpB	Chitin
Isotomidae	10	Gathering collectors	Detritus/Microphytes	Tegument	Cylindrical	EpB	Chitin
Nematoda	10	Gathering collectors	Detritus/Microphytes	Tegument	Cylindrical	EpB	Membranous
Oligochaeta	10	Gathering collectors	Detritus/Microphytes	Stigmata	Cylindrical	EpB	Membranous
Ostracoda	10	Gathering collectors	Detritus/Microphytes	Gill	Spherical	EpB	Calcium carbonate
Polymitarcyidae	10	Gathering collectors	Detritus/Microphytes	Gill	Flattened	EpB	Chitin/sclerotin
Psychodidae	10	Gathering collectors	Detritus/Microphytes	Stigmata	Cylindrical	EpB	Chitin
Hydrometridae	11	Predators/Scavengers	Decaying tissue	Stigmata	Cylindrical	SA	Chitin/sclerotin
Mesoveliidae	11	Predators/Scavengers	Decaying tissue	Stigmata	Flattened	SwS	Chitin/sclerotin
Saldidae	11	Predators/Scavengers	Microinvertebrates	Stigmata	Flattened	SA	Chitin/sclerotin
Veliidae	11	Predators/Scavengers	Microinvertebrates	Stigmata	Flattened	FwS	Chitin/sclerotin

Table S5.5 The results of all mixed effects models of functional groups and taxonomic groups at lentic sites, against the number of samples positive for MU. P-values were estimated from the model using the pval.fnc function of the language R package with 10,000 Markov chain Monte Carlo samples.

Taxon/group	P value	Coefficient	AIC
Veliidae	0.48982	2.12867	-16.4909
Tipulidae	0.86449	0.73419	-16.0792
Thiaridae	0.61275	-0.53219	-16.3629
Stratiomyiidae	0.12405	-9.84885	-18.5857
Simuliidae	0.73703	-2.66960	-16.199
Scirtidae	0.81700	-0.51156	-16.114
Sciomyzidae	0.44700	28.49018	-16.6872
Pyralidae	0.28167	9.91966	-17.2513
Psychodidae	0.56389	-28.90581	-16.3814
Protoneuridae	0.23673	-1.00881	-17.4853
Polymitarcyidae	0.28194	-8.58477	-17.2411
Pleidae	0.16409	0.86532	-18.0572
Planorbidae	0.35711	0.38209	-16.9293
Pilidae	0.76052	-4.72126	-16.1634
Physidae	0.78387	-1.38156	-16.1336
Ostracoda	0.05479	0.35631	-19.9854
Oligochaeta	0.24457	0.68361	-17.486
Notonectidae	0.31423	-0.44708	-17.1398
Noteridae	0.61871	0.75027	-16.3429
Nepidae	0.43795	-11.31370	-16.7392
Naucoridae	0.46006	3.10554	-16.6093
Mesoveliidae	0.18144	-3.13976	-18.0132
Lymnaeidae	0.20661	-6.12455	-17.677
Libellulidae	0.77476	0.26271	-16.1383

Lampyridae	0.90894	-1.57569	-16.0726
Isotomidae	0.24930	23.98491	-17.5108
Hydrozoida	0.01029	-9.30011	-22.4995
Hydropsychidae	0.16597	-12.55946	-18.0422
Hydrophilidae	0.75286	0.32483	-16.1351
Hydrometridae	0.05108	-39.42746	-20.1503
Hydrochidae	0.41819	20.71936	-16.731
Hydrobiidae	0.59372	1.25174	-16.3503
Hydraenidae	0.62004	0.19604	-16.2682
Hydracarinidae	0.78950	-0.22969	-16.1161
Hirudinae	0.79669	0.90992	-16.1305
Heptageniidae	0.14545	1.49316	-18.3572
Gyrinidae	0.50749	-5.70560	-16.4423
Gerridae	0.24690	-2.42153	-17.5166
Entomobryiidae	0.56945	-1.85681	-16.3565
Elmidae	0.91908	-2.37171	-16.084
Dytiscidae	0.42654	-0.88204	-16.8235
Dixidae	0.72661	-4.28328	-16.2098
Curculionidae	0.30422	15.12612	-17.1458
Culicidae	0.72133	-0.13209	-16.1899
Corixidae	0.83630	0.87210	-16.1004
Corduliidae	0.74289	-0.85833	-16.1717
Copepoda	0.01980	0.67333	-21.5231
Conchostraca	0.69374	-0.36147	-16.2214
Coenagrionidae	0.65943	-0.63123	-16.2549
Cladocera	0.35403	0.38659	-17.016
Chrysomelidae	0.25875	6.51649	-17.2395
Chironomidae	0.01236	-0.39818	-22.2938

Chaoboridae	0.17442	6.79236	-18.1145
Ceratopogonidae	0.76431	-0.40865	-16.1275
Caenidae	0.23390	1.02013	-17.5425
Bulininae	0.33147	-0.49709	-17.057
Belostomatidae	0.32231	-3.28153	-17.0732
Baetidae	0.14984	-0.30878	-18.2902
Atyidae	0.57245	0.16501	-16.3841
Araneae	0.17136	-3.02399	-17.8723
Ancylidae	0.95221	0.09899	-16.0588
group11	0.39250	-1.56082	-16.913
group10	0.00381	0.39177	-24.812
group9	0.60984	0.17127	-16.3566
group8	0.12904	4.88053	-18.3169
group7	0.31440	-0.51832	-17.1154
group6	0.29205	0.78542	-17.1465
group5	0.94227	0.02151	-16.062
group4	0.07513	-0.24505	-19.2738
group3	0.32080	0.46926	-17.1204
group2	0.05924	-0.72702	-19.8482
group1	0.25558	-0.22921	-17.4667

Table S5.6 The results of all mixed effects models of functional groups and taxonomic groups at lotic sites, against the number of samples positive for MU. P-values were estimated from the model using the pval.fnc function of the language R package with 10,000 Markov chain Monte Carlo samples.

Taxon/group	P value	Coefficient	AIC
Veliidae	0.41578	-2.67155	2.58986
Tricorythidae	0.53926	-2.54708	2.83946
Tipulidae	0.61276	-6.96441	3.08669
Thiaridae	0.70153	-0.12136	3.08342
Stratiomyiidae	0.41307	-15.48138	2.54304
Sphaeriidae	0.47404	1.65785	2.71471
Simuliidae	0.73635	-0.38594	3.15175
Scirtidae	0.31143	3.35269	2.14376
Pyralidae	0.06890	-32.47219	0.06326
Psychodidae	0.61726	-2.26733	2.96195
Protoneuridae	0.78194	0.39179	3.20125
Potamonautidae	0.94752	0.64140	3.28817
Polycentropodidae	0.70575	-6.70252	3.13541
Pleidae	0.10780	-2.22863	0.87891
Planorbidae	0.82718	0.22915	3.24765
Pilidae	0.33373	9.28538	2.26783
Physidae	0.19188	-10.07399	1.49513
Perlidae	0.70131	-6.31619	3.08443
Ostracoda	0.63737	0.26601	3.03906
Oligochaeta	0.86408	-0.17896	3.26314
Odonata	0.03104	8.49505	-1.60674
Notonectidae	0.21147	2.29795	1.64530
Noteridae	0.93348	-0.18927	3.29009
Nepidae	0.11875	-17.42698	0.89656

Nematoda	0.21320	-15.90279	1.75724
Naucoridae	0.10447	-17.72515	0.76636
Mesoveliidae	0.86965	-0.88215	3.28073
Lymnaeidae	0.97463	-0.84656	3.29269
Libellulidae	0.19458	3.34599	1.42204
Leptophlebiidae	0.47634	0.89976	2.82792
Leptoceridae	0.91799	0.27425	3.27454
Lampyridae	0.24357	15.38257	1.82858
Isotomidae	0.00387	41.77032	-5.44155
Hydroptilidae	0.03081	-24.93223	-1.37375
Hydropsychidae	0.96457	0.26143	3.28884
Hydrophilidae	0.14264	-4.57737	1.22270
Hydraenidae	0.30254	0.98391	2.24715
Hydracarinidae	0.51379	-0.98796	2.89067
Hirudinae	0.61741	-3.65542	3.05933
Heptageniidae	0.73543	0.26086	3.16810
Hebridae	0.77694	-4.30275	3.19957
Gyrinidae	0.19343	-14.52173	1.53332
Gomphidae	0.87193	1.08527	3.28204
Gerridae	0.10364	5.78681	0.53774
Ephydridae	0.49559	14.14675	2.75513
Ephemerythidae	0.86648	0.48233	3.27732
Ephemeridae	0.86162	1.44188	3.27494
Entomobryiidae	0.29585	-2.26701	2.30741
Empididae	0.07317	-30.35164	0.02563
Elmidae	0.39290	-0.28259	2.53624
Dytiscidae	0.75563	0.89877	3.18516
Dixidae	0.78803	2.34160	3.25942

Curculionidae	0.13333	65.57214	0.93557
Culicidae	0.73779	-0.26039	3.20568
Corixidae	0.36517	3.60023	2.40993
Corduliidae	0.26791	-2.37385	1.93891
Copepoda	0.22648	2.21525	1.76537
Conchostraca	0.86035	-2.63441	3.26909
Coenagrionidae	0.09881	2.72957	0.33460
Cladocera	0.01213	2.61125	-3.27727
Chironomidae	0.40283	-0.16740	2.55238
Ceratopogonidae	0.77217	-0.43345	3.18288
Caenidae	0.64034	0.13455	3.07852
Bulininae	0.32003	1.82142	2.25628
Bithyniidae	0.88216	-1.64788	3.27659
Belostomatidae	0.06307	-7.54488	0.04823
Baetidae	0.98319	-0.00516	3.29278
Atyidae	0.35780	0.38168	2.38764
Araneae	0.25177	-1.52394	2.14045
Ancylidae	0.85218	-0.82716	3.26654
group11	0.36874	-2.35430	2.49735
group10	0.73610	0.12871	3.16039
group9	0.00343	2.77855	-5.78703
group8	0.95278	0.16614	3.29068
group7	0.23103	0.97619	1.72352
group6	0.77769	-0.05966	3.20716
group5	0.59283	0.42005	3.00011
group4	0.72825	-0.06749	3.16484
group3	0.11180	-1.34441	0.89890
group2	0.94853	-0.06122	3.29282

group1	0.92049	-0.01886	3.27273

Appendix III

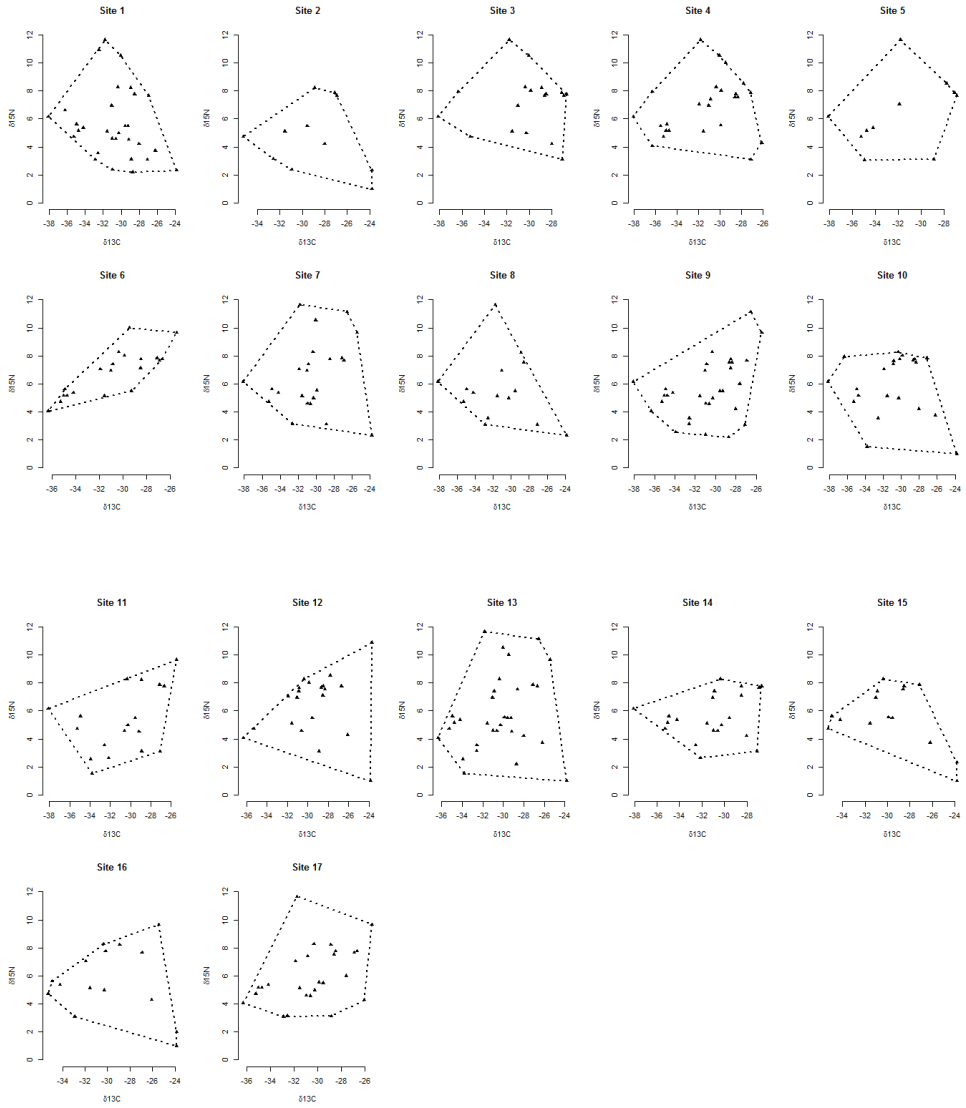


Figure S6.1: The average δ ^{15}N and a low δ^{13}C in bi-plot space for each taxonomic group at each site (1 to 17), the dashed line represents the convex hull around the extent of the occupied niche space, niche width is calculated as the volume within this area.

