 doi:10.20944/preprints201908.0031.v1

Aquatic Hemiptera in Southwest Cameroon: Biodiversity of Potential Reservoirs of *Mycobacterium ulcerans* and Multiple *Wolbachia* Sequence Types Revealed by Metagenomics

Seraphine N. Esemu[1,2], Xiaofeng Dong[3], Achah J. Kfusi[1,2], Catherine S. Hartley[3], Roland N. Ndip[1,2], Lucy M. Ndip[1,2], Alistair C. Darby[4], Rory J. Post[5,6] and Benjamin L. Makepeace[2]*

[1] Laboratory for Emerging Infectious Diseases, University of Buea, PO Box 63, Buea, Cameroon
[2] Department of Microbiology and Parasitology, University of Buea, PO Box 63, Buea, Cameroon
[3] Institute of Infection & Global Health, University of Liverpool, Liverpool L3 5RF, UK
[4] Institute of Integrative Biology, University of Liverpool, Liverpool L69 7ZB, UK
[5] School of Natural Sciences & Psychology, Liverpool John Moores University, Liverpool L3 5UG, UK
[6] Disease Control Department, London School of Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT
* To whom correspondence should be addressed: blm1@liverpool.ac.uk

Abstract: Buruli ulcer (BU), caused by *Mycobacterium ulcerans*, is a neglected tropical disease associated with freshwater habitats. A variety of limnic organisms harbor this pathogen, including aquatic bugs (Hemiptera: Heteroptera), which have been hypothesized to be epidemiologically important reservoirs. Aquatic Hemiptera exhibit high levels of diversity in the tropics, but species identification remains challenging. In this study, we collected aquatic bugs from emerging foci of BU in the Southwest Region of Cameroon, which were identified using morphological and molecular methods. The bugs were screened for mycobacterial DNA and a selection of 20 mycobacteria-positive specimens from the families Gerridae and Veliidae were subjected to next-generation sequencing. Only one individual revealed putative *M. ulcerans* DNA, but all specimens contained sequences from the widespread alpha-proteobacterial symbiont, *Wolbachia*. Phylogenetic analysis placed the *Wolbachia* sequences into supergroups A, B and F. Circularized mitogenomes were obtained for seven gerrids and two veliids, the first from these families for the African continent. This study suggests that aquatic Hemiptera may have a minor role (if any) in the spread of BU in Southwest Cameroon. Our metagenomic analysis provides new insights into the incursion of *Wolbachia* into aquatic environments and generates valuable resources to aid molecular taxonomic studies of aquatic Hemiptera.

Keywords: Buruli ulcer; symbiosis; *Limnogonus*; pond skater; riffle bug; *Rhagovelia*; *Metrocoris*; *Trepobates*

1. Introduction

Mycobacterium ulcerans is a slow-growing environmental pathogen that infects the skin and subcutaneous tissues, causing Buruli ulcer (BU; also known as Bairnsdale/Daintree ulcer in Australia) [1]. BU is an emerging, treatable but neglected skin infection that manifests as slowly developing, unspecific indolent nodules, papules, or induration by oedema, which can progress to necrotizing skin ulcers [2]. BU treatment outcomes can be good if the condition is diagnosed early, while late diagnosis can lead to extreme patient suffering and severe complications necessitating surgery or even amputation of limb(s) [3,4]. Although the incubation period of *M. ulcerans* infection has not been clearly defined, it is thought to vary from weeks to several months, with a median of 4 - 5 months [5,6].

BU has been reported in over 35 countries [7], almost half of which are in Africa [2]. In Cameroon, the first cases of BU were diagnosed in 1969 in 47 patients residing in a confined area near the villages of Ayos and Akonlinga [8,9]. This was two decades after BU was first described in Australia [10] and some years after the first case report in Africa [11]. Although the number of BU cases in Cameroon has declined following active surveillance in line with recent global trends [12], it is still a major public health problem in this country. New foci of infection continue to emerge, while Bankim and Akonolinga remain as disease hot spots. From 2001 to 2014, the number of BU endemic health districts rose from two to 64 [12-15].

Mycobacterium ulcerans DNA has been detected in aquatic bugs (Hemiptera: Heteroptera) [16,17], other limnic organisms (*e.g.*, aquatic plants, fish, tadpoles and mosquitoes [18-20]) and environmental samples (freshwater and soil [21,22]) in several ecological studies, but its route(s) of transmission from these sources has not yet been fully established [23-25]. Emergence of BU has been linked to acquisition of a megaplasmid, pMUM, by *Mycobacterium marinum*, a far less virulent environmental pathogen associated with aquatic habitats [26]. The megaplasmid encodes for polyketide synthases and accessory enzymes for the biosynthesis of a macrolide toxin, mycolactone [27], which is responsible for tissue damage and local immune suppression in BU [28]. Phylogenomic evidence indicates that all mycolactone-producing mycobacteria (MPM) form a monophyletic group comprising three clonal lineages, which should be considered ecovars of *M. ulcerans* rather than distinct species (the so-called "*Mycobacterium liflandii*", "*Mycobacterium pseudoshottsii*" and "*Mycobacterium shinshuense*") [26]. It has been hypothesized that *M. ulcerans* can be transmitted by human-biting aquatic bugs, notably those in the families Naucoridae, Belostomatidae, Notonectidae and Nepidae [29-31]. Although it has been clearly established that *M. ulcerans* specifically colonizes the salivary glands of biting bugs and they can transmit the pathogen to laboratory rodents [16,31], their role as disease vectors in nature remains controversial. Indeed, it is highly unlikely they are directly responsible for the majority of BU cases [32,33]. Nevertheless, as most families of aquatic bugs contain species (or morphs) with flight ability [34], the colonization of aquatic bugs by *M. ulcerans* may be important in the spread of the pathogen into new areas.

Previous studies in central Cameroon have implicated aquatic bugs as reservoirs and/or vectors of *M. ulcerans*, with colonization of the insects being detected in a BU-endemic area but not a proximate non-endemic area [16]. The current study investigated aquatic bugs from the previously under-sampled Southwest Region of Cameroon for signatures of *M. ulcerans* in order to map environmental sources and potential reservoirs of the pathogen, as evaluating the presence of *M. ulcerans* in the environment is important in assessing the risks for human infection [23,34,35]. Precise species identification and delimitation of aquatic bugs in tropical regions, where species diversity is very high, can be challenging. Several authors have proposed the use of DNA barcoding targeting a standard region of the mitochondrial gene, cytochrome *c* oxidase subunit I (*coi*), in the taxonomy of true bugs [36,37]. Therefore, our study extended the previous use of DNA barcoding for identification of Cameroonian aquatic bugs [38] in a more focused geographical area where BU appears to be emerging.

Recent reports of a high prevalence of the α-proteobacterial symbiont, *Wolbachia*, in aquatic insects [39] may also be of epidemiological relevance. Although *Wolbachia* is the most widespread animal symbiont on Earth, being found in >50% of terrestrial arthropod species [40], the extent of its penetration into aquatic environments remains unclear. *Wolbachia* is known to affect the susceptibility of arthropod hosts to colonization by other microorganisms (either positively or negatively, depending on specific combinations of pathogen, vector and *Wolbachia* strain [41-43]) and can profoundly impede vector competence if artificially introduced into naïve hosts [44]. Therefore, interactions between *M. ulcerans* and *Wolbachia* in aquatic bugs have the potential to impact on the dissemination of the pathogen in the host and perhaps the wider environment.

Here, we present the results of an intensive environmental and taxonomic survey of aquatic bugs in the Southwest Region of Cameroon. Through a combination of traditional morphology-based taxonomy and DNA barcoding coupled with metagenomics, we provide new insights into the diversity of aquatic Hemiptera in Cameroon, acquire the first complete mitochondrial genomes from bugs for the country, and demonstrate that while infection with *M. ulcerans* is virtually absent in bugs from the Southwest, they harbor surprisingly diverse *Wolbachia* sequence types.

2. Materials and methods

2.1. Study sites

A cross-sectional descriptive study, in which study sites were selected using a multi-stage sampling approach, was carried out in the Southwest Region of Cameroon. This region was chosen because it is BU-endemic, but very little information on the disease (especially environmental sources of *M. ulcerans*) exists for this part of the country. Localities endemic for BU have been identified in the Southwest Region, leading to the establishment of a BU detection and treatment center; however, research and control efforts have so far been focused only in the three

 doi:10.20944/preprints201908.0031.v1

main BU-endemic foci (located in other administrative Regions): Akonlinga (Centre), Ayos (Centre) and Bankim (Adamawa) [13,45].

Six health districts (HDs) from a total of 18 in the Southwest Region were selected for the study based on the cumulative number of BU cases reported within a 14-year period (2001-2014) [12]. Two of the HDs (Mbonge and Ekondo-Titi) are referred to as mesoendemic foci in the present study because they reported cumulative BU cases of 75 and 37 respectively; values much lower than those reported by the three major endemic foci in Cameroon [Akonolinga (1081), Bankim (557) and Ayos (485)] within the same 14-year period [12]. The other four HDs (Limbe, Buea, Muyuka and Kumba) are referred to as hypoendemic foci because each reported <10 cumulative BU cases within the stated period.

An initial scoping exercise was undertaken in the towns and accessible villages in each selected HD to i) introduce the study to the local population and ii) map the sampling points (also termed as water bodies or sampling sites). Some villages were poorly accessible by road and could only be reached by trekking. In each locality, there was formal self-introduction to the traditional ruler followed by briefings on the study's goals. Sampling points were mapped with the assistance of the population. Almost all water bodies sampled were shallow, slow-flowing (except after heavy rain) or stagnant water bodies with overhanging vegetation that were used by community members for a variety of purposes (drinking, washing of clothes and/or vehicles, recreation, cooking, fishing and bathing) and were persistent throughout the year. An exception with regard to depth was Lake Barombi Mbo, a volcanic lake with an area of 5 km^2 and an average depth of 69 m, although only the shallows accessible by wading were sampled (Figure S1C). The other sampling sites varied to a lesser degree with respect to depth, size and exposure to sunlight (Figure S1A-D).

2.2. Sample collection and transportation

Aquatic bugs were collected from each of 25 water bodies (Figure 1) from 13th – 25th July 2017. The same sampling methods were used at each site and the sampling was carried out between 9:00 am and 3:00 pm. The aquatic bugs were captured by hauling a three-piece collapsible professional hand net (250 mm wide frame, 1 mm in mesh size and 300 mm deep) (EFE & GB Nets, Lostwithiel, UK) from the water surface to the bottom substrate for comprehensive sampling of specimens in the water column, the depth of which never exceeded ~1.3 m (wading height). Epineustonic bugs observed moving across the water surface were also sampled opportunistically. At each site, several sweeps with nets were performed within 40 min by a team of four workers and samples were placed into a white plastic tray (EFE & GB Nets) in order to remove extraneous vegetable matter and non-hemipteran animals. Putative hemipterans were transferred into a 50 mL polypropylene tube containing absolute acetone, kept in a cooler and transported to the Laboratory for Emerging Infectious Diseases, University of Buea, for short-term storage. Subsequently, the acetone was decanted and all samples were shipped to the University of Liverpool, UK, under an import license issued by the UK Animal & Plant Health Agency. The acetone was replaced and the insects stored at 4°C.

2.3. Aquatic bugs identification and morphotyping

Each bug was numbered uniquely for taxa identification. Using a stereomicroscope, taxonomic keys, written descriptions, specimen comparisons, image verification and an extensive literature search [25,38,46-65] the aquatic bugs were placed into families, genera, morphotypes and, wherever possible, species. Confirmation of the identity of selected specimens was done by morphological comparisons with material in the collections of the Natural History Museum, London.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 August 2019 doi:10.20944/preprints201908.0031.v1

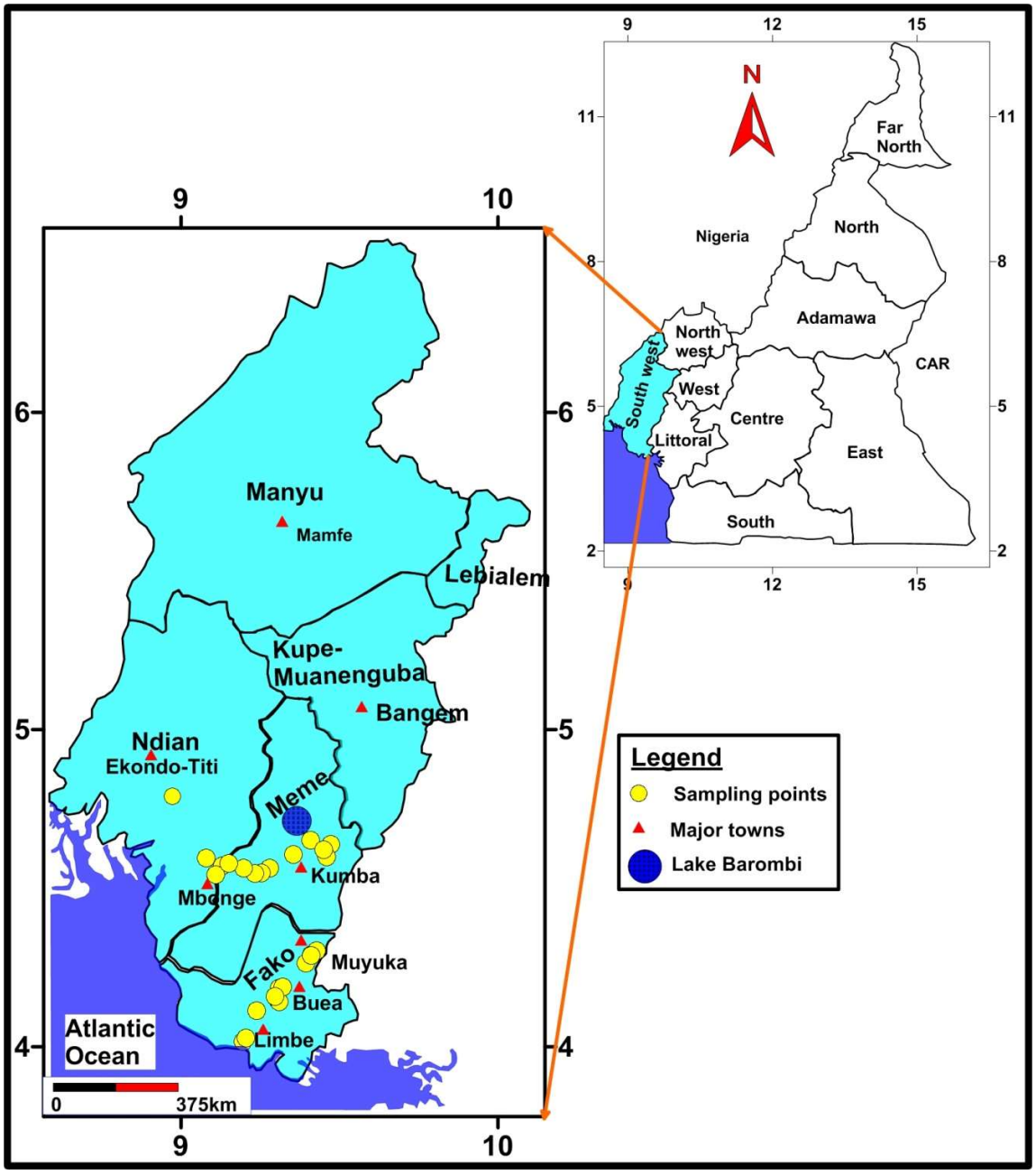

Figure 1. Sample collection sites in the Southwest Region of Cameroon.

2.4. DNA extraction from aquatic bugs

The bugs were rinsed with distilled water to remove acetone and each specimen was cut into several pieces with a separate sterile disposable scalpel (Swann-Morton, Sheffield, England) before transfer into a 1.5 mL snap cap tube. Ammonium hydroxide (150 µL, 1.5 M) (Sigma-Aldrich) was added to the tube and boiled in a heating block at 100°C for 15 min [66]. The tubes were centrifuged at 5,000 × g for 1 min and reheated at 100°C for 15 min with the lids open to evaporate ammonia and obtain a final volume of 70 – 100 µL. Insoluble material was removed after centrifugation at 10,000 × g for 10 min and discarded. The concentration of DNA was determined by Quant-iT PicoGreen dsDNA quantification assay using a microplate fluorimeter (Infinite F200, Tecan) and Magellan Data Analysis Software (Tecan). The DNA samples were stored at 4°C.

2.5. Coi amplification (DNA barcoding)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 August 2019 doi:10.20944/preprints201908.0031.v1

For DNA barcoding, the mitochondrial *coi* region was amplified. Unless stated otherwise, each conventional PCR run in this study had a final volume of 20 µL comprising 10 µL BioMix Red master-mix (2X), 1 µL each primer from a 10 µM working stock (final concentration, 0.5 µM), 1 µL DNA template and 7 µL nuclease free water. All PCR amplifications were conducted in a Biometra TRIO Thermal Cycler (Analytik, Jena, Germany) and electrophoretic separations of PCR products were on a 1% agarose gel stained with SYBR Safe DNA Gel stain (Invitrogen). The PCR products were visualized and photographed with a Safe Imager transilluminator (Invitrogen). A negative control, in which DNA template was replaced with nuclease-free water, was included in each PCR run.

The primers used for the amplification of the *coi* fragment (710 bp) were LCO1490: 5'-GGTCAACAAATCATAAAGATATTGG-3' (sense) and HC02198: 5'-TAAACTTCAGGGTGACCAAAAAATCA-3' (antisense) [67]. PCR amplifications comprised initial denaturation at 94°C for 5 min, followed by 35 cycles of denaturation at 94°C for 1 min, annealing at 50°C for 1.5 min and extension at 72 °C for 1 min. The final extension was at 72°C for 5 min and the reactions were then held at 4°C. Selected amplicons were directly sequenced in both directions by Macrogen (Seoul, South Korea) using the same primers used for PCR amplification.

2.6. Amplification of a mycobacterial rpoB gene fragment by quantitative PCR

Quantitative real-time PCR (qPCR) was used for rapid screening to detect and quantify mycobacterial DNA in a subset of aquatic bug specimens (see Results). The qPCR was designed to target a fragment (88 bp) of the RNA polymerase β-subunit (*rpoB*) gene that is 100% conserved in most non-tuberculous mycobacteria, including MPM. The following primers and probes were synthesized by Eurofins Genomics (Ebersberg, Germany): forward primer (Myco-F: 5′-GATCTCCGACGGTGACAAGC-3′), reverse primer (Myco-R: 5′-CAGGAACGGCATGTCCTCG-3′) and Myco-probe (5′-*6FAM*-ACGGCAACAAGGGCGTCATCGGCAAGATCCT-*BHQ-1*-3′). A 20 µL reaction mix was prepared containing 1 µL DNA template, 10 µL SYBR Green Master Mix (2X) (Bioline), and final concentrations of 0.5 µM for each primer and 0.25 µM for the probe in PCR-grade water. Amplification proceeded on a CFB-3220 DNA Engine Opticon 2 System (Bio-Rad) at 50°C for 3 min (holding), 95°C for 5 min (initial denaturation) followed by 40 cycles of 20 sec at 95°C and 40 sec at 60°C. To quantify the mycobacterial load of each sample, each run included a standard curve with DNA concentrations corresponding to 1,000,000 to 0.1 copies per reaction of a synthetic mycobacterial *rpoB* oligonucleotide standard (Eurofins Genomics), diluted in 100 ng/µl yeast tRNA (Invitrogen, Carlsbad, CA, USA) to prevent aggregation. Each DNA standard dilution and sample was assayed in duplicate. Opticon Monitor software (v. 3.1) was used for linear regression analysis based on tenfold dilutions of the standard. Each PCR run included a negative control (PCR-grade water) and a positive control (genomic DNA from *Mycobacterium ulcerans* NCTC 10417, Public Health England Culture Collections, Salisbury, UK).

The frequency of *rpoB*-positive bugs was compared between meso- and hypoendemic sites using a χ^2 analysis at www.socscistatistics.com.

2.7. Confirmatory conventional PCRs for bacteria

For specimens positive with the *rpoB* qPCR, a 400 bp fragment of the IS*2404* insertion sequence from MPM was amplified in a first-round PCR while the nested round amplified a 150 - 200 bp fragment. Both rounds of PCR were performed with primers as previously described [68]. The PCR reactions were prepared as described above for *coi*, except for the nested round in which 1 µL PCR product from the first-round PCR was used as a DNA template. The PCR amplifications were carried out under the following conditions: initial denaturation at 95 °C for 15 min followed by 25 cycles of denaturation at 94 °C for 30 sec, annealing at 64 °C for 1 min and extension at 72 °C for 1 min. The final extension was at 72 °C for 10 min. The cycling parameters for the nested round were identical to those of the first round, except the number of cycles was raised to 30.

To amplify *Wolbachia* surface protein (*wsp*), primers wsp-81F and wsp-691R [69] were used at a final concentration of 0.5 µM in 20-µl reactions with BioMix Red master-mix (see *coi* PCR above). The thermocycling conditions comprised 35 cycles at 94°C for 1 min, 55°C for 1 min and 72°C for 1 min.

2.8. Metagenomic sequencing

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 August 2019 doi:10.20944/preprints201908.0031.v1

Twenty gerrid and veliid specimens with the highest copy numbers or mycobacterial *rpoB* were subjected to next-generation sequencing on an Illumina HiSeq 4000 platform (Illumina, Inc., San Diego, CA, USA). Illumina fragment libraries were prepared from genomic DNA using the Nextera XT kit as previously described [70] and sequenced as paired-ends (2 × 150 bp) over two lanes, with a minimum of 280 million clusters per lane. Base-calling, de-multiplexing of indexed reads, and trimming of adapter sequences was conducted as reported previously [70]. Raw data were submitted to the Sequence Read Archive at the National Center for Biotechnology Information (NCBI) under BioProject PRJNA542604 and accession numbers SRX5884614 - SRX5884633.

2.9. Taxonomic assignment

Taxonomic read assignments were performed using *k*-mer based classifiers Kraken v. 2.0.7 [71] and CLARK-S v. 1.2.5 [72] with default settings. The output from Kraken2 or CLARK-S was visualized using the Krona v. 2.7 [73] tool, creating hierarchical interactive pie-charts. MetaPhlAn2 v. 2.6.0 [74] was also used to identify bacterial species and estimate their relative abundance across all 20 samples according to the database 'mpa_v20_m200'. Velvet v. 1.2.10 [75] wrapped by VelvetOptimiser v. 2.2.6 was used to assemble the 20 bug genomes to contig level. According to these assemblies, GC-coverage plots (proportion of GC bases and Velvet node coverage) were generated using BlobTools v. 1.0.1 [76]. Taxonomic assignment of these contigs was achieved using megablastn to search against the nr database.

2.10. Mitogenome assembly and annotation

The mitogenome of each sample was assembled *de novo* using NOVOPlasty v. 2.7.2 [77], with the *D. melanogaster* mitogenome as a seed sequence (NC_024511.2). Each circularized mitogenome assembly obtained in the first run was used iteratively as a seed sequence to run NOVOPlasty again. In all the assembly processes, the *k*-mer length was set to 39 and other operation parameters in the configuration file were left on the default setting or set to fit the features of the reads. Protein-coding, rRNA and tRNA genes were annotated with Mitos2 [78] by searching the Refseq 81 Metazoa database using the invertebrate mitochondrial genetic code. The protein-coding genes were manually added or edited, if necessary, after comparing with other hemipteran mitogenomes via tblastn. The tRNAs were identified with the tRNAscan-SE search server v. 2.0 [79] by setting sequence source to 'other mitochondrial' and genetic code for tRNA isotype prediction to 'invertebrate mito'. Mitogenome sequences were compared using BLAST Ring Image Generator v. 0.95 [80] to *Perittopus* sp. (JQ910988.1) [81], *Aquarius paludum* (NC_012841.1) [82] and *Gigantometra gigas* (NC_041084.1) [83] reference sequences. Gene rearrangement analysis of the assembled mitogenomes were visualized using the Mauve genome aligner v. 2015-02-25 [84]. MrBayes v. 3.2.6 [85] was run with the generation parameter set to '1 in 100 samples' and the substitution model set to 'GTR+G+I' (the first 25% of samples were discarded as burn-in) to build a mitogenome tree based on protein super-alignments as previously described [86]. The assembled mitogenomes were submitted to NCBI under accession numbers MN027271 - MN027279.

2.11. Barcode analysis of coi sequences and phylogenetics

All Sanger-sequenced *coi* reads were trimmed of ~50 poor-quality bases at the 5′ and 3′ end before assembly using the cap3 [87] program. The assemblies were then submitted to the Barcode of Life Data (BOLD) System [88] v. 4 under the project name "CBUG" and analyzed using the "Cluster Sequences" feature, which calculates pairwise distance following alignment with an amino acid-based hidden Markov model. Gaps were handled by pairwise deletion and the sequences were binned into operational taxonomic units (OTUs) by nearest-neighbor distance. The *coi* sequences were compared with reference gerrid, veliid and mesoveliid sequences from NCBI and from a previous Cameroonian barcoding study of aquatic bugs [38], in which the data were available in their supplementary information but not BOLD Systems or NCBI. Alignments were performed using Mafft v. 7.402 [89] with the '--auto' option and automatically trimmed with Gblocks v. 0.91b [90]. A phylogenetic tree for *coi* sequences was reconstructed using IQ-TREE v. 1.6.7 [91] with the parameters '-m MFP -alrt 1000 -bnni -bb 1000' to determine the best-fit model for ultrafast bootstrap, and both Shimodaira–Hasegawa-like (SH)-like approximate likelihood ratio tests and ultrafast bootstrapping were conducted to access branch supports within one single run.

2.12. Analysis of Wolbachia genes and phylogenetics

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 August 2019 doi:10.20944/preprints201908.0031.v1

Wolbachia reads assigned by Kraken2 were further assembled and scaffolded using SPAdes-3.7.1 [92] with the '--careful' function. *Wolbachia* protein-coding genes in these assemblies were predicted using Prokka v1.13 [93] with default settings. One-to-one orthologs of *Wolbachia* protein-coding genes in five reference *Wolbachia* strains (GCF_000008025.1 GCF_000073005.1, GCF_000306885.1, GCF_000008385.1, GCF_001931755.2 and GCF_000829315.1), each from a different taxonomic "supergroup" (higher-level clade), were identified by using OrthoFinder 2.2.3 [94]. To investigate the *Wolbachia* supergroups in our samples, concatenated phylogenetic trees were constructed for multiple orthologs (depending on the recovery of *Wolbachia* protein-coding genes in each assembly) at the amino-acid level using IQ-TREE as described above. A phylogenetic analysis for the *Wolbachia wsp* gene (trimmed PCR amplicons) was also performed using MrBayes with the same parameters as for the whole mitogenome tree. Velvet-assembled contigs classified as *Wolbachia* bacteriophage (phage WO) were aligned onto a complete phage genome (KX522565.1) from *w*VitA, a symbiont of the parasitoid wasp *Nasonia vitripennis* [95], to identify the presence of phage in sequenced bug samples.