Table S6.1 The mean regional metrics for all taxa, calculated from the site metrics weighted by the abundance of each taxa at each site. Bacteria is calculated as number of bacilli per dry mass in mg of organism per ml of supernatant during extraction.

Taxa	Total abundance	δ13C	δ15N	Bacteria	Niche Width	Vulnerability	Generality
Aequidens tetramerus	1	-29.45	10	0.00	**95.84**	**15.59**	**21.44**
Aeshnidae	5	-31.55	5.12	12.29	57.44	14.88	7.93
Ampullariidae	2	-23.83	1.99	0.00	59.13	7.00	8.00
Ancylidae	49	-32.6	3.15	89.22	75.89	12.70	16.28
Anura	11	-29.91	5.53	31.90	72.07	11.53	15.36
Araneae	21	-26.92	7.65	16.63	58.48	10.53	11.76
Argulidae	1	-23.73	10.86	0.00	67.97	13.40	9.57
Baetidae	16	-35.05	5.15	736.39	53.33	10.50	14.31
Belostomatidae	53	-29.57	5.48	30.44	63.03	10.91	14.02
Caenidae	78	-34.75	5.15	549.97	67.56	13.59	17.69
Ceratopogonidae	93	-32.57	3.56	77.42	77.38	12.22	16.33
Chichlidae sp.	1	-29.45	10	0.00	95.84	15.59	21.44
Chironomidae	1402	-35.27	4.73	183.17	62.68	11.60	14.11
Cleithracara maronii	1	-29.45	10	0.00	27.28	11.91	10.92
Coenagrionidae	33	-31.80	11.65	554.97	69.53	12.92	13.56
Copella carsevennensis	26	-27.12	7.85	287.74	56.03	9.38	11.75
Copepoda	17	-32.10	3.28	0.00	61.80	7.85	9.36
Corduliidae	1	-32.48	10.91	0.00	73.97	13.57	20.36
Corixidae	56	-34.21	5.37	349.14	58.65	11.74	14.51
Crenicichla saxatilis	2	-27.61	6	0.00	65.46	13.90	17.38
Culicidae	42	-30.99	4.59	121.96	69.78	13.49	17.96
Curimatopsis crypticus	15	-30.84	7.64	0.00	70.64	11.36	10.43
Dytiscidae	14	-28.03	4.22	49.96	58.31	11.22	12.61
Dytiscidae (larvae)	38	-30.62	4.56	5.06	62.49	11.57	15.79
Elmidae	7	-26.19	3.74	297.07	68.19	11.94	15.55
Elmidae (larvae)	3	-29.17	4.51	0.00	75.80	12.05	18.60

Euryhynchus amazoniensis	18	-29.89	8.01	52.14	58.89	10.01	11.76
Georyssidae	1	-32.36	-0.99	0.00	87.47	12.81	15.77
Gerridae	5	-38.11	6.16	3.54	62.94	9.12	12.85
Haliplidae	11	-33.92	2.55	0.00	68.08	14.18	18.86
Haliplidae (larvae)	7	-33.79	1.52	0.00	76.11	11.11	16.76
Hemigrammus ocellifer	19	-28.35	7.52	0.00	69.57	12.85	13.92
Hemigrammus rodwayi	54	-30.88	7.41	9.62	58.88	12.61	12.90
Hemigrammus unilineatus cayennensis	65	-28.5	7.75	1.11	58.85	12.75	13.47
Hirudinea	38	-25.45	9.65	2.10	79.32	14.14	18.82
Hydrometridae	1	-36.26	6.62	0.00	73.97	13.57	20.36
Hydrophilidae	12	-30.96	2.37	0.00	67.10	14.43	18.67
Hydrophilidae (larvae)	4	-32.14	2.63	0.00	50.02	9.42	13.42
Hydroptilidae	82	-36.33	4.07	0.00	68.84	12.49	16.07
Isopoda	5	-32.91	3.09	0.00	73.97	13.57	20.36
Krobia aff guianensis sp 1	11	-30.04	10.53	68.80	71.86	12.40	15.21
Krobia aff guianensis sp 2	3	-27.80	8.53	0.00	66.55	11.22	11.29
Leptophlebiidae	9	-34.75	5.15	1760.85	55.61	7.99	13.87
Lestidae	27	-28.62	7.53	0.00	63.91	12.32	15.69
Libellulidae	221	-31.55	5.12	91.93	67.76	12.22	13.61
Littorinidae	152	-23.83	0.99	0.00	51.12	6.41	6.24
Macrovelidae	10	-28.88	3.12	203.47	70.63	11.71	16.41
Moenkhausia grandisquamis	3	-30.18	7.75	0.00	59.13	7.00	8.00
Moenkhausia hemigramoides	3	-30.18	7.75	0.00	72.42	10.00	11.00
Moenkhausia surinamensis	23	-28.51	7.11	0.00	50.84	12.74	11.97
Nannostomus beckfordi	8	-26.69	7.77	0.00	59.49	12.33	13.34
Nepidae	4	-26.56	11.13	14.16	84.50	14.25	18.18
Noteridae	47	-30.28	4.97	20.46	67.17	11.18	14.91
Noteridae (larvae)	9	-27.09	3.09	287.19	61.79	11.27	14.54
Oligochaeta	169	-34.94	5.62	223.54	62.30	11.27	14.63

Ostracoda	19	-28.73	2.19	89.86	74.14	14.49	20.28
Palaemonetes	45	-31.92	7.04	34.09	67.31	9.18	10.84
Philopotamidae	1	-36.33	4.07	0.00	67.97	13.40	9.57
Physidae	16	-23.83	0.99	32.64	63.48	10.86	11.96
Planorbidae	57	-23.83	2.32	90.47	53.45	7.90	8.68
Pleidae	7	-34.21	5.37	0.00	68.25	12.25	14.78
Polycentrus punctatus	7	-36.33	7.93	9.50	62.61	12.34	9.78
Pristella maxilaris	113	-30.36	8.26	24.49	61.88	12.14	12.42
Protoneuridae	7	-28.62	7.53	38.37	86.53	13.37	18.38
Pyrrhulina filamentosa	17	-28.32	7.71	2.17	63.37	12.05	9.38
Rivulidae sp.	1	-28.91	8.21	0.00	57.44	14.88	7.93
Rivulus lungi	25	-28.91	8.21	316.53	57.94	7.97	9.13
rivulus ocellatus	2	-28.91	8.21	0.00	63.25	7.82	10.75
Scirtidae	1	-32.36	-0.99	0.00	67.22	15.80	19.75
Scirtidae (larvae)	2	-32.36	-0.99	0.00	57.60	7.00	14.00
Simuliidae	3	-35.55	5.5	2612.34	63.70	6.87	14.71
Sphaeriidae	24	-36.33	4.07	0.00	94.65	15.60	21.37
Tabanidae	10	-29.26	5.49	45.22	70.81	13.83	19.46
Tanypodinae	119	-31.04	6.94	191.94	66.70	12.97	15.07
Trichodactylidae	26	-26.09	4.28	0.00	59.82	7.43	8.59
Trichoptera	2	-36.33	4.07	0.00	27.28	11.91	10.92
Veliidae	13	-38.11	6.16	787.38	65.56	8.04	11.24

Table S6.2 Site location, biodiversity indices

Site	Northing	Easting	Total Organisms	Original	Jacknife estimate
1	5.631467	-53.7072	219	25	38
2	4.8608	-52.2568	169	10	14
3	4.736183	-52.327	205	19	31
4	4.838083	-52.3533	99	15	21
5	5.296233	-53.0514	31	9	13
6	5.362083	-53.0337	76	12	18
7	5.605467	-53.8277	121	17	25
8	4.834467	-52.3021	123	23	33
9	5.3941	-52.992	412	31	47
10	5.42875	-53.0888	183	16	24
11	5.3772	-52.9539	243	19	29
12	4.3342	-52.1525	126	15	23
13	4.300417	-52.1233	430	33	52
14	5.03535	-52.5165	320	18	26
15	4.929067	-52.4038	208	10	13
16	5.6666	-53.7799	279	16	25
17	5.602283	-53.8364	342	21	30

Table S6.3 Local food web metrics and niche width

Site	Niche Width	Average Generality	Average Vulnerability	Average Connectance
1	73.97	20.36	13.57	0.20
2	45.45	5.00	6.00	0.16
3	57.44	7.93	14.88	0.26
4	63.70	14.71	6.87	0.21
5	59.73	8.00	4.00	0.24
6	27.28	10.92	11.91	0.28
7	87.47	15.77	12.81	0.24
8	63.25	10.75	7.82	0.10
9	67.22	19.75	15.80	0.23
10	72.42	11.00	10.00	0.23
11	57.60	14.00	7.00	0.19
12	67.97	9.57	13.40	0.24
13	95.84	21.44	15.59	0.22
14	42.44	12.83	11.85	0.27
15	43.60	7.67	8.63	0.17
16	59.13	8.00	7.00	0.19
17	63.70	15.00	12.00	0.23

Appendix IV

OPEN ACCESS Freely available online

PLOS | NEGLECTED TROPICAL DISEASES

First Detection of *Mycobacterium ulcerans* DNA in Environmental Samples from South America

Aaron Morris[1,2]*, Rodolphe Gozlan[2,3], Estelle Marion[4], Laurent Marsollier[4], Demetra Andreou[1], Daniel Sanhueza[2,4], Rolland Ruffine[2], Pierre Couppié[5], Jean-François Guégan[2]

1 University of Bournemouth, School of Applied Sciences, Dorset, United Kingdom, **2** UMR MIVEGEC IRD-CNRS-Universités de Montpellier 1 et 2, Centre IRD de Montpellier, Montpellier, France, **3** UMR BOREA IRD-MNHN-Université Pierre et Marie Curie, Muséum National d'Histoire Naturelle, Paris, France, **4** Equipe Inserm Avenir ATOMycA, CRCNA INSERM U892 and CNRS U6299, Université et CHU d'Angers, Angers, France, **5** Institut Guyanais de Dermatologie Tropicale, EA 2188, Centre Hospitalier André Rosemon, Cayenne, French Guiana

Abstract

The occurrences of many environmentally-persistent and zoonotic infections are driven by ecosystem changes, which in turn are underpinned by land-use modifications that alter the governance of pathogen, biodiversity and human interactions. Our current understanding of these ecological changes on disease emergence however remains limited. Buruli ulcer is an emerging human skin disease caused by the mycobacterium, *Mycobacterium ulcerans*, for which the exact route of infection remains unclear. It can have a devastating impact on its human host, causing extensive necrosis of the skin and underlying tissue, often leading to permanent disability. The mycobacterium is associated with tropical aquatic environments and incidences of the disease are significantly higher on floodplains and where there is an increase of human aquatic activities. Although the disease has been previously diagnosed in South America, until now the presence of *M. ulcerans* DNA in the wild has only been identified in Australia where there have been significant outbreaks and in western and central regions of Africa where the disease is persistent. Here for the first time, we have identified the presence of the aetiological agent's DNA in environmental samples from South America. The DNA was positively identified using Real-time Polymerase Chain Reaction (PCR) on 163 environmental samples, taken from 23 freshwater bodies in French Guiana (Southern America), using primers for both IS2404 and for the ketoreductase-B domain of the *M. ulcerans* mycolactone polyketide synthase genes (KR). Five samples out of 163 were positive for both primers from three different water bodies. A further nine sites had low levels of IS2404 close to a standard CT of 35 and could potentially harbour *M. ulcerans*. The majority of our positive samples (8/14) came from filtered water. These results also reveal the Sinnamary River as a potential source of infection to humans.