3. Results

3.1. Distribution of aquatic bugs in the Health Districts

Of the 1,113 aquatic insects collected in the Southwest Region, 1,102 were identified as true bugs (Hemiptera) belonging to eight families and 29 putative genera. Specimens in each genus were further separated into morphotypes, generating a total of 110 (of which 34 were represented by only one specimen each). Subsequently, 331 individual specimens were retained for species confirmation and voucher specimens while the remaining 771 were subjected to molecular analysis. Of the 25 sampling points, 17 (68 %) were in hypoendemic zones while the remaing eight (32%) were in mesoendemic areas. Similarly, of the 1,102 aquatic bugs captured, 714 (64.8%) were from the hypoendemic sites and 388 (35.2%) from the mesoendemic sites. The distribution of the aquatic bugs by HD is shown in Figure 2.

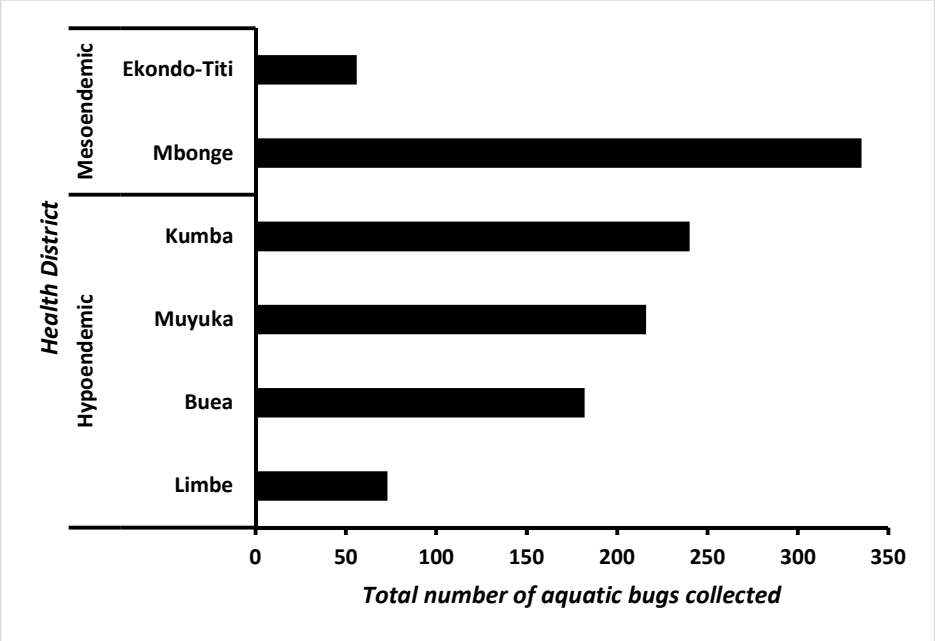

Figure 2. The distribution of the aquatic bugs collected in the Health Districts.

3.2. Composition of aquatic bug taxa

Most bugs collected belonged to just two families, with 498 (45.2%) from the Gerridae (water striders) and 426 (38.7%) from the Veliidae (riffle bugs). The Notonectidae, Naucoridae, Belostomatidae, Mesoveliidae, Hydrometridae and Nepidae were represented by 54 (4.9 %), 52 (4.7 %), 36 (3.3 %), 18 (1.6 %), 11 (1.0 %) and 7 (0.6

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 August 2019 doi:10.20944/preprints201908.0031.v1

%) specimens, respectively (Figure 3). Overall, the population of aquatic bugs originating from families known to bite humans was very low (149/1,102, 15.5%) and distributed in the HDs according to Table S1.

Visual inspection of the morphotypes indicated that Gerridae were dominated by *Limnogonus* spp. and the Veliidae by *Rhagovelia* spp. (Table 1 and Figure 4). Overall, the specimens were allocated to 42 putative species, some containing multiple morphotypes (Table 1). The family Gerridae was the most diverse, with 18 putative species, followed by Naucoridae (8 putative species), Notonectidae (6 putative species), Veliidae (4 putative species), Belostomatidae (3 putative species) and one putative species each for the families Mesoveliidae, Hydrometridae and Nepidae. The Gerridae, Veliidae and Hydrometridae were present in all HDs (Table S2); indeed, six sampling points (three in Buea HD and one each in Muyuka, Kumba and Mbonge) had only gerrids and veliids.

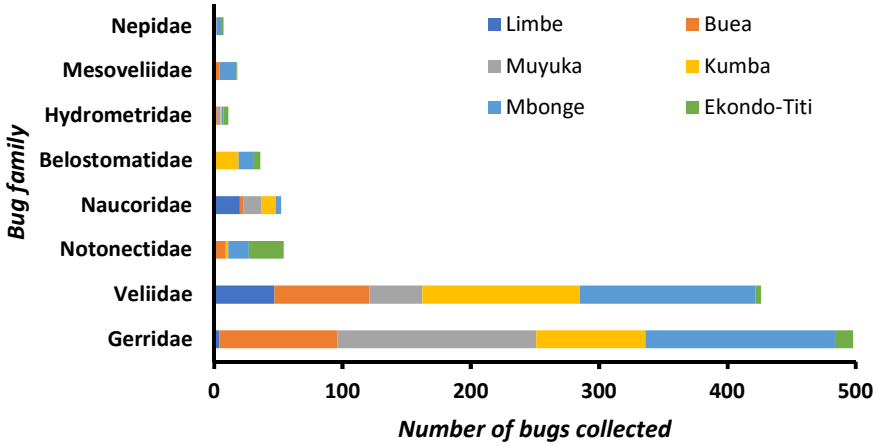

Figure 3. Aquatic bug families identified in this study and their distribution.

Table 1. Composition of aquatic bugs based on morphological characteristics and confirmation at the Natural History Museum, London.

Family (*n*)	Genus	Putative species	Total	No. of morphotypes
Gerridae (498)	*Limnogonus*	*Limnogonus (L.) curriei* Bergroth 1916	2	1
		Limnogonus (L.) *guttatus* Poisson 1948	45	1
		Limnogonus (L.) hypoleucus Gerstaecker 1873	4	1
		Limnogonus (L.) *intermedius* Poisson 1941	57	1
		Limnogonus (L.) *poissoni* Andersen 1975	83	1
		Limnogonus (*s. str.*) *cereinventris* Signoret 1862	30	1
		Limnogonus (*L.*) spp.	41	7
		Limnogonus (*s. str.*) sp.	1	1
	Aquarius	*Aquarius remigis*	3	1
		Aquarius spp.	6	3
	Tenagogonus	*Tenagogonus olbovitlotus*	6	1
		Tenagogonus spp.	8	4
	Trepobates	*Trepobates* spp.	89	13
	Metrobates	*Metrobates* spp.	55	10
	Metrocoris	*Metrocoris* spp.	16	4
	Neogerris	*Neogerris severini*	1	1

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 August 2019 doi:10.20944/preprints201908.0031.v1

	Rhagadotarsus	*Rhagadotarsus* sp.	1	1
	Eurymetra	*Eurymetra* sp.	1	1
	Eurymetropsis	*Eurymetropsis* sp.	50	1
Veliidae (426)	*Rhagovelia*	*Rhagovelia reitteri* Reuters, 1882	184	1
		Rhagovelia spp.	195	9
	Microvelia	*Microvelia* spp.	47	6
	Walambianisops	*Walambianisops* spp.	20	5
	Anisops	*Anisops* spp.	11	3
Notonectidae (54)	*Paranisops*	*Paranisops* sp.	1	1
	Enithares	*Enithares* sp.	5	2
	Notonecta	*Notonecta* sp.	2	1
	Nychia	*Nychia* sp.	15	4
	Aneurocoris	*Aneurocoris* sp.	25	1
	Illyocoris	*Illyocoris* sp.	6	2
	Laccocoris	*Laccocoris* sp.	5	1
Naucoridae (52)	*Maccrocolis*	*Maccrocolis laticollis*	2	1
	Naucoris	*Naucoris obscuratus*	2	1
		Naucoris sp.	9	3
	Neomaccrocoris	*Neomaccrocoris parviceps ocellatus*	2	1
		Neomaccrocoris parviceps parviceps	1	1
Belostomatidae (36)	*Belostoma*	*Belostoma* sp.	1	1
	Diplonychus	*Diplonychus* sp.	34	6
	Lethocerus	*Lethocerus* sp.	1	1
Mesoveliidae (18)	*Mesovelia*	*Mesovelia* sp.	18	3
Hydrometridae (11)	*Hydrometra*	*Hydrometra* sp.	11	1
Nepidae (7)	*Ranatra*	*Ranatra* sp.	7	1
	Total		**1,102**	**110**

Not all human-biting bugs (Table S1) were analyzed for the detection of *M. ulcerans* due to the high intraspecific morphological polymorphism observed. For example, the 54 notonectids, placed in six putative species, comprised 16 morphotypes with the most abundant morphotype having only five specimens. Similarly, bugs in the families Mesoveliidae (18, 1.6 %) and Hydrometridae (11, 1.0 %) were also not analyzed for M. ulcerans carriage because of their small numbers (Table 1).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 August 2019 doi:10.20944/preprints201908.0031.v1

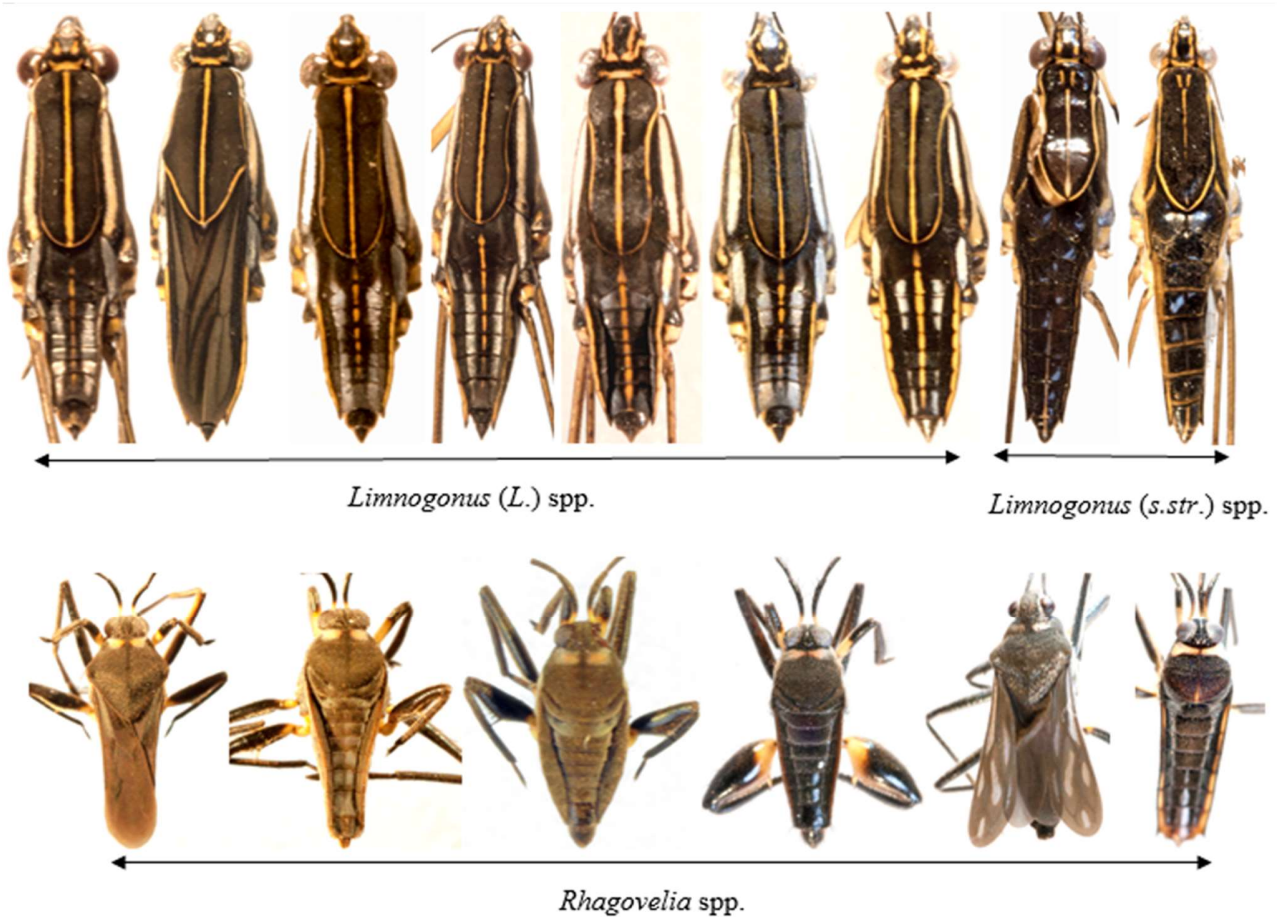

Figure 4. Representative specimens of the most common genera of aquatic bugs identified in this study: *Limnogonus* (Gerridae) and *Rhagovelia* (Veliidae).

3.3. Prevalence of mycobacterial DNA in the aquatic bugs

The mycobacterial *rpoB* gene was detected and quantified in 8.9% of the 771 specimens, comprising 13.2% (53/405) from Gerridae and 4.4% (16/366) from Veliidae. The distribution of the positive specimens among the putative species is detailed in Table 2. The positive specimens comprised 7.9% (21/266) from mesoendemic sites and 9.5% (48/505) from hypoendemic sites (p = 0.457). None of the specimens positive for the mycobacterial *rpoB* gene were positive for the *M. ulcerans*-specific IS*2404* sequence, although this region was successfully amplified from the positive control.

Table 2. Detection and quantification of *rpoB* gene fragment in aquatic bugs from the study sites.

Family (*n*)	Genus	Putative species	Total analyzed	*rpoB* positive (%)	Mean *rpo*B copy no.
Gerridae (405)	*Limnogonus*	*Limnogonus* (*L.*) *guttatus* Poisson 1948	39	7 (17.9)	63
		Limnogonus (*L.*) *poissoni* Andersen 1975	76	15 (19.7)	143
		Limnogonus (*L.*) *intermedius* Poisson 1941	50	7 (14)	82
		Limnogonus (*L.*) *curriei* Bergroth 1916	2	0 (0.0)	0
		Limnogonus (*s. str.*) *cereinventris* Signoret 1862	24	7 (29.2)	80

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 August 2019 doi:10.20944/preprints201908.0031.v1

		Limnogonus (*L.*) *hypoleucus* Gerstaecker 1873	2	1 (50)	58
		Limnogonus spp.	34	0 (0.0)	0
	Aquarius	*Aquarius* spp.	5	0 (0.0)	0
	Tenagogonus	*Tenagogonus* spp.	8	2 (25.0)	13
	Metrobates	*Metrobates* spp.	42	5 (11.9)	47
	Trepobates	*Trepobates* spp.	70	4 (5.7)	50
	Eurymetropsis	*Eurymetropsis* sp.	46	4 (8.7)	112
	Metrocoris	*Metrocoris* spp.	7	1 (14.3)	91
Veliidae	*Rhagovelia*	*Rhagovelia reitteri* Reuters 1882	171	3 (1.8)	44
(366)		*Rhagovelia* spp.	164	11 (6.7)	111
	Microvelia	*Microvelia* spp.	31	2 (6.5)	33
	Total		**771**	**69 (8.9)**	

3.4. Bacterial sequences in gerrid and veliid metagenomic datasets

As both epineustonic and nektonic bugs have been reported to have high infection rates with *M. ulcerans* in other BU-endemic regions of Cameroon (see Discussion), we applied next-generation genomic sequencing to the 20 bug specimens with the highest *rpoB* copy numbers to confirm that they were infected only with non-MPM species. Approximately 11 – 27 million reads were obtained per specimen (Table 3) and taxon-annotated GC-coverage plots ("blob analysis") revealed that several of the samples contained apparent bacterial sequences at relatively high coverage (Figure S2). In most samples, the CLARK-S tool identified ~10-fold fewer bacterial *k*-mers than did Kraken2 (data not shown), so was not used for downstream analyses of bacterial sequences. *K*-mers identified as originating from *Mycobacterium* spp. by Kraken2 exhibited low abundance (<0.5% of all bacterial read-pairs) in all specimens except for a *Metrocoris* sp. (Gerridae) from Muyuka, in which mycobacterial sequences comprised almost 5% of all bacterial read-pairs (Table 3). However, 98% of these mycobacterial *k*-mers were classified as deriving from *Mycobacterium* sp. WY10, a non-MPM soil-associated species [96], with no sequences classified as *M. ulcerans* in this specimen. Only one specimen, a *Limnogonus hypoleucus* from Limbe, contained *M. ulcerans* sequences at a proportion >0.1% of all bacterial sequences (84 read-pairs), accounting for ~50% of mycobacterial sequences (Table 3).

 doi:10.20944/preprints201908.0031.v1

315 **Table 3.** Distribution of sequences classified as originating from *Mycobacterium* spp. and *Wolbachia* in 20 aquatic bug genomes as determined by Kraken2.

Sample	Health District	Location Code	Species	Total reads	All bacterial read pairs (% total)	*Mycobacterium* read pairs (% bacterial)	*M. ulcerans* read pairs (% bacterial)	*Wolbachia* read pairs (% bacterial)	*wsp* PCR
1	Buea	B05	*Tenagogonus* sp. 3	14,767,852	58,114 (0.39%)	82 (0.14%)	0 (0.00%)	125 (0.22%)	-
2	Buea	B01	*Limnogonus guttatus*	13,155,711	75,982 (0.58%)	62 (0.08%)	1 (0.00%)	17,278 (22.74%)	+
3	Buea	B02	*Limnogonus cereiventris*	27,361,770	161,821 (0.59%)	53 (0.03%)	1 (0.00%)	30,423 (18.80%)	+
4	Buea	B04	*Limnogonus cereiventris*	12,617,275	57,802 (0.46%)	18 (0.03%)	0 (0.00%)	13,120 (22.70%)	+
5	Buea	B02	*Microvelia* sp. 1	15,199,889	156,452 (1.03%)	35 (0.02%)	0 (0.00%)	6,017 (3.85%)	+
6	Kumba	K02	*Metrobates* sp. 7	11,913,069	61,477 (0.52%)	73 (0.12%)	1 (0.00%)	29 (0.05%)	-
7	Kumba	K02	*Metrobates* sp. 6	11,847,970	62,737 (0.53%)	57 (0.09%)	0 (0.00%)	70 (0.11%)	-
8	Kumba	K02	*Metrobates* sp. 8	12,711,110	66,881 (0.53%)	68 (0.10%)	0 (0.00%)	32 (0.05%)	-
9	Kumba	K01	*Rhagovelia reitteri*	12,615,049	168,912 (1.34%)	148 (0.09%)	0 (0.00%)	106,556 (63.08%)	+
10	Limbe	L02	*Limnogonus hypoleucus*	13,728,598	69,491 (0.51%)	172 (0.25%)	89 (0.13%)	7,470 (10.75%)	+
11	Mbonge	Mb03	*Limnogonus poissoni*	13,897,927	87,297 (0.63%)	67 (0.08%)	0 (0.00%)	24,629 (28.21%)	+
12	Mbonge	Mb03	*Limnogonus poissoni*	13,479,480	64,224 (0.48%)	75 (0.12%)	0 (0.00%)	13,345 (20.78%)	+
13	Mbonge	Mb02	*Limnogonus intermedius*	15,587,895	113,606 (0.73%)	61 (0.05%)	0 (0.00%)	9,359 (8.24%)	+
14	Mbonge	Mb04	*Limnogonus poissoni*	14,300,937	82,433 (0.58%)	57 (0.07%)	0 (0.00%)	13,412 (16.27%)	+
15	Mbonge	Mb06	*Rhagovelia* sp. 2	14,197,736	129,978 (0.92%)	60 (0.05%)	0 (0.00%)	66,515 (51.17%)	+
16	Muyuka	Mu02	*Metrocoris* sp. 1	14,136,653	72,422 (0.51%)	3,598 (4.97%)*	0 (0.00%)	142 (0.20%)	-
17	Muyuka	Mu05	*Eurymetropsis* sp.	14,455,526	62,105 (0.43%)	90 (0.14%)	0 (0.00%)	48 (0.08%)	-
18	Muyuka	Mu01	*Trepobates* sp. 5	13,054,268	57,197 (0.44%)	37 (0.06%)	0 (0.00%)	5,326 (9.31%)	-
19	Muyuka	Mu05	*Rhagovelia* sp. 1	15,150,283	138,234 (0.91%)	59 (0.04%)	0 (0.00%)	5,186 (3.75%)	-
20	Muyuka	Mu05	*Rhagovelia* sp. 1	14,978,846	131,153 (0.88%)	100 (0.08%)	0 (0.00%)	65,354 (49.83%)	+

316 *98% of these reads were classified as deriving from *Mycobacterium* sp. WY10.

 doi:10.20944/preprints201908.0031.v1

Strikingly, read-pairs classified as *Wolbachia* were more abundant than mycobacterial sequences in 16/20 specimens, and in many cases, the differential in read-pair count between the two bacterial genera was 2 – 3 logs (Table 3). However, *Wolbachia* read-pairs accounted for a very wide range of bacterial sequences (0.05 – 63.08%) between specimens. Using a conventional PCR targeting *wsp*, we were able to amplify this *Wolbachia*-specific gene from all 12 specimens with >6,000 *Wolbachia* read-pairs, but from none of the samples with a lower *Wolbachia* DNA content (Table 3).

In order to classify the *Wolbachia* strains in the bug specimens, we applied three approaches. First, Sanger sequencing of the *wsp* amplicons and a Bayesian phylogenetic analysis suggested that most *Wolbachia* infections belonged to supergroup B, represented in the tree by strain *w*Pip (Figure 5). Two specimens apparently deviated from this pattern, with a supergroup A sequence related to *w*Mel being amplified from *Rhagovelia* sp. 2, while a more divergent *wsp* sequence was recovered from *Microvelia* sp. 1, which was nested between representative sequences from supergroups F and D (Figure 5). Second, as *wsp* is known to recombine [97], especially between supergroups A and B that co-infect numerous arthropod hosts, we also built trees from protein-coding *Wolbachia* orthologues recovered from the metagenomic datasets. These supported the supergroup assignments suggested by the *wsp* analysis in all cases, with the *Wolbachia* sequences from *Microvelia* sp. 1 indicating a placement in supergroup F (Figure S3). Third, the metagenomic phylogenetic analysis tool MetaPhlAn2 identified matches with supergroup B *Wolbachia* genomes in almost all samples with amplifiable *wsp*, although a strong signal assigned to a supergroup A genome (*w*MelPop) was apparent in the *Rhagovelia* sp. 2 specimen (Figure 6). For the *Microvelia* sp. 1 sample, *Wolbachia* reads were detected but not classified as being associated with a specific sequenced strain. The classification of the *Wolbachia* reads by MetaPhlAn2 suggested that more than one strain might be present in several samples, especially the *Rhagovelia* spp (Figure 6). Finally, it was noteworthy that sequences matching the bacteriophage of *Wolbachia* (phage WO), which is known to carry the genes responsible for *Wolbachia*-mediated reproductive manipulations in arthropods [98], were identified on numerous contigs from the *Rhagovelia* sp. 2 specimen (Figure S4).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 August 2019 doi:10.20944/preprints201908.0031.v1

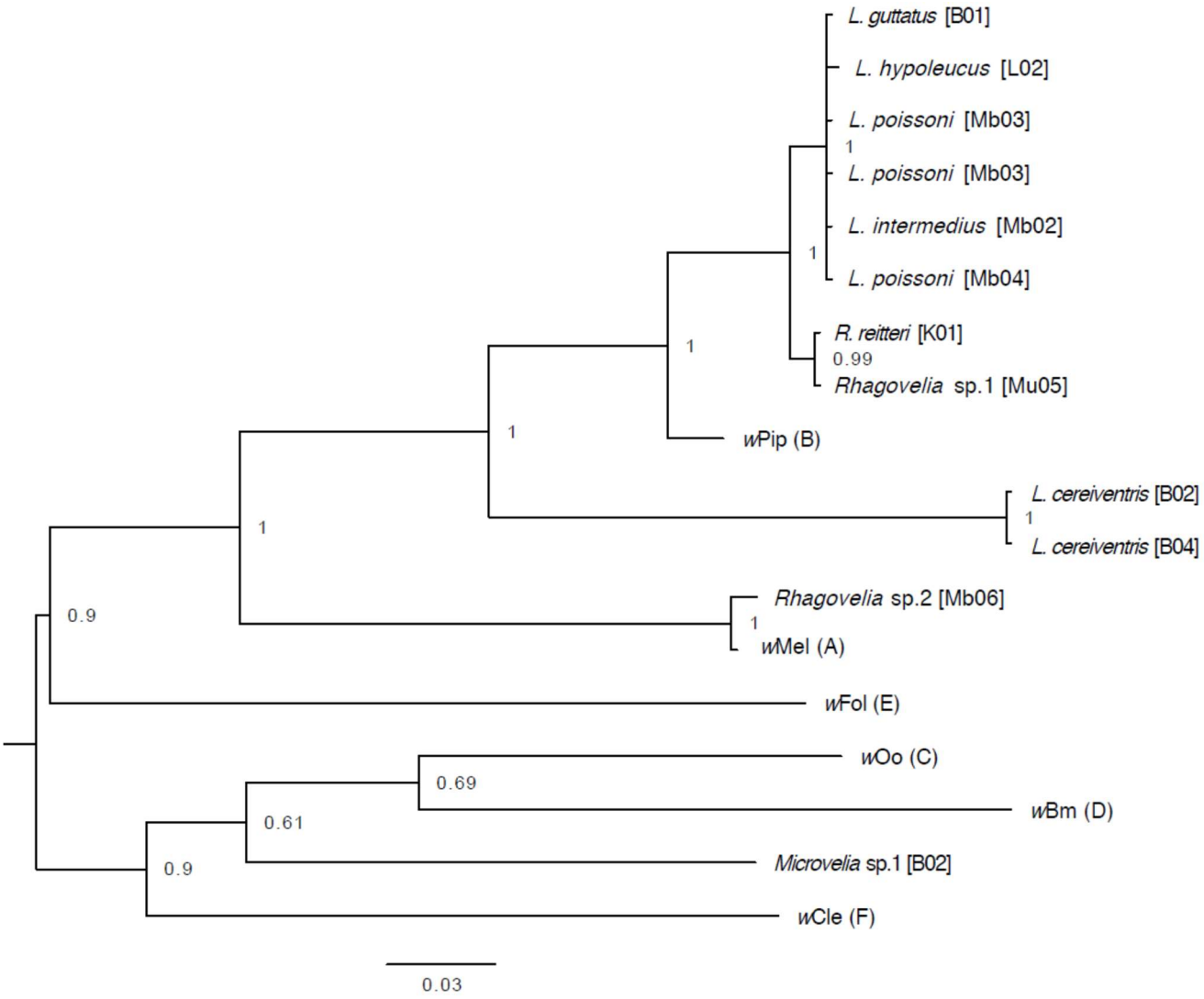

Figure 5. Bayesian tree of *Wolbachia* surface protein gene sequences from 12 aquatic bug specimens compared with those representing five major *Wolbachia* supergroups: *w*Mel (from *Drosophila melanogaster*), *w*Pip (from *Culex pipiens*), *w*Oo (from *Onchocerca ochengi*), *w*Bm (from *Brugia malayi*), *w*Fol (from *Folsomia candida*) and *w*Cle (from *Cimex lectularius*). Letters in parentheses after *Wolbachia* strain names refer to supergroup designations; suffixes after bug names in square brackets refer to location codes (see Table 3).