Citation: Morris A, Gozlan R, Marion E, Marsollier L, Andreou D, et al. (2014) First Detection of *Mycobacterium ulcerans* DNA in Environmental Samples from South America. PLoS Negl Trop Dis 8(1): e2660. doi:10.1371/journal.pntd.0002660

Editor: Jessica N. Ricaldi, Universidad Peruana Cayetano Heredia, Peru

Received May 27, 2013; **Accepted** December 9, 2013; **Published** January 30, 2014

Funding: This work is supported by Laboratoire d'Excellence Centre d'Etude de la Biodiversité Amazonienne (grant ANR-10-LABX-25-01) through its integrative research programme BIOHOPSYS on Biodiversity and infectious diseases, and a 3-years Bournemouth University PhD fellowship grant to AM. This research partially received technical support and funding from ANR SEST Programme EREMIBA (Ref. 0010005) for JFG. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: amorris@bournemouth.ac.uk

Introduction

Buruli ulcer is an emerging human skin disease caused by the mycobacterium, *Mycobacterium ulcerans*. In the same genus as other high profile infectious agents which cause tuberculosis and leprosy, the prevalence of Buruli ulcer has been rapidly increasing in developing parts of the world [1]. Typically found in moist, tropical areas, the disease can have a devastating impact on a human host, causing extensive necrosis of the skin and underlying tissue, often leading to permanent disability. Whilst the exact route of infection remains unclear, the causative mycobacterium is strongly associated with aquatic environments and incidences of the disease are significantly higher on floodplains, or where people come into continual contact with rivers, ponds, swamps and lakes [2,3]. Since the discovery of specific PCR primers, which are sensitive enough to detect *M. ulcerans* from the environment, it has been found on, or within numerous aquatic species and in environmental samples from aquatic systems [1,4–7]. As the mechanism of infection remains inconclusive, it is not possible to definitely advocate the optimum aquatic conditions necessary for the disease to flourish. We can however speculate on preferred habitats, based on the current evidence and physiological traits. Sequencing of the genome reveals a lack of *crtI* which is responsible for the production of phytoene dehydrogenase, an enzyme that within *M. marinum* is necessary for the synthesis of light-inducible carotenoid pigments [8]. The loss of need for these pigments, which give protection against UV-induced damage, suggests that *M. ulcerans* lives in conditions where it does not require this ability. Culturing *in vitro* has also shown the mycobacterium to have a preference for a low oxygen environment [9]. These traits therefore indicate that the mycobacteria have a preference for environments with low light and oxygen levels.

Identification of the bacteria in the environment has been generally isolated to parts of Australia where there have been significant outbreaks [10–12] or tropical regions of Africa where

Author Summary

This study provides the first ever recorded extraction of *Mycobacterium ulcerans* DNA from the environment in South America, specifically from French Guiana an ultra-peripheral French territory. *M. ulcerans* is the causative agent responsible for the devastating necrotic skin infection Buruli ulcer, which is prevalent in many tropical countries, notably in western and central Africa, and continues to present outbreaks in the developing world. Despite this, our understanding of the disease remains limited, routes of infection, environmental sources and ubiquity within the environment are still uncertain and only within the past decade have we begun to understand more about this emerging disease.

the disease is persistent [4,5,13–17]. Whilst in South America human cases of the disease have been definitively present since 1969 as for example in French Guiana [1,18,19] and previous data suggests the presence of *M. ulcerans* DNA in environmental water sources [18], the DNA has never before been definitively identified in the environment. To ascertain whether environmental conditions where the pathogen is present in French Guiana are comparable to other continents and laboratory analysis extractions of DNA were taken from numerous freshwater bodies in addition to abiotic readings of the water. This has led to the first ever detection of *M. ulcerans* from environmental samples in South America. Further to this we undertook analysis of whether the construction of the Petit-Saut Dam, an hydro-electric installation upstream from an endemic area of French Guiana is correlated with a decline in the number of cases downstream near the city of Sinnamary subsequent to its construction.

Materials and Methods

French Guiana is a French ultra-peripheral territory within South America, bordering the countries of Brazil to the east and south and Suriname to the west. It is 83,534 km^2 and has a low population density, with most inhabitants residing along a 50 km wide coastal strip dominated by swampy areas. The altitude in the coastal strip is *c.a.* 10–30 m. The rest of the country is almost purely pristine primary tropical rain forest. During the period of 1969 to 2011, 242 cases of Buruli ulcer were reported (Couppie, personal communication) the low population (274,652) makes this number of infections relatively high, with an average number of new notified cases of 2.09/100,000 persons per year (Figure 1).

To identify the presence of *M. ulcerans*, 163 environmental samples were taken from 23 water bodies in French Guiana

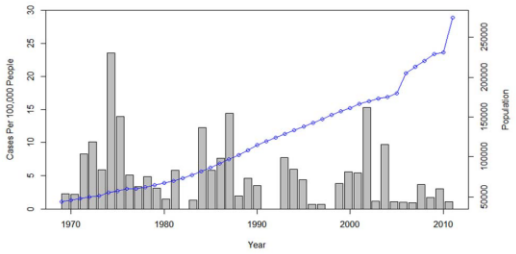


Figure 1. Cases of Buruli ulcer from 1969 to 2012 per 100,000 people, the dotted line represents the increasing human population living in French Guiana.
doi:10.1371/journal.pntd.0002660.g001

(Figure 2). Sites where determined by looking for water-bodies in areas where the disease was prevalent and had a high chance of human contact. Indications of this included: trails near or through the site, presence of fishing or boating activities or other recreational uses and close proximity to human settlement. In addition, we looked for similar sites in areas where there have been no cases of disease and sites which had little human contact. Preference was given to sites which remained present during the dry season, as this was indicative of permanence.

Samples taken included: water, soil/sediment and detritus and when present dominant aquatic plant species, algae, biofilms and samples of the semi-aquatic plant *Montrichardia arborescens* (Araceae), a plant species which is characteristic of Amazonian swamps (Table S1 in Text S1). These represent a range of previously described habitats for *M. ulcerans* where positive samples have been identified in other continents [6,10,13]. In certain cases when water samples were taken the location contained dense aquatic vegetation, biofilms were unavoidably disturbed from the leaves and stems and it was not possible to take a sample without also capturing biofilms, it is clearly stated in cases where these samples are positive that they contain both water and biofilms.

Water was collected (50 ml) from various sampling points at each site with an attempt to cover all meso-habitats, this generally included: bank-side, the centre of the water body, water from within aquatic vegetation, shaded areas and exposed areas. Water was taken to the laboratory and samples were filtered through 1.6 µmglass microfiber filters before being filtered through 0.4 µm cellulose nitrate membrane and the residue collected and frozen at −20°C for future analysis.

Soil was similarly taken from aquatic areas around the site representing various habitats. Plants, algae and biofilms were collected at various locations when present at the site.

Dissolved oxygen, *pH*, conductivity and water temperature where measured at approximately the locations where each sample was taken.

DNA extraction was carried out using PowerSoil DNA extraction kits (Mo Bio Lab., Carlsbad, USA). Using Real-Time Polymerase Chain Reactions, two primer pairs were used to positively identify the bacillus following standardised methods [10,20] Table 1.

Accession numbers

IS2404: AF003002
KR: BX649209

Number of cases since the building of Petit-Saut Dam

To help identify the biological niche of *M.ulcerans* in French Guiana we performed a series of statistical tests to identify whether the DNA was predominantly present within certain abiotic conditions. To account of non-normal distributions of data a Wilcoxon-rank sum test was performed on the number of cases in Sinnamary district per month, per 100,000 people for the 18 years since the Dam was constructed and impoundment of water started in January 1994 against the 18 years prior to this period. This was repeated for comparison for cases across the whole of French Guiana.

Differences between abiotic parameters at sites

To account for non-normal distribution, Wilcoxon tests were used to determine statistical differences between abiotic measurements (*pH*, conductivity, dissolved oxygen and water temperature) at *M. ulcerans* positive sites against the same measurements at negative sites. These tests were also repeated to identify any

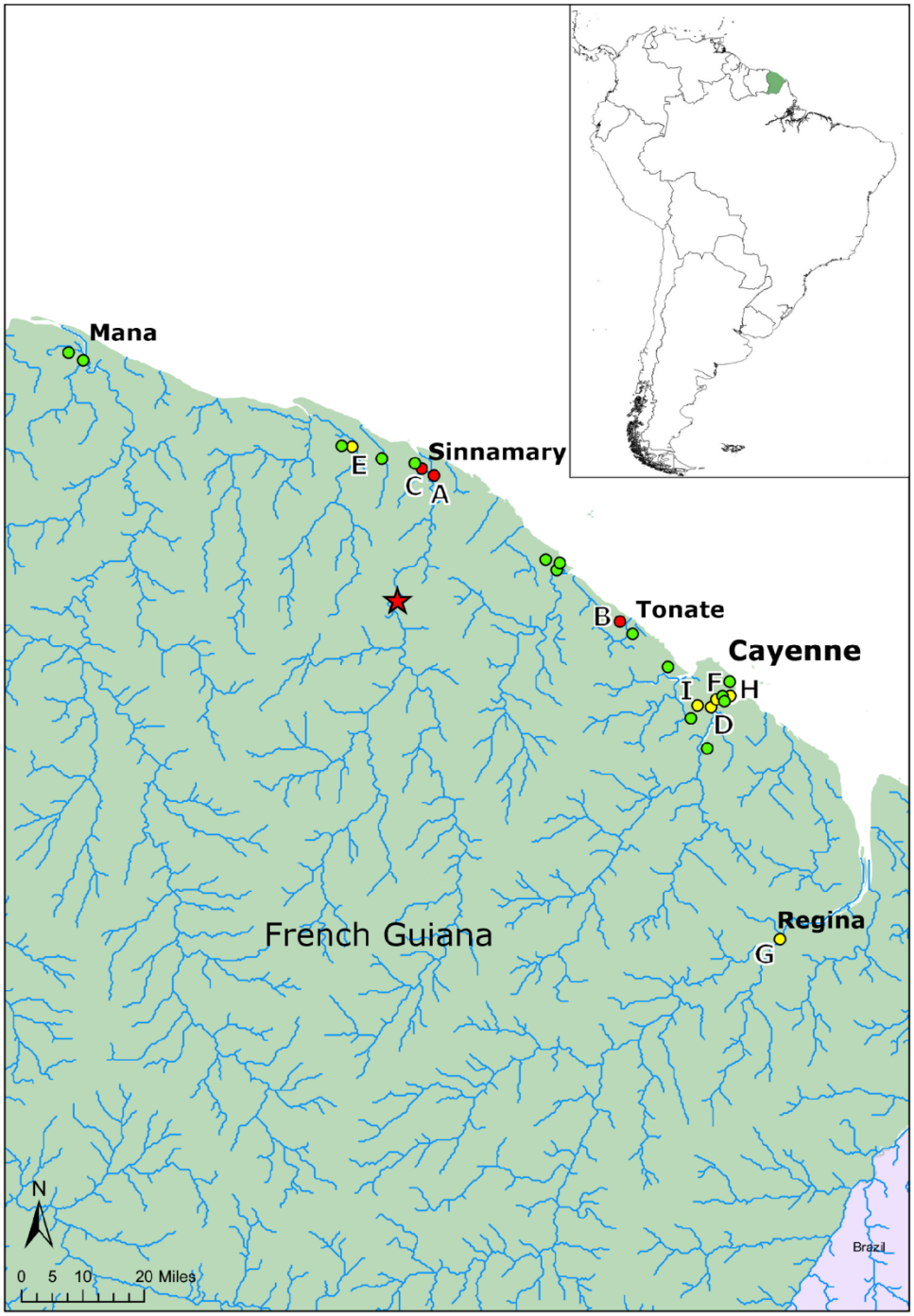


Figure 2. Map showing the approximate location of the twenty-three sampling sites (Text S1). In green are the negative sites to both *Mycobacterium ulcerans* during this survey, in yellow are the site positive to IS2404 alone and in red are the sites positive to both IS2404 and KR. The dam of Petit-Saut built in 1994 is indicated by a star [21].
doi:10.1371/journal.pntd.0002660.g002

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

Primer or probe	Sequence (5′-3′)	Nucleotide positions	Amplicon size (bp)
IS*2404* TF	AAAGCACCACGCAGCATCT	27746–27762	59
IS*2404* TR	AGCGACCCCAGTGGATTG	27787–27804	
IS*2404* TP	6 FAM CGTCCAACGCGATC MGBNFQ	27768–27781	
KRTF	TCACGGCCTGCGATATCA	3178–3195	65
KRTR	TTGTGTGGGCACTGAATTGAC	3222–3242	
KRTP	6 FAM-ACCCCGAAGCACTG-MGBNFQ	3199–3212	

TF, forward primer; TR, reverse primer; TP, probe. Nucleotide position based on the first copy of the amplicon in pMUM001.
doi:10.1371/journal.pntd.0002660.t001

statistical differences between abiotic factors between sites which were positive for *M. ulcerans*.

Results

Five samples out of 163 were identified as positive for both IS2404 and KR (CT<35) from three different water bodies in French Guiana; Site A (lat 5.3772, long −52.953883) in Sinnamary (1/19 samples positive), site B (lat 5.3941, long −52.992017) near Tonate (3/5 samples positive) and site C (lat 5.03535, long −52.516483) on Route JoJo a road just outside Sinnamary (1/6 samples positive) (see Table 2 and Figure 2). A further nine sites had low levels of IS2404 close to a standard CT of 35 (Table 3) and could potentially harbour *M. ulcerans*, however it is not possible to conclude this definitively as they were negative for KR.