3.5. Assembly of complete mitogenomes from gerrids and veliids and DNA barcoding

Of the 20 bug specimens subjected to next-generation sequencing, it was possible to assemble circularized mitogenomes in nine cases, of which only two were from veliids. Annotation of these genomes showed a gene order in both bug families that was typical of insects (Figure S5). The veliid mitogenomes appeared to be approximately 100 bp shorter than those from gerrids (Figure S5, Figure 7). However, due to a well-recognized limitation of Illumina sequencing technology [99], it was apparent that we achieved only low coverage of the control region (AT-rich region; see breaks in Figure 7 at ~15 kbp) compared with three available references from gerrids and veliids, especially for Metrocoris and Trepobates. Although putative differences in this region between species must be confirmed using alternative sequencing technologies, it appears to be potentially useful for discriminating between Limnogonus spp. (Figure 7A, B).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 August 2019 doi:10.20944/preprints201908.0031.v1

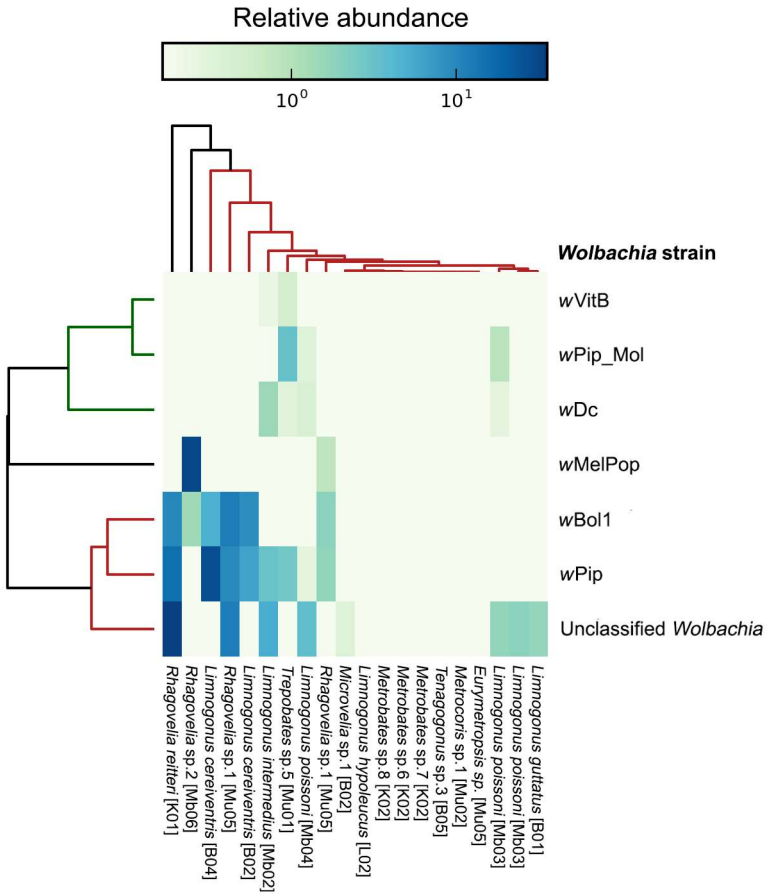

Figure 6. Metagenomic phylogenetic analysis of *Wolbachia* sequences from 20 aquatic bug genomes as determined by MetaPhlAn2 [74]. *Wolbachia* reference genomes are from symbionts of *N. vitripennis* (*w*VitB), *Culex pipiens molestus* (*w*Pip_Mol), *Diaphorina citri* (*w*Dc), *D. melanogaster* (*w*MelPop), *Hypolimnas bolina* (*w*Bol1), and *Culex pipiens pipiens* (*w*Pip).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 August 2019 doi:10.20944/preprints201908.0031.v1

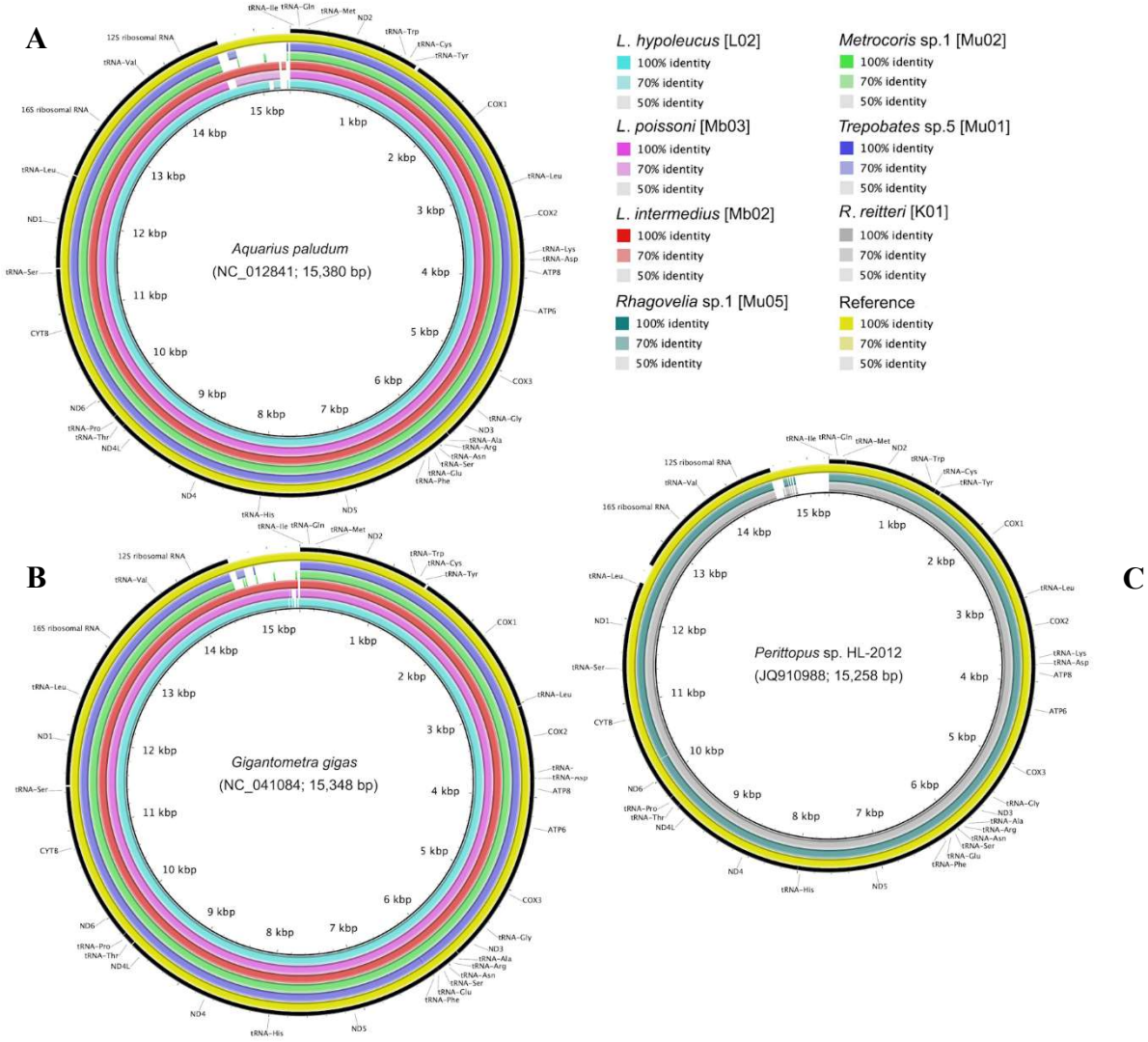

Figure 7. BLAST Ring Image Generator [80] plots for circularized mitogenomes of Gerridae (A, B) and Veliidae (C). Details of reference genomes from NCBI are provided in the center of each map.

To evaluate the concordance between morphological identifications and DNA barcodes, 50 gerrid, veliid, and mesoveliid samples that produced bright *coi* amplicons by PCR were sequenced, and 43 samples generated unambiguous alignments of forward and reverse reads. These were supplemented by an additional nine *coi* sequences obtained from the Illumina sequencing effort, producing a total of 52 sequences across 20 morphotypes. Twenty-one Sanger-sequenced samples were flagged as containing stop codons by BOLD Systems but were retained for downstream analyses. Sequence clustering by BOLD Systems produced 19 putative OTUs. A phylogenetic analysis of the *coi* sequences from gerrids, veliids and mesoveliids incorporating sequences from the study of Ebong et al. and top BLASTn hits from NCBI revealed incomplete resolution of bug families [*e.g.*, note the position of "V5" veliid sequences from Ebong et al. and references for *Stridulivelia* spp., *Microvelia americana* and *Perittopus horvathi* (all Veliidae) on the tree; Figure 8]. However, clustering of gerrid genera from our study (*Trepobates*, *Eurymetropis*, *Metrocoris*, *Metrobates* and *Limnogonus*) and often species (*L. hypoleucus*, *L. cereiventris*, and to a large extent, *L. intermedius*) was clearly evident. The inclusion of barcodes containing stop codons frequently, but not always, generated spurious OTUs containing a single specimen in the BOLD cluster analysis as highlighted previously (Figure 8)

 doi:10.20944/preprints201908.0031.v1

[100]. This had a comparatively greater impact on the analysis of the smaller veliid dataset, although a more likely explanation for the aberrant placement of our *Mesovelia* specimens is specimen misidentification, since support for their positioning within a clade of *Rhagovelia* spp. was strong (Figure 8). In summary, the 20 morphotypes were collapsed into 12 high-confidence OTUs once those that contained only sequences containing stop codons (seven OTUs) were disqualified.

With the exception of the "V5" sequences from the Ebong *et al*. (2016) analysis [38] and our "*Mesovelia*" specimens, clustering at the family level in both studies was consistent (Figure 8). A phylogenetic analysis at the whole mitogenome level also clearly separated gerrid and veliid specimens (including references from NCBI) into family-specific groups, while *Limnogonus* spp. were resolved as a monophyletic clade with unambiguous segregation by species (Figure 9). The "V4" sequences from Ebong *et al*. [38] appear to belong to our *Rhagovelia* spp. OTU 5, and the "Ge6" sequences identified as *L. hypoleucus* in Ebong *et al*. [38] were positioned proximal to our own *L. hypoleucus* (OTU 10) (Figure 8). However, we did not identify any *Angilia* spp. ("V1") in our study (Figure 8). Additional comparisons between the two works were not possible, as Ebong *et al*. [38] did not provide putative identifications for several sequence clusters and their data are not held in BOLD Systems. Nevertheless, it is striking that despite the much greater geographic coverage of the Ebong *et al*. study [38], extending across most of Cameroon's administrative Regions, they reported the initial identification of five gerrid genera, of which only three (all observed by us) were corroborated by DNA barcoding in their analysis. In our study located within one Region, eight gerrid genera were recorded, of which five were selected for DNA barcoding and received strong phylogenetic support. Four of these genera (*Trepobates*, *Eurymetropis*, *Metrocoris*, and *Metrobates*) were not observed in the Ebong *et al*. study [38].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 August 2019 doi:10.20944/preprints201908.0031.v1

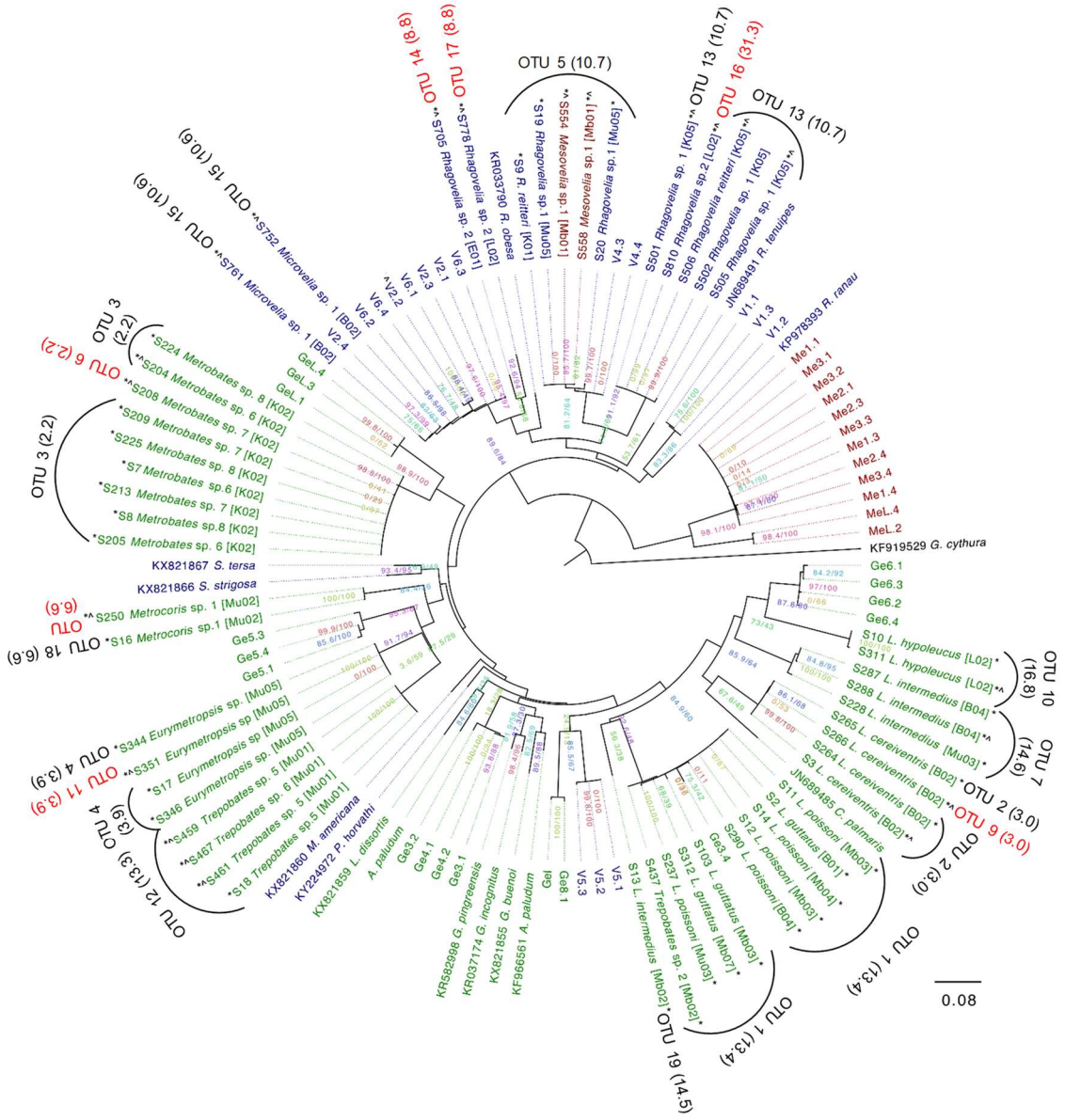

Figure 8. (previous page). Phylogenetic tree of *coi* sequences from gerrids (green), veliids (blue) and mesoveliids (dark red). Sequences from the current study are marked with an asterisk (*) and show location codes in square brackets, references from NCBI show accession numbers, and sequences from the prior study of Ebong *et al*. [38] are represented by short alphanumeric codes. The "Cluster Sequences" feature at BOLD Systems [88] was used to generate the OTUs and nearest-neighbor distances (in parentheses). Dubious OTUs (labelled in red text) are associated with stop codons in *coi* sequences, indicated by a caret symbol (^) after the putative species name. Numbers on the nodes of the tree are in the format "SH-like approximate likelihood ratio test support (%)/ultrafast bootstrap support (%)".

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 August 2019 doi:10.20944/preprints201908.0031.v1

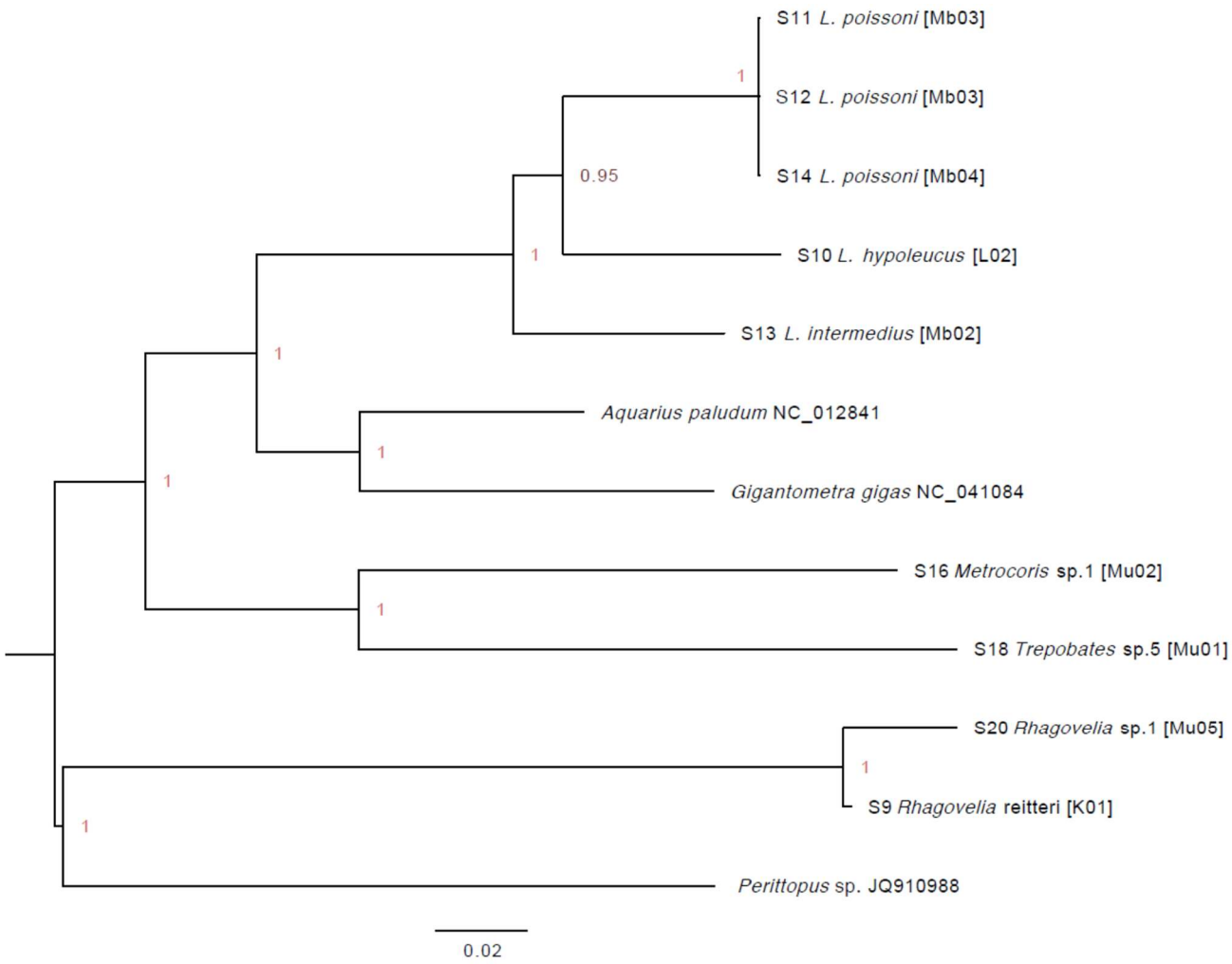

Figure 9. Bayesian tree of whole mitogenomes from nine Cameroonian aquatic bugs (Gerridae: *Limnogonus* spp., *Metrocoris* sp., *Trepobates* sp.; Veliidae; *Rhagovelia* spp.) plus references from NCBI (Gerridae: *Aquarius paludum*, *Gigantometra gigas*; Veliidae, *Perittopus* sp.). The tree was constructed using protein-coding sequences only.

4. Discussion

Despite raised awareness of BU, reflected by increased public health concern and intensive basic research efforts on *M. ulcerans* since the late 1990s, transmission of the pathogen and its maintenance in the environment remain poorly understood. The role of aquatic Hempitera in the transmission of *M. ulcerans* was first hypothesized in 1999 [29]. Subsequent studies demonstrated that *M. ulcerans* has a predilection for the salivary glands of biting aquatic bugs (most studies have focused on Naucoridae and Belostomatidae; [16,101]) and can transmit the pathogen to laboratory rodents [31]. It has also been demonstrated that laboratory mice can develop BU lesions through contamination of miniscule puncture injuries, including mosquito bites, with viable *M. ulcerans* [102]. This suggests that *M. ulcerans* in the aquatic environment primarily infects humans via existing minor skin trauma but could be inoculated mechanically by insect vectors through either infected arthropod saliva or transfer of *M. ulcerans* on the skin surface into the bite wound. However, except perhaps for sporadic cases [32], evidence for a vector role for biting aquatic bugs (or mosquitoes [103]) in BU transmission in nature is weak [23].

A second potential role for aquatic Hemiptera in the epidemiology of BU is as maintenance hosts or reservoirs that might serve to amplify *M. ulcerans* populations in aquatic environments and/or spread the pathogen between water sources, especially as most aquatic Hemiptera are capable of

 doi:10.20944/preprints201908.0031.v1

flight. Hemiptera are one among many groups of aquatic organisms that can harbor *M. ulcerans*, with other limnic invertebrates, fish, amphibians, and biofilms on aquatic plants all contributing to the distribution of the pathogen in freshwater bodies and waterlogged soils [18-20,22,33,35]. In a very large survey involving screening of >200,000 pooled specimens of aquatic animals for *M. ulcerans* in two BU-endemic areas of Cameroon, Hemiptera did not have the highest prevalence rate of infection but was the only taxon shown to be positive almost all year round [104]. Also in Cameroon, *M. ulcerans* was detected in pools of aquatic Hemiptera at rates exceeding 30% in a BU-endemic area, but was absent in bugs from a non-endemic zone 100 km away on the same river (the Nyong [16]). Countering these findings, a collection of 22,832 invertebrates in Ghana failed to find any association between *M. ulcerans* prevalence in hemipterans and BU endemicity between several villages [33]. A subsequent metanalysis of the Ghanaian data identified oligochaete worms rather than Hemiptera as key hosts in aquatic transmission networks [105], although the distribution of *M. ulcerans* was later correlated more generally with the presence of benthic detritivores and filter–feeders rather than any one taxonomic group [106].

The findings of our study contrast markedly with others from Cameroon in unearthing little evidence for the presence of *M. ulcerans* in aquatic bugs. We sought to screen the bug specimens rapidly with a general mycobacterial *rpoB* qPCR assay that allowed us to assess bacterial load, not only qualitative presence, for downstream next-generation sequencing, which was originally intended to provide data on the environmental strain(s) of *M. ulcerans* in the Southwest Region. However, the fact that the *rpoB*-positive specimens were negative with the nested IS*2404* assay, which targets a highly-repeated insertion sequence [26], strongly suggested that the bugs contained non-MPM DNA. This lack of *M. ulcerans* was confirmed in most cases following whole genome sequencing of 20 bug specimens with relatively high *rpoB* copy numbers. Although ~80 read-pairs were classified as originating from *M. ulcerans* in one gerrid from a hypoendemic site (Limbe), this extremely low signal (not confirmed by the IS*2404* assay) could originate from degraded *M. ulcerans* DNA; for instance, from gut contents or nucleic acids bound to the insect cuticle. Nevertheless, it is suggestive of the presence of *M. ulcerans* at that site.

There are several possible reasons for the discrepancy between our findings and previous surveys of *M. ulcerans* in aquatic bugs in Cameroon. The most obvious is that our study was much smaller than those of Marion *et al*. [16] and Garchitorena *et al*. [104] in terms of the total number of bugs collected, and we assayed insects individually rather than in small pools. The previous studies were also undertaken in highly endemic areas for BU, with >10-fold more recorded human cases than even our "mesoendemic" sites. The influence of seasonality is also potentially important, although in areas of high endemicity, the positivity rate for *M. ulcerans* has been reported to be greatest in July – August [16,104], which was a similar period to our field survey in the Southwest. Following the publication of the hypothesis that biting bugs are vectors of *M. ulcerans* [29], subsequent studies have tended to focus on biting taxa at the expense of non-biting bugs, whereas we paid greater attention to gerrids and veliids, as they were much more abundant at our sampling sites. While a recent ecological analysis linked infection of aquatic Hemiptera with *M. ulcerans* to the nektonic lifestyle of predatory species that feed on animals their own size or even larger, occupy aquatic vegetation and live towards the bottom of the water column [34], *M. ulcerans* infection of gerrids has been reported from Cameroon previously at rates similar to other bug taxa [16]. Moreover, the pathogen has been isolated from a gerrid from Benin [17].

The sequencing of the bug specimens in our study provided an opportunity to investigate other components of the microbiome in these relatively little-studied insects. In this small sample, *Wolbachia* was found not only to be the dominant symbiont but also to be represented by multiple sequence types. While *Wolbachia* is ubiquitous in arthropods from terrestrial environments, very few data on its distribution in aquatic species existed until recently, apart from analyses of holometabolous, medically-important vectors with aquatic larval stages (principally mosquitoes [107], but also blackflies [108,109] and some limited surveys in aquatic crustaceans [110,111]). However, a 2017 metanalysis of *Wolbachia* prevalence in aquatic insects (n = 228 species) estimated that >50% of species are infected, with Hemiptera exhibiting a particularly high prevalence (69%)

 doi:10.20944/preprints201908.0031.v1

[39]. Importantly, assessing *Wolbachia* infection by molecular methods alone (including metagenomics) leads to potential pitfalls in interpretation, primarily due to high rates of lateral gene transfer from *Wolbachia* into the host genome, which can lead to *Wolbachia* relics or "genomic fossils" in species that may have lost the infection [112-114]. Host genomes containing sequences from more than one supergroup have even been reported [115]. Therefore, our putative *Wolbachia* infections in Cameroonian gerrids and veliids require confirmation by microscopy for whole bacteria or deeper sequencing to recover closed bacterial genomes.