All three sites (five samples) positive for both IS2404 and KR were typologically similar, with areas that were highly stagnant forming shallow water bodies and with high levels of *M. arborescens* plant growth (Table 2). Dissolved oxygen levels at sites A and B, at the locations where the mycobacterium was found were low (<1 mg/l), at site C dissolved oxygen was higher (1.93 mg/l); however other locations within the site and within a few meters were also similarly low to sites A and C (<1 mg/l). There was no statistical significance for differences in *pH*, conductivity, dissolved oxygen and temperature between positive sites and negative sites.

Number of cases since the building of Petit-Saut Dam

The number of cases in the 18 years after the construction of the Petit-Saut Dam were significantly lower per 100,000 people per year in Sinnamary at 0.6 than during the 18 years before at 10.1 (Test statistic (W) = 210, p-value = 0.0296), the number of cases in the whole of French Guiana remain statistically similar before and after construction (Test statistic (W) = 132.5, p-value = 0.359).

Discussion

It is difficult yet to draw definitive conclusions about potential abiotic or biotic factors affecting mycobacterium levels in French Guiana from our current information, whilst our results showed no statistical differences; this may be because of too few positive sites to draw reliable statistical comparisons. Our results extend the range of geographical distribution of environmental *M. ulcerans* to another continent, i.e. south America, where previously the DNA has not been identified in the environment. Similarly we have found that the majority of our positive samples came from filtered water (8/14), which has been the case in other countries [13]. Whilst it could be concluded that the mycobacteria are more prevalent in the water column, this may be an artefact of the sampling methods used. When sampling water we are able to concentrate a large volume onto filters for extraction, whereas with soil, biofilms and plants, the limitations of the extraction kits mean we can only utilise a relatively small fraction of material from a site (0.25 g). In addition the freshwater bodies that we found positive samples in this area were on floodplains, suggesting this category of environment constitutes the source for environmentally-persistent mycobacteria or a "receptacle" concentrating these bacilli from further upstream. From our results we can also recommend the importance of taking multiple abiotic readings from a single site, because we are assessing the ecology of an organism that is living in a microscopic environment, we must consider the large variation in abiotic parameters within a few meters or less in the same water-body (for example Site B, Tables 2 and 3).

Table 2. Details of positive samples for both IS2404 and KR (CT<35).

Site	Sample Type	Locality	CT Value (IS204)	CT Value (KR)	Bact/ml	Dissolved Oxygen (mg/l)	*pH*	Conductivity (µS/m)	Water Temp (°C)	Season
A	Water & Biofilms	Sinnamary	34.70	33.69	18.33	0.74	5.55	197	24.9	Dry
B	Water filtrand	Nr Tonate	31.93	33.48	138.9	0.93	6.237	150.8	28.9	Dry
B	Water filtrand	Nr Tonate	29.77	30.92	674.6	0.93	6.237	150.8	28.9	Dry
B	Water filtrand	Nr Tonate	32.11	34.5	121.8	0.03	5.896	92.4	26.4	Dry
C	Water filtrand	Route JoJo	34.14	34.47	27.49	1.93	5.342	32	26.8	Dry

Abiotic parameters taken from where the sample was collected within the water body is also included. Mycobacteria per ml of 50 ml water sample, or 0.25grams of solid sample.
doi:10.1371/journal.pntd.0002660.t002

Table 3. Details of samples with sites positive for IS2404 (CT>35).

Site	Sample Type	Locality	CT Value	Bact/ml	Dissolved Oxygen (mg/l)	*pH*	Conductivity (µS/m)	Water Temp (°C)	Season
A	Water filtrand	Sinnamary	36.2	10.53	0.46	5.3	201	24.7	Dry
E	Soil/sediment	Nr Iracoubu	35.41	17.9	NA	NA	NA	NA	Dry
F	Soil/sediment	Matoury	37.2	4.877	4.94	6.764	99.50	28.8	Dry
D	Water filtrand	Matoury	37.43	2.48	0.05	6.442	14.32	26.6	Dry
G	Water filtrand	Regina	36.33	5.553	3.23	5.50	15.6	27.3	Dry
G	Water filtrand	Regina	37.39	2.56	0.95	5.35	13.5	26.0	Dry
G	Soil/sediment	Regina	35.96	7.297	1.26	5.27	13.9	26.2	Dry
H	*M. arborescens* skin	Montjolly	35.53	15.39	0.57	5.62	280	30	Dry
I	Filamentous algae	Matoury	36.61	7.33	4.3	4.87	34.8	25.6	Dry

Abiotic parameters taken from where the sample was collected within the water body is also included. Bacteria per ml of 50 ml water sample, or 0.25grams of solid sample.

doi:10.1371/journal.pntd.0002660.t003

The low levels of cases combined with a low population density in a territory such as French Guiana suggest identification from the environment would be difficult; however we were able to find positive sites with less than 200 samples, which in themselves are relatively small components of a system. This would reinforce the possibility that *M. ulcerans* is a fairly common and widely distributed mycobacterium, and it is other factors, e.g. socio-economic (i.e. levels of human contact with water, sanitation etc), transmission related (i.e. presence of potential vectors, hosts or reservoirs), or habitat modifications (deforestation, dam construction, etc), that might be the primary drivers of cases.

The finding of the majority of positive sites from the area around Sinnamary River downstream to the dam is of interest, the first cases of Buruli ulcer in French Guiana were recorded here and approximately 10% of human cases of Buruli ulcer in French Guiana concern the inhabitants of the Sinnamary. In the Sinnamary region the number of cases has been very low since 1994, with significantly less cases in the 18 years after 1994, despite an increasing human population. Whilst changes in the behaviour of people may have an influence, the building of the Petit-Saut Dam (Figure 2) may also be playing a role. The dam has profound effects on the level of water which comes into the area, possibly reducing flooding or regulating water flows, or potentially limiting mycobacteria being brought upstream from the rainforest and riverine swamp areas.

The results presented here suggest there remains a potential for future infection in French Guiana, as our knowledge of transmission improves we hope to be able to identify at which point a water body harbouring the bacteria can become a site of transmission to people. They also provide a basis for future studies of *M. ulcerans* in South America.

Supporting Information

Text S1 Supporting information includes details of the number and type of samples taken for all sites.
(DOC)

Acknowledgments

We would like to thank the anonymous reviewers for their constructive feedback and help shaping this manuscript.

Author Contributions

Conceived and designed the experiments: AM RG JFG DA. Performed the experiments: AM DA EM LM RR DS. Analyzed the data: AM. Contributed reagents/materials/analysis tools: AM EM LM DA PC. Wrote the paper: AM RR JFG LM PC.

References

1. Merritt RW, Walker ED, Small PLC, Wallace JR, Johnson PDR, et al. (2010) Ecology and Transmission of Buruli Ulcer Disease: A Systematic Review. PLoS Negl Trop Dis 4: e911.
2. Merritt RW, Benbow ME, Small PLC (2005) Unraveling an Emerging Disease Associated with Disturbed Aquatic Environments: The Case of Buruli Ulcer. Frontiers in Ecology and the Environment 3: 323–331.
3. Brou T, Broutin H, Elguero E, Asse H, Guegan J-F (2008) Landscape Diversity Related to Buruli Ulcer Disease in Côte d'Ivoire. PLoS Negl Trop Dis 2: e271.
4. Mosi L, Williamson H, Wallace JR, Merritt RW, Small PLC (2008) Persistent Association of Mycobacterium ulcerans with West African Predaceous Insects of the Family Belostomatidae. Applied and Environmental Microbiology 74: 7036–7042.
5. Marsollier L, Robert R, Aubry J, Saint Andre J-P, Kouakou H, et al. (2002) Aquatic Insects as a Vector for Mycobacterium ulcerans. Appl Environ Microbiol 68: 4623–4628.
6. Marsollier L, Stinear T, Aubry J, Saint André JP, Robert R, et al. (2004) Aquatic Plants Stimulate the Growth of and Biofilm Formation by Mycobacterium ulcerans in Axenic Culture and Harbor These Bacteria in the Environment. Applied and Environmental Microbiology 70: 1097–1103.
7. Marsollier L, Sévérin T, Aubry J, Merritt RW, Saint André J-P, et al. (2004) Aquatic Snails, Passive Hosts of Mycobacterium ulcerans. Applied and Environmental Microbiology 70: 6296–6298.
8. Ramakrishnan L, Tran HT, Federspiel NA, Falkow S (1997) A crtB homolog essential for photochromogenicity in Mycobacterium marinum: isolation, characterization, and gene disruption via homologous recombination. Journal of Bacteriology 179: 5862–5868.
9. Palomino JC, Obiang AM, Realini L, Meyers WM, Portaels F (1998) Effect of Oxygen on Growth of Mycobacterium ulcerans in the BACTEC System. Journal of Clinical Microbiology 36: 3420–3422.
10. Fyfe JAM, Lavender CJ, Johnson PDR, Globan M, Sievers A, et al. (2007) Development and Application of Two Multiplex Real-Time PCR Assays for the Detection of Mycobacterium ulcerans in Clinical and Environmental Samples. Applied and Environmental Microbiology 73: 4733–4740.
11. Fyfe JAM, Lavender CJ, Handasyde KA, Legione AR, O'Brien CR, et al. (2010) A Major Role for Mammals in the Ecology of *Mycobacterium ulcerans*. PLoS Negl Trop Dis 4: e791.
12. Johnson PDR, Azuolas J, Lavender CJ, Wishart E, Stinear TP, et al. (2007) Mycobacterium ulcerans in Mosquitoes Captured during Outbreak of Buruli Ulcer, Southeastern Australia. Emerging Infectious Diseases 13: 1653–60.

13. Williamson HR, Benbow ME, Nguyen KD, Beachboard DC, Kimbirauskas RK, et al. (2008) Distribution of *Mycobacterium ulcerans* in Buruli Ulcer Endemic and Non-Endemic Aquatic Sites in Ghana. PLoS Negl Trop Dis 2: e205.
14. Williamson HR, Benbow ME, Campbell LP, Johnson CR, Sopoh G, et al. (2012) Detection of *Mycobacterium ulcerans* in the Environment Predicts Prevalence of Buruli Ulcer in Benin. PLoS Negl Trop Dis 6: e1506.
15. Portaels F, Meyers WM, Ablordey A, Castro AG, Chemlal K, et al. (2008) First Cultivation and Characterization of *Mycobacterium ulcerans* from the Environment. PLoS Negl Trop Dis 2: e178.
16. Debacker M, Aguiar J, Steunou C, Zinsou C, Meyers WM, et al. (2004) Mycobacterium ulcerans Disease (Buruli ulcer) in Rural Hospital, Southern Benin, 1997–2001. Emerging Infectious Diseases 10: 1391–8.
17. Marston BJ, Diallo MO, Horsburgh CR, Diomande I, Saki MZ, et al. (1995) Emergence of Buruli ulcer disease in the Daloa region of Cote d'Ivoire. The American journal of tropical medicine and hygiene 52: 219–224.
18. Guerra H, Palomino JC, Falconi E, Bravo F, Donaires N, et al. (2008) Mycobacterium ulcerans disease, Peru. Emerging infectious diseases 14: 373.
19. McGann H, Stragier P, Portaels F, Gascoyne-Binzi D, Collyns T, et al. (2009) Buruli ulcer in United Kingdom tourist returning from Latin America. Emerging infectious diseases 15: 1827.
20. Marion E, Eyangoh S, Yeramian E, Doannio J, Landier J, et al. (2010) Seasonal and Regional Dynamics of *M. ulcerans* Transmission in Environmental Context: Deciphering the Role of Water Bugs as Hosts and Vectors. PLoS Negl Trop Dis 4: e731.
21. Lehner B, Verdin K, Jarvis A (2006) HydroSHEDS. World Wildlife Fund US, Washington, DC.