Whether all of the *Wolbachia* infections in the aquatic bugs are extant or not, in our very small sample of fully-sequenced specimens, evidence for three *Wolbachia* supergroups was apparent. Supergroups A and B are the most prevalent clades in arthropods worldwide, forming the basis of the so-called *Wolbachia* "pandemic" [116,117]. Coinfection by both clades in a single host has been described repeatedly [118,119], with some evidence for this in our own specimens. The presence of supergroup F-like sequences in *Microvelia* sp. 1 was a more unusual finding, as this clade is less widespread than the A and B supergroups, although it is also the only one known from both arthropod and nematode hosts [120,121]. As the *Wolbachia* signal in this specimen was relatively low (~6,000 read-pairs, or 0.04% of the total), the possibility that the clade F sequences originate from gut contents or an arthropod or nematode parasite or parasitoid cannot be discounted at this stage [122,123]. A priority for future research on association of aquatic bugs with *M. ulcerans* will be to determine if *Wolbachia* influences this relationship and whether the presence of the symbiont might explain the limited evidence of *M. ulcerans* infection in our specimens compared with other locations in Cameroon. Notably, while most research on pathogen protection by *Wolbachia* in arthropods has focused on artificial infections for vector-borne disease control [44], *Wolbachia* has also been reported to influence susceptibility to pathogens in some natural systems [42,43].

A final contribution of our study was to extend the integrative taxonomy of aquatic Hemiptera in Cameroon following a previous published work [38]. Although the Ebong *et al*. study [38] included a much greater number of specimens used for DNA barcoding (188) than our own from across the whole country, our intensive field survey in the Southwest produced a similar number of barcodes to Ebong *et al.* [38] for the Gerridae and Veliidae, and they were selected from a larger sample size for these taxa. We obtained barcodes from four gerrid genera not observed in the Ebong et al. [38] study (*Trepobates, Eurymetropis, Metrocoris*, and *Metrobates*) and complete mitogenomes for *Trepobates*, *Metrocoris* and *Limnogonus*; the first to be published worldwide. The barcoding effort showed some limitations, particularly with respect to the amplification of putative nuclear mitochondrial transfers (which accumulate mutations) and/or the impact of high rates of heteroplasmy [100]. However, with stringent interpretation of barcode segregation, biases caused by these artifacts can be minimized, and the effects of mixing Sanger and next-generation sequencing technologies should also be considered carefully [83].

We noted that separation of aquatic bug families by *coi* sequences was not perfect in all cases, especially when considering reference sequences deposited in NCBI. Misidentifications are an obvious explanation for this discrepancy, and probably also explain the aberrant placement of *Mesovelia* spp. in our study. Unfortunately, although previous studies using a larger number of characters (*e.g*., whole mitogenomes) have been published for Hemiptera, representation of aquatic taxa has been patchy, with no inclusion of mesoveliid specimens for instance [81,82]. While our whole mitogenome analysis for the small number of gerrid and veliid assemblies available separated the two families unambiguously, future efforts should focus on obtaining complete mitogenomes from mesoveliids too.

At the genus and species level, the *coi* barcoding proved its usefulness for gerrids in particular, but sampling of a more diverse range of genera will be required before its utility with veliid specimens can be firmly established. A very large barcoding effort for Hemiptera found that the minimum interspecific distance was >3% for 77% of congeneric species-pairs, although a case of identical barcodes between different genera of mirid bugs was uncovered [37]. An important consideration in species infected with symbionts that can cause cytoplasmic incompatibility, such as *Wolbachia*, is introgression of mitochondrial haplotypes following symbiont-driven selective sweeps

[124]. In some cases, this introgression can penetrate across hybrid zones during breakdown of reproductive barriers [125-127]. In this respect, it is noteworthy that a phylogenetic analysis of *Limnogonus* spp. using two mitochondrial and one nuclear marker found the two subgenera (*Limnogonus sensu stricto* and the exclusively Afrotropical *Limnogonoides*) to be paraphyletic, as well as weak bootstrap support for clades and species within *Limnogonoides* [128]. Although *L. guttatus* was not included in that analysis, it is thought to be closely related to *L. hypoleucus* and *L. poissoni* [128], and to the extent to which these species might be able to hybridize in nature, potentially sharing *Wolbachia* strains, should be explored in future.

Acknowledgments: This work was funded by the Medical Research Foundation, UK, under an Africa Research Excellence Fund (AREF) Research Development Fellowship Grant (identifier ESEMUMRF-157-0010-F-ESEMU) awarded to S.N.E. We thank Alan Bannister (University of Liverpool) for assistance with the photography of the bugs (Figure 4).

Author contributions: Conceptualization, S.N.E, R.J.P., L.M.N., A.C.D., B.L.M.; Methodology, S.N.E., X.D., R.J.P., A.C.D., B.L.M.; Validation, S.N.E., R.J.P., X.D., B.L.M.; Formal Analysis, S.N.E., X.D., B.L.M.; Investigation, S.N.E., A.J.K., C.S.H., R.J.P., B.L.M.; Resources, S.N.E., A.C.D., L.M.N., R.N.N., B.L.M.; Data Curation, S.N.E., X.D., B.L.M.; Writing – Original Draft Preparation, S.N.E., X.D., B.L.M.; Writing – Review & Editing, S.N.E., A.C.D., R.J.P., X.D., B.L.M.; Visualization, S.N.E., X.D., B.L.M.; Supervision, S.N.E., L.M.N., R.N.N., R.J.P., B.L.M.; Project Administration, S.N.E., L.M.N., R.M.N., B.L.M.; Funding Acquisition, S.N.E., L.M.N., B.L.M.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Supplementary materials. The following are available online in the main supplement file (PDF). Figure S1: Representative sampling points, Figure S2: Representative blob-plot from sample 15 (*Rhagovelia* sp. 2 from Mbonge), Figure S3: Concatenated tree based on 26 protein-coding *Wolbachia* orthologs for sample 5, Figure S4: Contigs from sample 15 (*Rhagovelia* sp. 2 from Mbonge) mapping to a reference WO phage genome, Figure S5: Alignment of two veliid (*Rhagovelia* spp.) and seven gerrid (*Limnogonus, Metrocoris* and *Trepobates* spp.) mitogenomes, Table S1: The distribution of human-biting bugs captured in the Health Districts. Table S2: Distribution of aquatic Hemiptera in the Southwest Region of Cameroon (provided as an Excel table).

References

1. Tai, A.Y.C.; Athan, E.; Friedman, N.D.; Hughes, A.; Walton, A.; O'Brien, D.P. Increased severity and spread of mycobacterium ulcerans, southeastern australia. *Emerg Infect Dis* **2018**, *24*.
2. Röltgen, K.; Pluschke, G. Epidemiology and disease burden of buruli ulcer: A review. *Research and Reports in Tropical Medicine* **2015**, *6*, 59–73.
3. Toutous Trellu, L.; Nkemenang, P.; Comte, E.; Ehounou, G.; Atangana, P.; Mboua, D.J.; Rusch, B.; Njih Tabah, E.; Etard, J.F.; Mueller, Y.K. Differential diagnosis of skin ulcers in a mycobacterium ulcerans endemic area: Data from a prospective study in cameroon. *PLoS Negl Trop Dis* **2016**, *10*, e0004385.
4. Jacobsen, K.H.; Padgett, J.J. Risk factors for mycobacterium ulcerans infection. *Int J Infect Dis* **2010**, *14*, e677-681.
5. Loftus, M.J.; Trubiano, J.A.; Tay, E.L.; Lavender, C.J.; Globan, M.; Fyfe, J.A.M.; Johnson, P.D.R. The incubation period of buruli ulcer (mycobacterium ulcerans infection) in victoria, australia - remains similar despite changing geographic distribution of disease. *PLoS Negl Trop Dis* **2018**, *12*, e0006323.
6. Trubiano, J.A.; Lavender, C.J.; Fyfe, J.A.; Bittmann, S.; Johnson, P.D. The incubation period of buruli ulcer (mycobacterium ulcerans infection). *PLoS Negl Trop Dis* **2013**, *7*, e2463.
7. Bretzel, G.; Beissner, M. Pcr detection of mycobacterium ulcerans-significance for clinical practice and epidemiology. *Expert Rev Mol Diagn* **2018**, *18*, 1063-1074.
8. Pouillot, R.; Matias, G.; Wondje, C.M.; Portaels, F.; Valin, N.; Ngos, F.; Njikap, A.; Marsollier, L.; Fontanet, A.; Eyangoh, S. Risk factors for buruli ulcer: A case control study in cameroon. *PLoS Negl Trop Dis* **2007**, *1*, e101.

9. Ravisse, P. [skin ulcer caused by mycobacterium ulcerans in cameroon. I. Clinical, epidemiological and histological study]. *Bull Soc Pathol Exot Filiales* **1977**, *70*, 109-124.
10. P., M.; Tolhurst, J.C.; Buckle, G.; Sissons, H.A. A new mycobacterial infection in man. *J Pathol Bacteriol* **1948**, *60*, 93-122.
11. Dodge, O.G.; Lunn, H.F. Buruli ulcer: A mycobacterial skin ulcer in a uganda child. *J Trop Med Hyg* **1962**, *65*, 139-142.
12. Tabah, E.N.; Nsagha, D.S.; Bissek, A.C.; Njamnshi, A.K.; Bratschi, M.W.; Pluschke, G.; Um Boock, A. Buruli ulcer in cameroon: The development and impact of the national control programme. *PLoS Negl Trop Dis* **2016**, *10*, e0004224.
13. Akoachere, J.F.; Nsai, F.S.; Ndip, R.N. A community based study on the mode of transmission, prevention and treatment of buruli ulcers in southwest cameroon: Knowledge, attitude and practices. *Plos One* **2016**, *11*, e0156463.
14. Marion, E.; Landier, J.; Boisier, P.; Marsollier, L.; Fontanet, A.; Le Gall, P.; Aubry, J.; Djeunga, N.; Umboock, A.; Eyangoh, S. Geographic expansion of buruli ulcer disease, cameroon. *Emerg Infect Dis* **2011**, *17*, 551-553.
15. Noeske, J.; Kuaban, C.; Rondini, S.; Sorlin, P.; Ciaffi, L.; Mbuagbaw, J.; Portaels, F.; Pluschke, G. Buruli ulcer disease in cameroon rediscovered. *Am J Trop Med Hyg* **2004**, *70*, 520-526.
16. Marion, E.; Eyangoh, S.; Yeramian, E.; Doannio, J.; Landier, J.; Aubry, J.; Fontanet, A.; Rogier, C.; Cassisa, V.; Cottin, J., *et al.* 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* **2010**, *4*, e731.
17. Portaels, F.; Meyers, W.M.; Ablordey, A.; Castro, A.G.; Chemlal, K.; de Rijk, P.; Elsen, P.; Fissette, K.; Fraga, A.G.; Lee, R., *et al.* First cultivation and characterization of mycobacterium ulcerans from the environment. *PLoS Negl Trop Dis* **2008**, *2*, e178.
18. Marsollier, L.; Stinear, T.; Aubry, J.; Saint Andre, J.P.; Robert, R.; Legras, P.; Manceau, A.L.; Audrain, C.; Bourdon, S.; Kouakou, H., *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.
19. Willson, S.J.; Kaufman, M.G.; Merritt, R.W.; Williamson, H.R.; Malakauskas, D.M.; Benbow, M.E. Fish and amphibians as potential reservoirs of mycobacterium ulcerans, the causative agent of buruli ulcer disease. *Infect Ecol Epidemiol* **2013**, *3*.
20. Johnson, P.D.; Azuolas, J.; Lavender, C.J.; Wishart, E.; Stinear, T.P.; Hayman, J.A.; Brown, L.; Jenkin, G.A.; Fyfe, J.A. Mycobacterium ulcerans in mosquitoes captured during outbreak of buruli ulcer, southeastern australia. *Emerg Infect Dis* **2007**, *13*, 1653-1660.
21. Ross, B.C.; Johnson, P.D.; Oppedisano, F.; Marino, L.; Sievers, A.; Stinear, T.; Hayman, J.A.; Veitch, M.G.; Robins-Browne, R.M. Detection of mycobacterium ulcerans in environmental samples during an outbreak of ulcerative disease. *Appl Environ Microbiol* **1997**, *63*, 4135-4138.
22. Tian, R.B.; Niamke, S.; Tissot-Dupont, H.; Drancourt, M. Detection of mycobacterium ulcerans DNA in the environment, ivory coast. *Plos One* **2016**, *11*, e0151567.
23. Garchitorena, A.; Ngonghala, C.N.; Texier, G.; Landier, J.; Eyangoh, S.; Bonds, M.H.; Guegan, J.F.; Roche, B. Environmental transmission of mycobacterium ulcerans drives dynamics of buruli ulcer in endemic regions of cameroon. *Sci Rep* **2015**, *5*, 18055.
24. Ohtsuka, M.; Kikuchi, N.; Yamamoto, T.; Suzutani, T.; Nakanaga, K.; Suzuki, K.; Ishii, N. Buruli ulcer caused by mycobacterium ulcerans subsp shinshuense: A rare case of familial concurrent occurrence and detection of insertion sequence 2404 in japan. *JAMA Dermatol* **2014**, *150*, 64-67.
25. Ebong, S.M.; Eyangoh, S.; Marion, E.; Landier, J.; Marsollier, L.; Guegan, J.F.; Legall, P. Survey of water bugs in bankim, a new buruli ulcer endemic area in cameroon. *J Trop Med* **2012**, *2012*, 123843.
26. Doig, K.D.; Holt, K.E.; Fyfe, J.A.; Lavender, C.J.; Eddyani, M.; Portaels, F.; Yeboah-Manu, D.; Pluschke, G.; Seemann, T.; Stinear, T.P. On the origin of mycobacterium ulcerans, the causative agent of buruli ulcer. *BMC Genomics* **2012**, *13*, 258.
27. George, K.M.; Chatterjee, D.; Gunawardana, G.; Welty, D.; Hayman, J.; Lee, R.; Small, P.L. Mycolactone: A polyketide toxin from mycobacterium ulcerans required for virulence. *Science* **1999**, *283*, 854-857.
28. Fraga, A.G.; Cruz, A.; Martins, T.G.; Torrado, E.; Saraiva, M.; Pereira, D.R.; Meyers, W.M.; Portaels, F.; Silva, M.T.; Castro, A.G., *et al.* Mycobacterium ulcerans triggers t-cell immunity followed by local and regional but not systemic immunosuppression. *Infect Immun* **2011**, *79*, 421-430.
29. Portaels, F.; Elsen, P.; Guimaraes-Peres, A.; Fonteyne, P.A.; Meyers, W.M. Insects in the transmission of mycobacterium ulcerans infection. *Lancet* **1999**, *353*, 986.
30. Mosi, L.; Williamson, H.; Wallace, J.R.; Merritt, R.W.; Small, P.L. Persistent association of mycobacterium ulcerans with west african predaceous insects of the family belostomatidae. *Appl Environ Microbiol* **2008**, *74*, 7036-7042.