Appendix V

OPEN

Emerging Microbes and Infections (2014) 3, e56; doi:10.1038/emi.2014.56

www.nature.com/emi

ORIGINAL ARTICLE

Complex temporal climate signals drive the emergence of human water-borne disease

Aaron Morris[1,2], Rodolphe E Gozlan[1,3], Hossein Hassani[1], Demetra Andreou[1], Pierre Couppié[4] and Jean-François Guégan[2]

Predominantly occurring in developing parts of the world, Buruli ulcer is a severely disabling mycobacterium infection which often leads to extensive necrosis of the skin. While the exact route of transmission remains uncertain, like many tropical diseases, associations with climate have been previously observed and could help identify the causative agent's ecological niche. In this paper, links between changes in rainfall and outbreaks of Buruli ulcer in French Guiana, an ultraperipheral European territory in the northeast of South America, were identified using a combination of statistical tests based on singular spectrum analysis, empirical mode decomposition and cross-wavelet coherence analysis. From this, it was possible to postulate for the first time that outbreaks of Buruli ulcer can be triggered by combinations of rainfall patterns occurring on a long (i.e., several years) and short (i.e., seasonal) temporal scale, in addition to stochastic events driven by the El Niño-Southern Oscillation that may disrupt or interact with these patterns. Long-term forecasting of rainfall trends further suggests the possibility of an upcoming outbreak of Buruli ulcer in French Guiana.
Emerging Microbes and Infections (2014) **3,** e56; doi:10.1038/emi.2014.56; published online 6 August 2014

Keywords: climate; coherence analysis; El Niño/La Niña; *Mycobacterium ulcerans*; rainfall; singular spectrum analysis; Southern America

INTRODUCTION

The identification of cohering patterns between climate and infectious disease using time series analysis is an important component in understanding the ecological niche of disease causing agents and in predicting future outbreaks. Such correlations can occur with both local and large-scale climatic oscillations.[1–9] The mechanisms behind these relationships often vary and have been attributed both to the direct and indirect effects of changing climate, notably for vector-borne and reservoir-borne diseases for which a component of their life-cycle may be highly sensitive to any rainfall or temperature variation. For example, decreases in precipitation can create pools of stagnant water which are breeding grounds for vectors,[4,10] flooding may cause contamination of surface water and wells through overflow of sewage systems and the failure of septic tanks,[11] or the loss of crops or water supplies may cause habitual changes or immunological deterioration in the population.[12] A key problem with time series analysis in long-term datasets is the separation of signals and stochastic noise. Noise can hide cohering patterns, while within a series, there may be a number of competing signals of varying strength. For example, rainfall measures over time have a number of seasonal changes; over a long period, an ecological process such as disease outbreaks may only be linked to changes in one of these components.

Buruli ulcer (BU) is an emerging human skin disease caused by the mycobacterium *Mycobacterium ulcerans*. Related to high-profile infections tuberculosis and leprosy, prevalence has been increasing in certain developing parts of the world and in some areas, it is more common than its aforementioned relations.[13] Despite this, knowledge of the infection remains limited. Typically found in moist, tropical areas, the disease can have a devastating impact on its host, causing extensive necrosis of the skin and underlying tissue, often leading to permanent disability if left untreated.[14,15] While the route of infection remains unclear, the bacillus is strongly associated with aquatic environments and incidences of the disease are significantly higher on floodplains, or where people come into continual contact with rivers, ponds, swamps and lakes.[16–18] In addition, DNA and cultures of the mycobacterium have been found on, or within numerous aquatic species.[13,19–22] This relationship with aquatic systems makes it an interesting candidate to look for coherent patterns with environmental parameters like rainfall and changes in large climatic drivers such as the El Niño-Southern Oscillation (ENSO).

MATERIALS AND METHODS

Environmental and disease data

The only accurate long-term (decadal) dataset for cases of BU is from French Guiana in South America, with records going back to 1969. French Guiana also has a well-recorded history of rainfall during this period making it highly suitable for this study. French Guiana is a French ultraperipheral territory bordering the countries of Brazil to the east and south and Suriname to the west. Although large at 83 534 km^2, the population density is very low with almost all inhabitants located in a thin strip along the coastline. The rest of the country is predominantly pristine primary tropical rain forest and

[1]Bournemouth University, Dorset BH12 5BB, UK; [2]UMR MIVEGEC, IRD-CNRS-Universités de Montpellier 1 et 2, Centre IRD de Montpellier, 34394 Montpellier cedex 5, France; [3]UMR BOREA, IRD-MNHN-Université Pierre et Marie Curie, Muséum National d'Histoire Naturelle, 375231 Paris cedex 5, France and [4]Institut Guyanais de Dermatologie Tropicale, EA 3593, Centre Hospitalier André Rosemon, Cayenne, French Guiana
Correspondence: A Morris
E-mail: amorris@bournemouth.ac.uk
Received 28 February 2014; revised 3 June 2014; accepted 16 June 2014

contains the 33 900 km^2 Guiana Amazonian national park, as well as a wealth of important ecosystems ranging from marshland to coastal mangroves. The numbers of BU cases were obtained from Cayenne Central hospital records dating back to 1969 till 2012, with identification based on a combination of histopathological, microbiological, clinical and genetic analysis. This dataset is the most accurate long-term data on BU to our knowledge, which can be used for coherence with climatic factors. A potential issue with lesion causing diseases is the variation in time between appearance of symptoms and seeking of medical attention, reflected in lesion size. As French Guiana is part of the European Union and is a low-population French territory, case reporting and assessment of lesions incurred a minimal delay; active surveillance of the disease was being undertaken with health-care professionals who are trained to recognize BU being present in all towns and villages. Disease cases are distributed across the territory in line with the distribution of the population and are present almost ubiquitously where there are people. Rainfall data were obtained from Météo-France and were recorded as the average rainfall in millimeters per month from 17 weather stations (Figure 1) across the populated coastal area of the territory from 1969 to 2012. Due to the restricted range of inhabited areas, a small human population and therefore, a relatively low number of cases in each locality, an average rainfall reading from the stations along the coastal area was taken and compared to data on all BU cases across French Guiana. ENSO data for the period were taken from the American National Climatic Data Center and measured as the sea surface temperature (SST) of the equatorial Pacific Ocean.

Ethical provisions

Written consent for participation to the study was obtained from patients in all instances. BU cases received treatment appropriately according to the French laws in public health, which are also under application in this territory. The study protocol was authorized by Cayenne General Hospital authorities according to French ethical rules. The database was declared to the Commission National Informatique et Libertés (CNIL NO 3X#02254258) following French law requirements. The database did not include names or any variable that could allow the identification of patients.

Singular spectrum analysis (SSA) decomposition and reconstruction

Since the introduction of SSA by Broomhead and King,[23,24] it has been applied successfully to several economic, financial and industrial time series [25–27] and has also been used previously in the analysis of coherence between disease and climate.[1,6–8,28] Consider the real-valued non-zero time series $Y_T=(\gamma_1, \gamma_2, ..., \gamma_T)$ of sufficient length T. The main purpose of SSA is to decompose the original series into a sum of series, so that each component in this sum can be identified as either a trend, periodic or quasiperiodic component (perhaps, amplitude-modulated), or noise. This is followed by a reconstruction of the original series. Each corresponding stage involves two primary steps, for decomposition; embedding and singular value decomposition and for reconstruction; grouping and reconstruction. For a detailed description of each stage, see Golyandina *et al* (2001).[29] In short, decomposition breaks the time series down into its constituent components (in this instance, repeating seasonal and long-term patterns in rainfall). Once isolated, it is possible to identify stochastic noise within the leftover signal and remove it before reconstructing a new noise-free time series.

Each seasonal component of the time series was first identified using periodograms, graphical representations of the distribution of power (or variance) among different frequencies. Independence of each seasonal component was also tested. The main concept in studying SSA component properties is 'separability', which characterizes how well different components can be separated from each other. SSA decomposition of the series Y_T can only be successful if the resulting additive components of the series are approximately separable. A natural measure of dependence between two time series $Y_T^{(1)}$ and $Y_T^{(2)}$ is the weighted correlation or 'ω-correlation'.[29] To identify correlations between all the components within the time series, a ω-correlation matrix was created. This shows the ω-correlation for the components

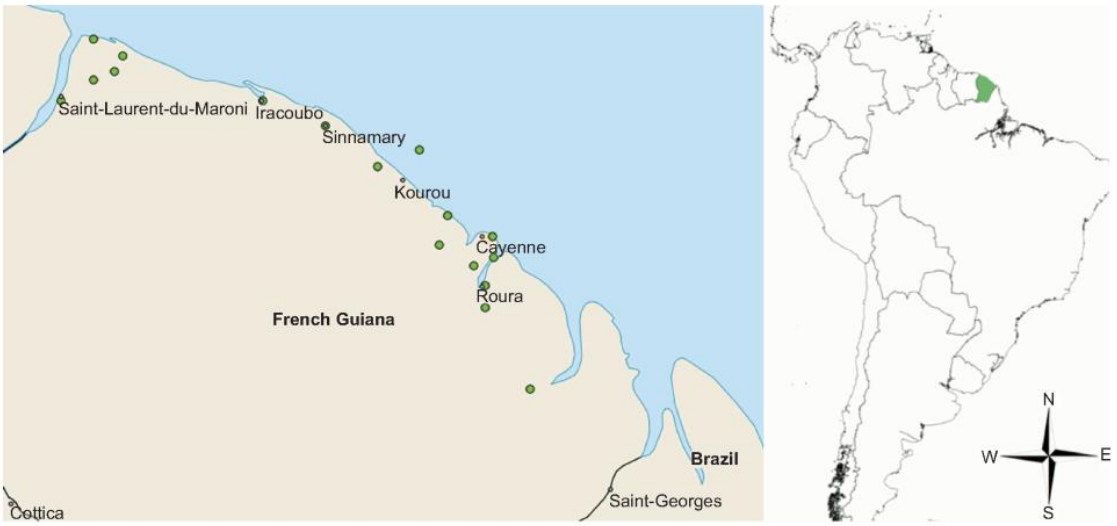


Figure 1 Map of French Guiana showing the location of 17 weather stations along the coast of French Guiana and the position of French Guiana within South America.

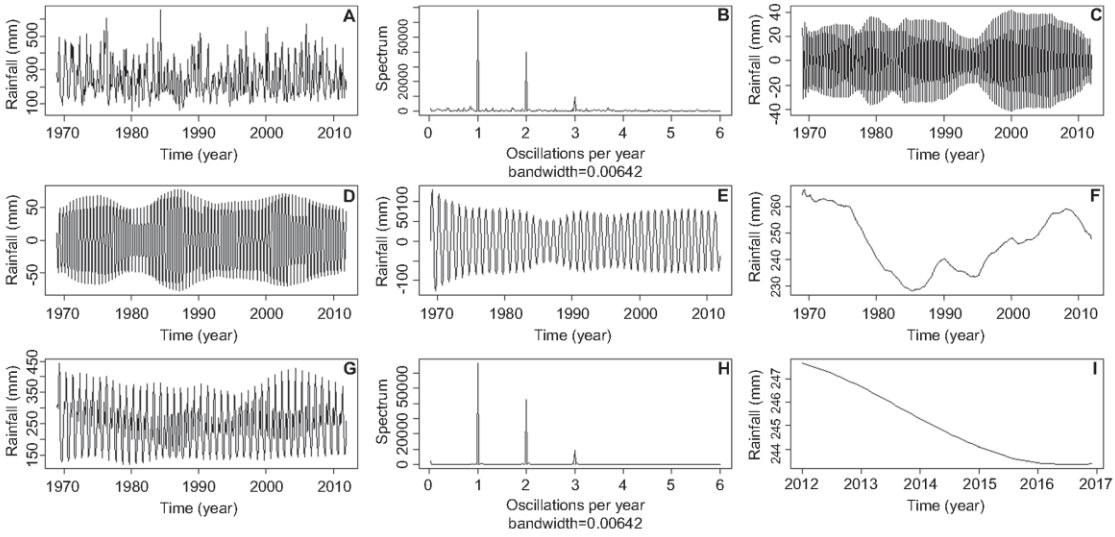


Figure 2 Monthly time series data showing the decomposition, reconstruction and forecasting of datapoints using SSA. (**A**) Original rainfall time series average for 17 weather stations along the coast of French Guiana, from 1969 to 2012. (**B**) Periodograms of the rainfall time series identifying significant repeating patterns once per year, twice per year and three times per year. SSA extracted component corresponding to periodogram spike of (**C**) three times per year, 4-month component, (**D**) twice per year, 6-month component and (**E**) once per year, 12-month component. (**F**) The extracted rainfall trend. (**G**) The reconstructed rainfall time series after the removal of stochastic noise. (**H**) A second periodogram of the reconstructed rainfall series showing less stochastic noise around the three main repeating patterns. (**I**) Forecasting of the rainfall trend to 2017 using sequential SSA.

in a 20-grade gray scale from white to black corresponding to the absolute values of correlations from 0 to 1 (with 0 being no correlation and 1 being absolute correlation).

The series was further analyzed using sequential SSA, which refers to an analysis of the residuals.[25] As a result of sequential SSA, it is possible to identify signal components which were incorrectly classified as noise through the earlier decomposition and then combine them together in order to build a signal which is shown as total rainfall residual. It is then possible to include the total signal extracted from the residuals in to the earlier reconstruction and use the post-sequential SSA reconstruction for forecasting new data points.

Empirical mode decomposition (EMD)

To explore relationships between seasonal components of the rainfall time series and any corresponding seasonal changes in BU, repeating intra- and inter-annual patterns in BU incidence were extracted using EMD. EMD is a technique developed specifically to decompose non-stationary and nonlinear time series and has been successfully applied to a number of climatic and epidemiological datasets.[1,30] The method is an iterative process which builds a number of oscillatory signals called intrinsic mode functions (IMFs), which are repeatedly subtracted from the time series; each iteration results in an IMF with a longer periodic component until just a trend signal is present. For a detailed description of EMD, see Huang *et al.*[31] A periodogram was created for each IMF to assess the frequency of the repeating pattern through time.

Trend coherence analysis

To identify the relationship between the rainfall trend obtained via SSA, SST and the number of human BU cases in French Guiana, continuous wavelet transforms were used. Wavelets have been utilized previously in various branches of ecological theory[32–36] and to identify relationships between disease and long-term climatic patterns.[37] They have the benefit of being unaffected by non-stationary time series, which are often found in ecology.[35] Cross-wavelet coherence analysis was performed using the biwavelet R package [38] with the Morlet mother wavelet and 2000 Monte Carlo randomizations. In order to further characterize the association between the time series, phase analysis was also undertaken to identify the phasing difference between the two, for example, whether one precedes the other,[39] this is indicated by arrows on the wavelet plots. To further assess the statistical significance of the patterns exhibited by the wavelet approach, null models were tested. To create time series for the null models, bootstrapping was used to construct from the observed time series, control datasets, which share properties of the original series under the following null hypothesis: the variability of the observed time series or the association between two time series is not different to that expected from outbreaks independent of the rainfall trend.

Seasonal coherence analysis

To identify correlations between seasonal components of the rainfall time series extracted via SSA, SST and seasonal changes in BU cases represented by the extracted IMF signals, cross-correlation functions were used.[40] These measure the similarity between two oscillating time series as a function of a time lag applied to one of them and can be used with stationary time series (i.e., series which statistical properties including mean, variance and autocorrelation are consistent over time).

RESULTS

Singular spectrum analysis

Periodograms of the raw rainfall time series (Figure 2A) identified seasonal components which oscillated yearly for 4-, 6- and 12-month periods (repeating patterns occurring, tri-annually, bi-annually and once

per year, Figure 2B) the signals of each repeating oscillation were isolated along with the long-term trend of rainfall (Figures 2C–2F). Remaining data were considered noise and the components were reconstructed into a noise-free signal (Figure 2G), a second periodogram of the new reconstructed signal shows the repeating patterns are intact, while stochastic noise has been removed (Figure 2H). Figure 2I shows the resulting output of forecasting from sequential SSA, with rainfall beginning to decline as French Guiana enters a trough of dryer years after several years of high rainfall. The 4-month component corresponds to two rainy seasons, one long rainy season and one short, and also to the main dry season from August to December (Figure 3A). The 6-month bi-annual component corresponds to the two rainy seasons and is a reflection of the strength of these two seasons (Figure 3B), while the 12-month component shows the overall rainfall level for the year (Figure 3C). Separability of these components was confirmed with a ω-correlation matrix showing that these seasonal components did not show high levels of correlation with other components (Figure 4).

Empirical mode decomposition

Seven IMF series were produced by EMD of repeating patterns from the BU case data (Figures 5A–5H); periodograms of these show that the first IMF has high levels of variation across the spectrum and therefore, is likely noise, the second is a bi-yearly repeating pattern and the third is a measure of cases per year, the fourth over 2 years and the fifth over 4 years. Subsequent IMF series are long-term trends (Figures 5I–5O).

Trend coherence analysis

During the period 1969–2012, the series was dominated by four inter-annual peaks in rainfall followed by three inter-annual periods of recessions, with three corresponding peaks and recessions of BU disease cases. The results of the wavelet coherence analyses showed a statistically significant correlation between the two time series for 1979–2000, with cohering peaks and troughs over periods of approximately 8 years (Figure 6A). Phase analysis, indicated by the arrows pointing downwards suggests a preceding relationship of rainfall change occurring before BU cases. In essence, after a peak in rainfall, during a dry period, the number of BU cases increases. Null models showed no significant relationships between rainfall and disease cases. The analysis between cases and SST showed less coherence; however, there were two short periods during the mid-1970s and early 1990s which weakly corresponded (Figure 6B).

Seasonal coherence analysis

Cross-correlation functions between the IMF signals representing intra- and inter-annual patterns in BU cases (first, second, third, fourth and fifth IMFs) and both SST and the reconstructed rainfall pattern from SSA analysis showed a number of corresponding signatures.

SST did not correlate with the SSA derived rainfall series (Figure 7A), but did corresponded with rainfall before SSA was applied (Figure 7B), suggesting that SST spikes cause higher levels of unpredictable rainfall anomalies, which during SSA analysis were classified as 'noise' or stochastic events. SST also corresponded with the first IMF of BU cases (Figure 7C), which was similarly classified as noise. This may mean that SST creates rainfall anomalies, which cause high levels of BU but do not follow any set seasonal or long-term patterns (i.e., random one off events). SST fluctuations further matched with inter-annual variation of BU cases with long lag periods, suggesting that the total number of cases over these periods is increased by the influence of SST-driven anomalies (Figures 7D–7G). In particular the fourth IMF, where a below average SST (La Niña) value

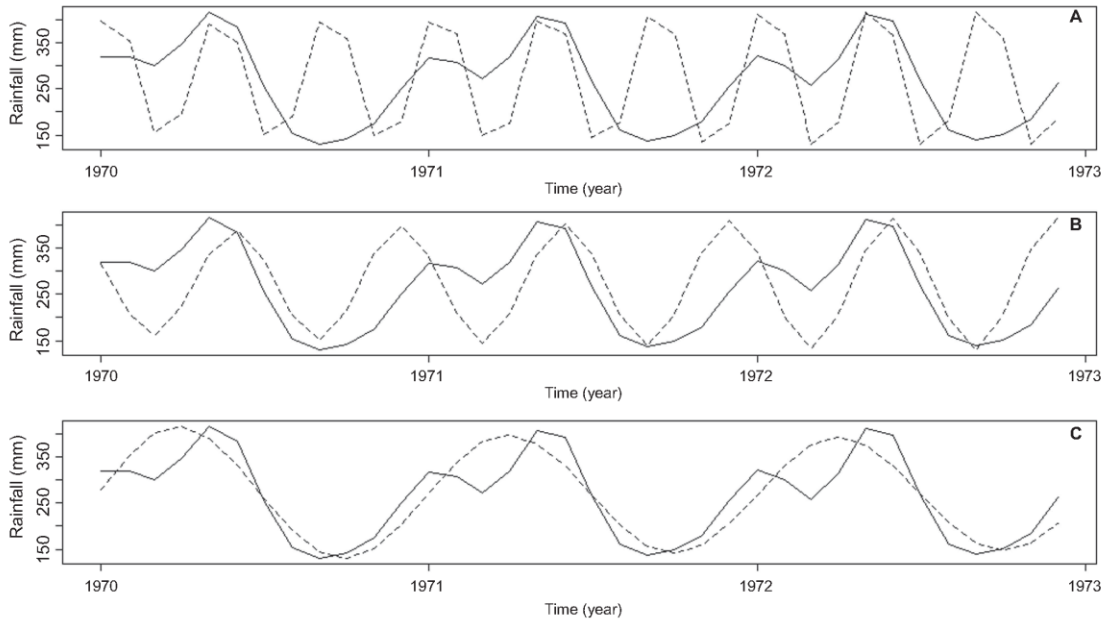


Figure 3 Three-year period of extracted components (dashed line) plotted against the same period of the reconstructed rainfall (solid line), showing which parts of the rainfall series the components represent. (**A**) Four-month component corresponds to both rainy seasons and the dry season. (**B**) Six-month component corresponds to the two rainy seasons. (**C**) Twelve-month component represents the rainfall for the whole year.

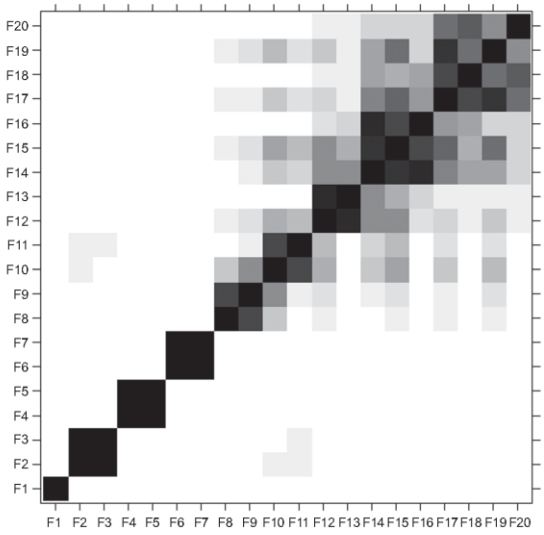


Figure 4 ω-correlation matrix, the F values represent oscillating components within a year (i.e. F6 is the 6-month bi-annual component). The level of correlation can be found by finding the component of interest along the X-axis and looking up the Y-axis to see where it corresponds with other rainfall components. Large values of ω-correlation between reconstructed components indicate that they should possibly be gathered into one group and correspond to the same component in SSA decomposition. The matrix uses a 20-grade gray scale from white to black corresponding to the absolute values of correlations from 0 to 1 (with 0 being no correlation and 1 being absolute correlation).

produces a higher than average level of BU cases over 2 years after a lag time of approximately 18 months (Figure 7F). Conversely, a peak in SST (El Niño) creates a decline in the 4-year oscillation of BU cases (Figure 7G).

The reconstructed rainfall series has a corresponding relationship with both the second and third IMF (Figures 7I and 7J), suggesting that an above average level of precipitation contributes to an above average spike in bi-annual BU cases, and also the overall yearly number of cases. By looking at the SSA-derived components of rainfall against the IMF signatures of BU, it is possible to identify the exact rainfall components which are driving these spikes. The SSA-derived 4-month component correlated with no IMF signatures (Figures 8A–8C). Figures 8D–8F show the correlations of the first, second and third IMF signatures with the SSA-derived 6-month component (i.e., the strength of both the two wet seasons); a significant correlation is identified with the bi-annual BU cases (second IMF) with an approximate 5-month lag; therefore, because of the average reported incubation periods of 3–5 months[41] and the lag time of several weeks before diagnosis, the two spikes in cases per year are most likely to be driven by the strength of the two spikes in rainfall per year rather than the two dry seasons. The 12-month component which is a measure of the total rainfall per year did not correlate with any seasonal IMF signatures (Figures 8G and 8H), but as expected correlated with the overall level of BU per year (Figure 8I).

DISCUSSION

The identification of more than one temporal correlation between rainfall and BU disease, in addition to wet weather anomalies outside of usual seasonal peaks in rainfall being driven by SST, highlights the complexity of using environmental changes in predicting disease outbreaks. The stochastic SST-driven incidences and the influence of long-term rainfall in addition to seasonal drivers is a first for BU in the world. While a similar long- and short-term climatic pattern has already been linked to cholera outbreaks,[28,42] the results of this paper show that such patterns are likely to be more spread among aquatic infectious pathogens such as with *M. ulcerans* and that SST relationships may be driving stochastic cases.

While further study will be required to fully understand what niche *M. ulcerans* occupies within French Guiana, in this paper, the identification of a relationship with rainfall provides important testable insights into its disease ecology. It also exemplifies the importance of using long-term datasets when trying to establish relationships between the environment and infectious disease, and the use of techniques such as wavelets, SSA and EMD to look deeper into time series.

By removing noise and decomposing the time series using SSA, it was possible to look at the influence of each individual seasonal pattern on BU and to identify a cause of important rainfall anomalies, which occur outside the regular seasonal patterns. By further extracting a long-term trend and relating this to cases, the effects of several components of rainfall become apparent, something which would perhaps be lost without decomposition.

During the analysis, EMD had some advantages and disadvantages over SSA, which makes it suitable for differing time series, dependent on the application. In this instance, the ability to successfully identify noise and a seasonal component, in addition to a hierarchy of increasing long-term periodic components in disease cases, was beneficial, particularly when linking disease patterns to SST. SSA, however, provided a more accurate separation of seasonal components within the rainfall data.

The increasing use of wavelets[32–35,37] is an important development for time series analysis in epidemiology; previously in this field, non-stationarity presented a serious problem for relating ecological, epidemiological and climatic datasets.[43–45] Wavelets provide the advantage of being localized in both time and frequency, whereas the standard Fourier transform, traditionally used in time series analysis, is only localized in frequency,[35] which, although useful for identifying constant periodic components, is not able to account for changes in frequency over time.

The results of the relationship between SST and BU cases broadly agrees with similar observations in Australia where it was found that periods of wet weather approximately 16–17 months prior to an outbreak, followed by a period of dry weather for 5 months to be the most suitable for BU emergence.[46] The results presented here suggest that outbreaks of BU over a long period (at least 2 years, as represented by the fourth IMF), correspond with a decline in SST (i.e., a La Niña event) 17 months prior, while the opposite is true for El Niño, with the effect causing a below average decrease in BU cases over the preceding 4 years. A La Niña event in the north of South America corresponds to a marked increase in wet weather anomalies (as corroborated by the significant cross-correlation function between the pre-SSA rainfall and SST), while El-Niño signals an increase in dryer weather. As SST also correlates with high rainfall events and BU cases classified as stochastic, it is possible that SST driven rainfall anomalies create random outbreaks in BU, independent of the usual seasonal cycles over a 2-year period.

The long-term and seasonal relationships between rainfall and cases, i.e., the long-term peaks in BU driven by a recession in rainfall over several years, and the bi-annual peaks in cases driven by spikes in

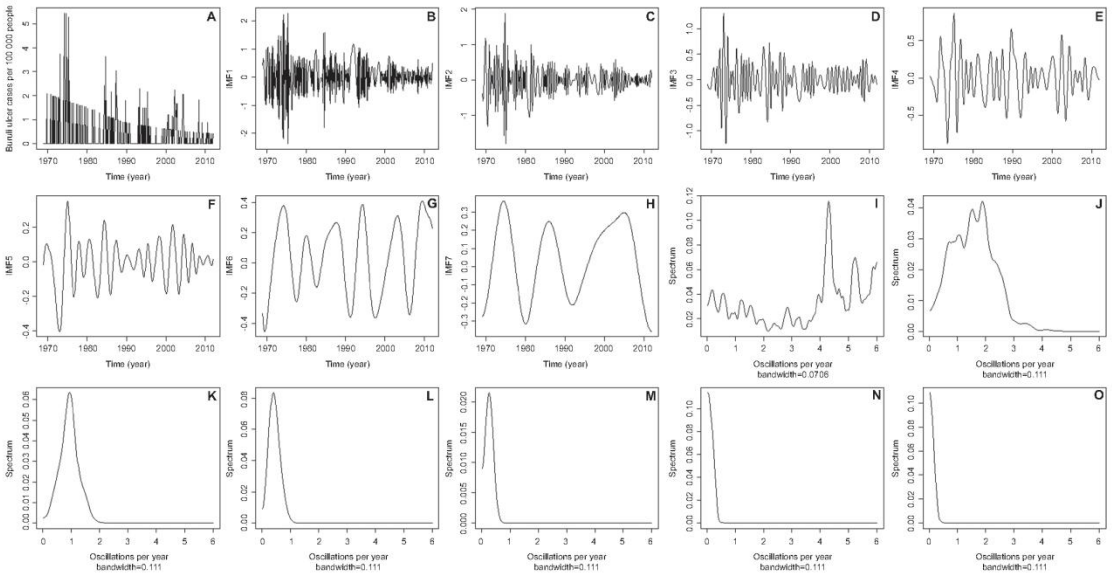


Figure 5 Monthly time series and EMD of Buruli ulcer cases per 100 000 people. (**A**) Monthly time series of Buruli ulcer cases per 100 000 people from 1969 to 2012. (**B**) First IMF. (**C**) Second IMF. (**D**) Third IMF. (**E**) Fourth IMF. (**F**) Fifth IMF. (**G**) Sixth IMF. (**H**) Seventh IMF. Periodograms for (**I**) first IMF showing a high level of variation across the spectra and therefore, should be considered white noise; (**J**) second IMF which has its highest power at two cycles per year; (**K**) third IMF with its highest power at one cycle per year; (**L**) fourth IMF with the highest power approximately at one cycle every 2 years; (**M**) fifth IMF with the highest power approximately at one cycle every 4 years; (**N**) sixth IMF with a low level of cycles representing very long-term trends; (**O**) seventh IMF also representing long-term trends.

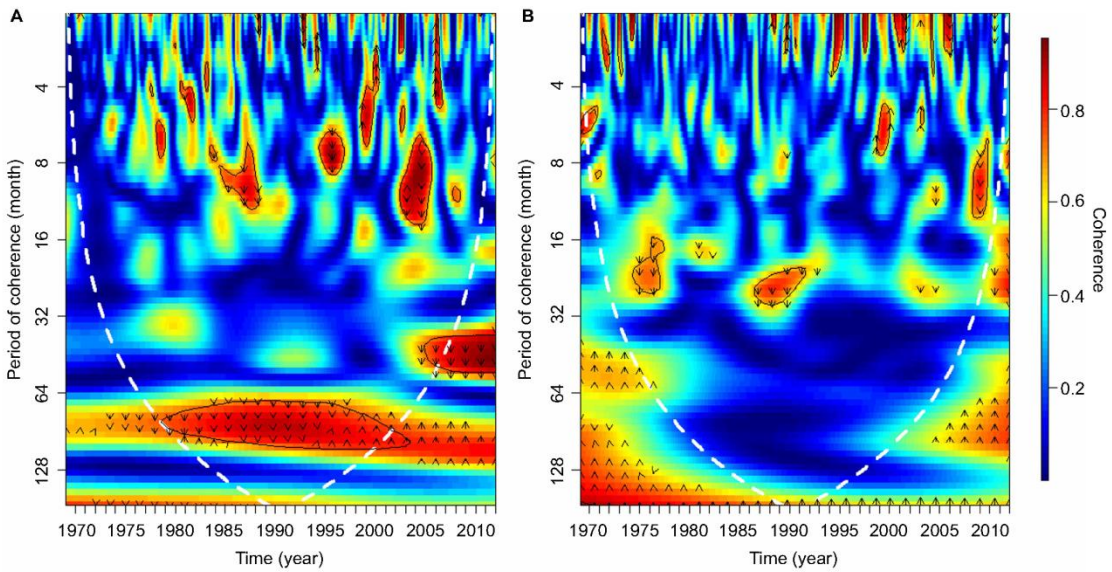


Figure 6 Wavelet coherence between (**A**) Buruli ulcer incidences per 100 000 people and the rainfall trend obtained from SSA. (**B**) Buruli ulcer incidences per 100 000 people and ENSO. The colors are coded from dark blue to dark red with dark blue representing low coherence through to high coherence with dark red. The solid black lines around areas of red show the α=5% significance levels computed based on 2000 Monte Carlo randomizations. The dotted white lines represent the cone of influence; outside this area, coherence is not considered as it may be influenced by edge effects. The black arrows represent the phase analysis and adhere to the following pattern: arrows pointing to the right mean that rainfall and cases are in phase, arrows pointing to the left mean that they are in antiphase, arrows pointing up mean that cases lead rainfall and arrows pointing down mean that rainfall leads cases.

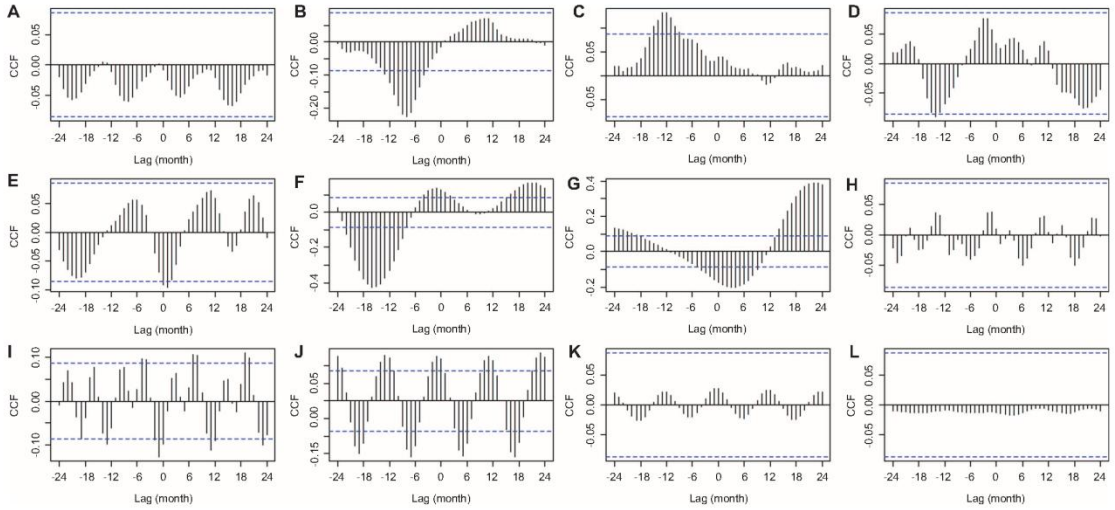


Figure 7 Cross-correlation analysis between (**A**) ENSO and reconstructed rainfall time series from SSA. (**B**) ENSO and original rainfall time series. ENSO and (**C**) first IMF from Buruli ulcer cases, (**D**) second IMF, (**E**) third IMF, (**F**) fourth IMF, (**G**) fifth IMF. Reconstructed rainfall series and (**H**) first IMF; (**I**) second IMF; (**J**) third IMF; (**K**) fourth IMF; (**L**) fifth IMF. Dashed horizontal blue lines in all panels represent the 95% confidence limit; black vertical lines which go beyond the dashed line can be considered non-random cohering oscillations between the two time series being assessed, with the lag period between an above average oscillation in the first time series and a subsequent above average oscillation in the second shown on the X-axis.

the two rainy seasons, could be explained in several ways. While the limited knowledge of BU disease transmission makes it only possible to speculate, the analysis does present an opportunity to shed further light on the ecological niche occupied by *M. ulcerans* and its environmental heterogeneity in space and time. Long periods of wet weather followed by a decrease in rainfall may increase the number of stagnant water bodies and swampland, flowing rivers may recede into a series of isolated pools,[47] while newly formed channels and wetlands will be cut off. This could spark an outbreak of disease vectors, host carriers and other aquatic species that have been identified to

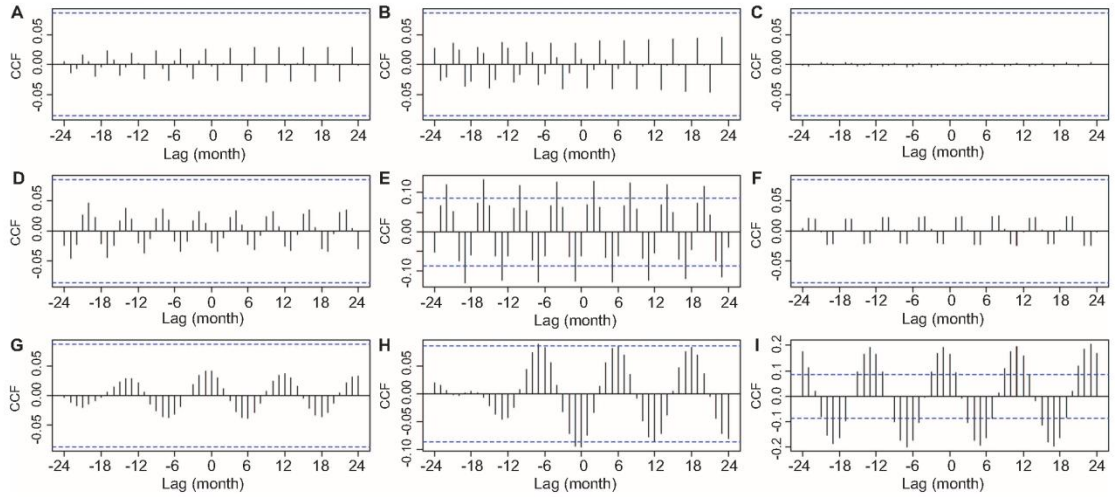


Figure 8 Cross-correlation analysis between rainfall seasonal components extracted by SSA and seasonal IMF series of Buruli ulcer. Four-month component and (**A**) first IMF, (**B**) second IMF, (**C**) third IMF. Six-month component and (**D**) first IMF, (**E**) Second IMF, (**F**) third IMF. Twelve-month component and (**G**) first IMF, (**H**) second IMF, (**I**) third IMF. Dashed horizontal blue lines in all panels represent the 95% confidence limit; black vertical lines which go beyond the dashed line can be considered non-random cohering oscillations between the two time series being assessed, with the lag period between an above average oscillation in the first time series and a subsequent above average oscillation in the second shown on the X-axis.

have a relationship with BU and which thrive in these conditions,[19,22,48–50] or beneficially change the aquatic community structure. A second hypothesis, which assumes that the disease agent does not necessarily require a true symbiotic relationship, would be that the initial rainfall washes the agent into new territory, and once established, the dry weather will again increase the number of preferential water bodies, thereby increasing the number of cases. Both events have the potential to cooccur over a long and short period of time and could potentially be working at differing spatial scales, driven by either long- or short-term rainfall patterns. Previous studies have shown how community assemblages and system dynamics can be significantly altered over varying periods of time at sites with differing sizes, hydrologies and landscape parameters.[18,51–55] The majority of the BU cases occur when both the long- and short-term temporal changes coincide, suggesting weather conducive to providing an increase in infectious habitats or human contact with these habitats during such a period. Using the argument that the disease agent is most abundant in stagnant water, it is possible to hypothesize that large bodies of water created after several years of high rainfall may recede slowly, stagnate and undergo significant changes in biotic community over several dry years. Nested within this period, high peaks in rainy seasons as suggested by an increase in the biannual component, or SST-driven rainfall anomalies, will cause these stagnating areas to swell, potentially flooding or seeping onto pathways where human contact is increased. In addition, periods of high rainfall during periods of long-term dryer climates could also be conducive to the occurrence of potential invertebrate vectors. Previous work on mosquito populations, for example, which have been linked to BU,[48] shows that they are at their highest during periods of high short-term climatic variability, cited as the amount of short-term fluctuations around a mean climate state on a fine time scale.[3,56–58] This may add weight to the idea of mosquito-based vectors, although it is possible that several less studied aquatic species exhibit similar responses. It must also be considered that the relationship could be influenced more by human behavior, for example, dryer years will often induce an increase in recreational use of local water bodies, notably for fishing and hunting.

The results of the forecasting are of potential importance in predicting future disease emergence. The most recent highest rainfall in French Guiana has occurred in 2009 and it is being followed by a 5-year period of dry weather (Figure 2I). Based on the analysis and the identification of a relationship with rainfall, it is possible to predict that this should also be followed by an increase in BU cases in the region.

Caution must be expressed when using time series data, particularly over long periods, as changes in diagnostic capabilities (for example, the introduction of polymerase chain reaction techniques allowing accurate genetic ifentification of the bacteria, knowledge of the disease and improvements in equipment and recording) and the accuracy in weather data collection will cause significant temporal discrepancies. These factors are difficult to address and are inherent to all such studies. It is also important to note that unpredictable external factors such as socioeconomical changes can also be having an unknown influence. The use of long-term weather data also presents the problem of how to include landscape parameters, as these also have an influence on a disease, but are often not available or poorly recorded early on in time. In this instance, however, it seems unlikely that a cyclical pattern is related to a steady change in population and landscape. The robust analysis shows that French Guiana time series for BU cases reveals interesting non-random patterns, which are vital for understanding the ecological niche of this aquatic microbial agent.

ACKNOWLEDGMENTS

This work was supported by Aaron Morris, Rodolphe E Gozlan, Pierre Couppié and Jean-François Guégan, who have benefited from an 'Investissement d'Avenir' grant managed by Agence Nationale de la Recherche called LabEx CEBA (Centre d'Etude sur la Biodiversit' Amazonienne) (ref. ANR-10-LABX-2501). This work was partially supported by a grant from ANR EREMIBA (05 SEST 008-02). Jean-François Guégan is sponsored by the Institut de Recherche pour le Développement and the Centre National de la Recherche Scientifique. This work has benefited from a PhD studentship awarded to Aaron Morris from the Bournemouth University.

1 Chaves LF, Satake A, Hashizume M, Minakawa N. Indian ocean dipole and rainfall drive a moran effect in East Africa malaria transmission. *J Infect Dis* 2012; **205**: 1885–1891.
2 Pascual M, Cazelles B, Bouma MJ, Chaves LF, Koelle K. Shifting patterns: malaria dynamics and rainfall variability in an African highland. *Proc Biol Sci* 2008; **275**: 123–132.
3 Zhou G, Minakawa N, Githeko AK, Yan G. Association between climate variability and malaria epidemics in the East African highlands. *Proc Natl Acad Sci USA* 2004; **101**: 2375–2380.
4 Gagnon AS, Bush AB, Smoyer-Tomic KE. Dengue epidemics and the El Niño Southern Oscillation. *Clim Res* 2001 2001; **19**: 35–43.
5 Hanf M, Adenis A, Nacher M, Carme B. The role of El Nino Southern Oscillation (ENSO) on variations of monthly Plasmodium falciparum malaria cases at the cayenne general hospital, 1996–2009, French Guiana. *Malar J* 2011; **10**: 100.
6 Pascual M, Rodó X, Ellner SP, Colwell R, Bouma MJ. Cholera dynamics and El Niño-Southern Oscillation. *Science* 2000; **289**: 1766–1769.
7 Rodó X, Pascual M, Fuchs G, Faruque ASG. ENSO and cholera: a nonstationary link related to climate change? *Proc Natl Acad Sci USA* 2002; **99**: 12901–12906.
8 Chaves LF, Pascual M. Climate cycles and forecasts of cutaneous leishmaniasis, a nonstationary vector-borne disease. *PLoS Med* 2006; **3**: e295.
9 Mahamat A, Dussart P, Bouix A *et al*. Climatic drivers of seasonal influenza epidemics in French Guiana, 2006–2010. *J Infect* 2013; **67**: 141–147.
10 Gagnon A, Smoyer-Tomic K, Bush A. The El Niño Southern Oscillation and malaria epidemics in South America. *Int J Biometeorol* 2002; **46**: 81–89.
11 Lipp EK, Huq A, Colwell RR. Effects of global climate on infectious disease: the cholera model. *Clin Microbiol Rev* 2002; **15**: 757–770.
12 Patz JA, Campbell-Lendrum D, Holloway T, Foley JA. Impact of regional climate change on human health. *Nature* 2005; **438**: 310–317.
13 Merritt RW, Walker ED, Small PL *et al*. Ecology and transmission of Buruli ulcer disease: a systematic review. *PLoS Negl Trop Dis* 2010; **4**: e911.
14 World Health Organization. *Buruli ulcer. Diagnosis of Mycobacterium ulcerans disease. A manual for health care providers*. Geneva: WHO, 2001. Available at http://apps.who.int/iris/bitstream/10665/67000/1/WHO_CDS_CPE_GBUI_2001.4.pdf?ua=1 (assessed 1 June 2013).
15 World Health Organization. *Buruli ulcer. Management of Mycobacterium ulcerans disease. A manual for health care providers*. Geneva: WHO, 2001. Available at http://whqlibdoc.who.int/hq/2001/WHO_CDS_CPE_GBUI_2001.3.pdf (assessed 1 June 2013).
16 Merritt RW, Benbow ME, Small PL. Unraveling an emerging disease associated with disturbed aquatic environments: the case of Buruli ulcer. *Front Ecol Environ* 2005; **3**: 323–331.
17 Morris A, Gozlan R, Marion E *et al*. First detection of *Mycobacterium ulcerans* DNA in environmental samples from South America. *PLoS Negl Trop Dis* 2014; **8**: e2660.
18 Garchitorena A, Roche B, Kamgang R *et al*. *Mycobacterium ulcerans* ecological dynamics and its association with freshwater ecosystems and aquatic communities: results from a 12-month environmental survey in Cameroon. *PLoS Negl Trop Dis* 2014; **8**: e2879.
19 Marsollier L, Robert R, Aubry J *et al*. Aquatic insects as a vector for *Mycobacterium ulcerans*. *Appl Environ Microbiol* 2002; **68**: 4623–4628.
20 Marsollier L, Stinear T, Aubry J *et al*. Aquatic plants stimulate the growth of and biofilm formation by *Mycobacterium ulcerans* in axenic culture and harbor these bacteria in the environment. *Appl Environ Microbiol* 2004; **70**: 1097–1103.
21 Marsollier L, Sévérin T, Aubry J *et al*. Aquatic snails, passive hosts of *Mycobacterium ulcerans*. *Appl Environ Microbiol* 2004; **70**: 6296–6298.
22 Mosi L, Williamson H, Wallace JR, Merritt RW, Small PL. Persistent association of *Mycobacterium ulcerans* with West African predaceous insects of the family belostomatidae. *Appl Environ Microbiol* 2008; **74**: 7036–7042.
23 Broomhead DS, King GP. Extracting qualitative dynamics from experimental data. *Physica D* 1986; **20**: 217–236.
24 Broomhead DS, King GP. On the qualitative analysis of experimental dynamical systems. In: Sarkar S (ed). *Nonlinear phenomena and chaos*. Bristol: Adam Hilger, 1986: 113–144.
25 Hassani H. *Singular spectrum analysis: methodology and comparison*. Munich: University Library of Munich, 2007.
26 Hassani H, Heravi S, Zhigljavsky A. Forecasting European industrial production with singular spectrum analysis. *Int J Forecast* 2009; **25**: 103–118.

27 Hassani H, Thomakos D. A review on singular spectrum analysis for economic and financial time series. *Stat Interface* 2010; **3**: 377–397.
28 Koelle K, Rodo X, Pascual M, Yunus M, Mostafa G. Refractory periods and climate forcing in cholera dynamics. *Nature* 2005; **436**: 696–700.
29 Golyandina N, Nekrutkin V, Zhigljavsky A. *Analysis of time series structure: SSA and related techniques*. Boca Raton, FL: Chapman and Hall/CRC, 2001.
30 Hurtado LA, Cáceres L, Chaves LF, Calzada JE. When climate change couples social neglect: malaria dynamics in Panamá. *Emerg Microbes Infect* 2014; **3**: e28.
31 Huang NE, Shen Z, Long SR *et al*. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc R Soc Lond Ser A Math Phys Eng Sci* 1998; **454**: 903–995.
32 Cazelles B, Chavez M, Berteaux D *et al*. Wavelet analysis of ecological time series. *Oecologia* 2008; **156**: 287–304.
33 Maraun D, Kurths J. Cross wavelet analysis: significance testing and pitfalls. *Nonlinear Processes Geophys* 2004; **11**: 505–514.
34 Maraun D, Kurths J, Holschneider M. Nonstationary Gaussian processes in wavelet domain: synthesis, estimation, and significance testing. *Phys Rev E Stat Nonlin Soft Matter Phys* 2007; **75**(1 Pt 2): 016707.
35 Torrence C, Compo GP. A practical guide to wavelet analysis. *Bull Am Meteorol Soc* 1998; **79**: 61–78.
36 Klvana I, Berteaux D, Cazelles B. Porcupine feeding scars and climatic data show ecosystem effects of the solar cycle. *Am Nat* 2004; **164**: 283–297.
37 Chowell G, Munayco C, Escalante A, McKenzie FE. The spatial and temporal patterns of falciparum and vivax malaria in Peru: 1994–2006. *Malar J* 2009; **8**: 142.
38 Gouhier TC. *Biwavelet: conduct univariate and bivariate wavelet analyses*. 0.14 ed. Loverpool: Gouhier TC, 2012. Available at http://biwavelet.r-forge.r-project.org (accessed 16 August 2013).
39 Cazelles B, Stone L. Detection of imperfect population synchrony in an uncertain world. *J Anim Ecol* 2003; **72**: 953–968.
40 Shumway R, Stoffer D. Time series regression and exploratory data analysis. In: *Time series analysis and its applications*. New York: Springer, 2011: 47–82.
41 Trubiano JA, Lavender CJ, Fyfe JA, Bittmann S, Johnson PD. The incubation period of Buruli ulcer (*Mycobacterium ulcerans* infection). *PLoS Negl Trop Dis* 2013; **7**: e2463.
42 Pascual M, Bouma MJ, Dobson AP. Cholera and climate: revisiting the quantitative evidence. *Microbes Infect* 2002; **4**: 237–245.
43 Benton TG, Plaistow SJ, Coulson TN. Complex population dynamics and complex causation: devils, details and demography. *Proc Biol Sci* 2006; **273**: 1173–1181.
44 Cazelles B, Hales S. Infectious diseases, climate influences, and nonstationarity. *PLoS Med* 2006; **3**: e328.
45 Hastings A. Transient dynamics and persistence of ecological systems. *Ecol Lett* 2001; **4**: 215–220.
46 van Ravensway J, Benbow ME, Tsonis AA *et al*. Climate and landscape factors associated with Buruli ulcer incidence in Victoria, Australia. *PLoS ONE* 2012; **7**: e51074.
47 Cazelles B, Chavez M, McMichael AJ, Hales S. Nonstationary influence of El Niño on the synchronous dengue epidemics in Thailand. *PLoS Med* 2005; **2**: e106.
48 Fyfe JA, Lavender CJ, Handasyde KA *et al*. A Major role for mammals in the ecology of *Mycobacterium ulcerans*. *PLoS Negl Trop Dis* 2010; **4**: e791.
49 Portaels F, Elsen P, Guimaraes-Peres A, Fonteyne PA, Meyers WM. Insects in the transmission of *Mycobacterium ulcerans* infection. *Lancet* 1999; **353**: 986.
50 Williamson HR, Benbow ME, Nguyen KD *et al*. Distribution of *Mycobacterium ulcerans* in Buruli ulcer endemic and non-endemic aquatic sites in Ghana. *PLoS Negl Trop Dis* 2008; **2**: e205.
51 Arthington AH, Balcombe SR, Wilson GA, Thoms MC, Marshall J. Spatial and temporal variation in fish-assemblage structure in isolated waterholes during the 2001 dry season of an arid-zone floodplain river, Cooper Creek, Australia. *Marine Freshwater Res* 2005; **56**: 25–35.
52 Harris GP. Temporal and spatial scales in phytoplankton ecology. Mechanisms, methods, models, and management. *Can J Fish Aquat Sci* 1980; **37**: 877–900.
53 Kratz TK, Frost TM, Magnuson JJ. Inferences from spatial and temporal variability in ecosystems: long-term zooplankton data from lakes. *Am Nat* 1987; **129**: 830–846.
54 Ruetz CR, Trexler JC, Jordan F, Loftus WF, Perry SA. Population dynamics of wetland fishes: spatio-temporal patterns synchronized by hydrological disturbance? *J Anim Ecol* 2005; **74**: 322–332.
55 Maio JD, Corkum LD. Relationship between the spatial distribution of freshwater mussels (Bivalvia: Unionidae) and the hydrological variability of rivers. *Can J Zool* 1995; **73**: 663–671.
56 Githeko A, Ndegwa W. Predicting malaria epidemics in the Kenyan Highlands using climate data: a tool for decision makers. *Glob Change Hum Health* 2001; **2**: 54–63.
57 Loevinsohn ME. Climatic warming and increased malaria incidence in Rwanda. *Lancet* 1994; **343**: 714–718.
58 Paaijmans KP, Blanford S, Bell AS, Blanford JI, Read AF, Thomas MB. Influence of climate on malaria transmission depends on daily temperature variation. *Proc Natl Acad Sci USA* 2010; **107**: 15135–15139.