31. Marsollier, L.; Robert, R.; Aubry, J.; Saint Andre, J.P.; Kouakou, H.; Legras, P.; Manceau, A.L.; Mahaza, C.; Carbonnelle, B. Aquatic insects as a vector for mycobacterium ulcerans. *Appl Environ Microbiol* **2002**, *68*, 4623-4628.
32. Marion, E.; Chauty, A.; Yeramian, E.; Babonneau, J.; Kempf, M.; Marsollier, L. A case of guilt by association: Water bug bite incriminated in m. Ulcerans infection. *Int J Mycobacteriol* **2014**, *3*, 158-161.
33. Benbow, M.E.; Williamson, H.; Kimbirauskas, R.; McIntosh, M.D.; Kolar, R.; Quaye, C.; Akpabey, F.; Boakye, D.; Small, P.; Merritt, R.W. Aquatic invertebrates as unlikely vectors of buruli ulcer disease. *Emerg Infect Dis* **2008**, *14*, 1247-1254.
34. Ebong, S.M.A.; Garcia-Pena, G.E.; Pluot-Sigwalt, D.; Marsollier, L.; Le Gall, P.; Eyangoh, S.; Guegan, J.F. Ecology and feeding habits drive infection of water bugs with mycobacterium ulcerans. *Ecohealth* **2017**, *14*, 329-341.
35. Garchitorena, A.; Guegan, J.F.; Leger, L.; Eyangoh, S.; Marsollier, L.; Roche, B. Mycobacterium ulcerans dynamics in aquatic ecosystems are driven by a complex interplay of abiotic and biotic factors. *Elife* **2015**, *4*.
36. Jung, S.; Duwal, R.K.; Lee, S. Coi barcoding of true bugs (insecta, heteroptera). *Molecular Ecology Resources* **2011**, *11*, 266-270.
37. Park, D.S.; Foottit, R.; Maw, E.; Hebert, P.D.N. Barcoding bugs: DNA-based identification of the true bugs (insecta: Hemiptera: Heteroptera). *Plos One* **2011**, *6*.
38. Ebong, S.M.A.; Petit, E.; Le Gall, P.; Chen, P.P.; Nieser, N.; Guilbert, E.; Njiokou, F.; Marsollier, L.; Guegan, J.F.; Pluot-Sigwalt, D., *et al.* Molecular species delimitation and morphology of aquatic and sub-aquatic bugs (heteroptera) in cameroon. *Plos One* **2016**, *11*.
39. Sazama, E.J.; Bosch, M.J.; Shouldis, C.S.; Ouellette, S.P.; Wesner, J.S. Incidence of wolbachia in aquatic insects. *Ecol Evol* **2017**, *7*, 1165-1169.
40. Weinert, L.A.; Araujo-Jnr, E.V.; Ahmed, M.Z.; Welch, J.J. The incidence of bacterial endosymbionts in terrestrial arthropods. *Proc Biol Sci* **2015**, *282*, 20150249.
41. Teixeira, L.; Ferreira, A.; Ashburner, M. The bacterial symbiont wolbachia induces resistance to rna viral infections in drosophila melanogaster. *PLoS Biol* **2008**, *6*, e2.
42. Zele, F.; Nicot, A.; Berthomieu, A.; Weill, M.; Duron, O.; Rivero, A. Wolbachia increases susceptibility to plasmodium infection in a natural system. *Proc Biol Sci* **2014**, *281*, 20132837.
43. Graham, R.I.; Grzywacz, D.; Mushobozi, W.L.; Wilson, K. Wolbachia in a major african crop pest increases susceptibility to viral disease rather than protects. *Ecol Lett* **2012**, *15*, 993-1000.
44. Kamtchum-Tatuene, J.; Makepeace, B.L.; Benjamin, L.; Baylis, M.; Solomon, T. The potential role of wolbachia in controlling the transmission of emerging human arboviral infections. *Curr Opin Infect Dis* **2017**, *30*, 108-116.
45. Bolz, M.; Bratschi, M.W.; Kerber, S.; Minyem, J.C.; Um Boock, A.; Vogel, M.; Bayi, P.F.; Junghanss, T.; Brites, D.; Harris, S.R., *et al.* Locally confined clonal complexes of mycobacterium ulcerans in two buruli ulcer endemic regions of cameroon. *PLoS Negl Trop Dis* **2015**, *9*, e0003802.
46. Tchakonté, S.; Ajeagah, G.A.; Tchatcho, N.L.N.; Camara, A.I.; Diomandé, D.; Ngassam, P. Stream's water quality and description of some aquatic species of coleoptera and hemiptera (insecta) in littoral region of cameroon. *Biodiversity Journal* **2015**, *6*, 27–40.
47. Mbogho, A.Y.; Sites, R.W. Naucoridae leach, 1815 (hemiptera: Heteroptera) of tanzania. *Afr Invertebr* **2013**, *54*, 513-542.
48. Chen, P.P.; Nieser, N.; Zettel, H. *The aquatic and semiaquatic bugs (heteroptera - nepomorpha & gerromorpha) of malesia*. Brill: Leiden, 2005; p 546.
49. Andersen, N.M.; Weir, T.A. *Australian water bugs. Their biology and identification (hemiptera-heteroptera, gerromorpha & nepomorpha)*. Apollo Books, CSIRO Publishing: Collingwood, Victoria, 2004; Vol. 14, p 344.
50. Andersen, N.M. *The semiaquatic bugs (hemiptera, gerromorpha). Phylogeny, adaptations, biogeography, and classification*. Scandinavian Science Press Ltd.: Ganløse, Denmark, 1982; Vol. 3, p 454.
51. De Moor, I.J.; Day, J.A.; De Moor, F.C. *Insecta ii. Hemiptera, megaloptera, neuroptera, trichoptera and lepidoptera*. Water Research Commission: Pretoria, 2003; Vol. 8, p 220.
52. Chen, P.P.; Zettel, H. Five new species of the halobatinae genus metrocoris mayr, 1865 (insecta: Hemiptera: Gerridae) from continental asia. *Annalen des Naturhistorischen Museums in Wien, Serie B* **1999**, *101*, 13-32.
53. Hecher, C.; Zettel, H. Faunistical and morphological notes on limnogonus subgenus limnogonoides poisson 1965 (heteroptera: Gerridae). *Linzer biologische Beiträge* **1996**, *28*, 325-333.
54. Schuh, R.T.; Slater, J.A. *True bugs of the world (hemiptera: Heteroptera) classification and natural history*. Cornell University Press: Ithaca, 1995; p 416.
55. Dethier, M. Introduction à la morphologie, la biologie et la classification des hétéroptères. *Bulletin Romand d'Entomologie* **1981**, *1*, 11-16.

56. Poisson, R.A. Contribution à l'étude des hydrocorises de madagascar (mission j. Millot et r. Paulian 1949). *Mémoires de l'Institut Scientifique de Madagascar* **1952**, *1*, 24-70.
57. Poisson, R.A. Contribution à la faune du cameroun hémiptères aquatiques. *Faune des colonies françaises* **1929**, *3*, 135-164.
58. Poisson, R.A. De voyage ml. Chopard en côte d'ivoire (1938–39)—hémiptères aquatiques. *Rev Fr Entomol* **1941**, *8*.
59. Poisson, R.A. Hydrocorises du cameroun. Mission j. Carayon 1947. *Rev Fr Entomol* **1948**, *3*, 167-177.
60. Poisson, R.A. Voyage de m. P. P. Grasse en afrique occidentale française. Hémiptères aquatiques. *Ann Soc Entomol Fr* **1937**, *106*.
61. Poisson, R.A. Quelques hydrocorises nouveaux de l'afrique du sud (mission suédoise brink et rüdebeck). *Bulletin de la Société Scientifique de Bretagne* **1955**, *30*, 129-140.
62. Poisson, R.A. Contribution a l'etude des hydrocorises de madagascar (heteroptera). 4e memoire. . *Mémoires de l'Institut Scientifique de Madagascar* **1956**, *7*, 243-265.
63. Poisson, R.A. Catalogue des hétéroptères hydrocorises africano-malgaches de la famille des nepidae (ltreille) 1802. *Bulletin de l' IFAN Tome* **1965**, *27*, 229–269.
64. Poisson, R.A. Catalogue des insectes hétéroptères gerridae leach, 1807 africano-malgaches. *Bulletin de l' IFAN Tome* **1965**, *27*, 1466-1503.
65. Poisson, R.A. Catalogue des insectes hétéroptères notonectidae leach, 1815 africano-malgaches. *Bulletin de l' IFAN Tome* **1966**, *28*, 729–768.
66. Ammazzalorso, A.D.; Zolnik, C.P.; Daniels, T.J.; Kolokotronis, S.O. To beat or not to beat a tick: Comparison of DNA extraction methods for ticks (ixodes scapularis). *PeerJ* **2015**, *3*, e1147.
67. Folmer, O.; Black, M.; Hoeh, W.; Lutz, R.; Vrijenhoek, R. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit i from diverse metazoan invertebrates. *Mol Mar Biol Biotechnol* **1994**, *3*, 294-299.
68. Ablordey, A.; Amissah, D.A.; Aboagye, I.F.; Hatano, B.; Yamazaki, T.; Sata, T.; Ishikawa, K.; Katano, H. Detection of mycobacterium ulcerans by the loop mediated isothermal amplification method. *PLoS Negl Trop Dis* **2012**, *6*, e1590.
69. Zhou, W.; Rousset, F.; O'Neil, S. Phylogeny and pcr-based classification of wolbachia strains using wsp gene sequences. *Proc Biol Sci* **1998**, *265*, 509-515.
70. Dong, X.F.; Armstrong, S.D.; Xia, D.; Makepeace, B.L.; Darby, A.C.; Kadowaki, T. Draft genome of the honey bee ectoparasitic mite, tropilaelaps mercedesae, is shaped by the parasitic life history. *Gigascience* **2017**, *6*.
71. Wood, D.E.; Salzberg, S.L. Kraken: Ultrafast metagenomic sequence classification using exact alignments. *Genome Biol* **2014**, *15*, R46.
72. Ounit, R.; Lonardi, S. Higher classification sensitivity of short metagenomic reads with clark-s. *Bioinformatics* **2016**, *32*, 3823-3825.
73. Ondov, B.D.; Bergman, N.H.; Phillippy, A.M. Interactive metagenomic visualization in a web browser. *BMC Bioinformatics* **2011**, *12*, 385.
74. Truong, D.T.; Franzosa, E.A.; Tickle, T.L.; Scholz, M.; Weingart, G.; Pasolli, E.; Tett, A.; Huttenhower, C.; Segata, N. Metaphlan2 for enhanced metagenomic taxonomic profiling. *Nat Methods* **2015**, *12*, 902-903.
75. Zerbino, D.R.; Birney, E. Velvet: Algorithms for de novo short read assembly using de bruijn graphs. *Genome Res* **2008**, *18*, 821-829.
76. Laetsch, D.R.; Blaxter, M.L. Blobtools: Interrogation of genome assemblies. *F1000Research* **2017**, *6*.
77. Dierckxsens, N.; Mardulyn, P.; Smits, G. Novoplasty: De novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res* **2017**, *45*, e18.
78. Bernt, M.; Donath, A.; Juhling, F.; Externbrink, F.; Florentz, C.; Fritzsch, G.; Putz, J.; Middendorf, M.; Stadler, P.F. Mitos: Improved de novo metazoan mitochondrial genome annotation. *Mol Phylogenet Evol* **2013**, *69*, 313-319.
79. Lowe, T.M.; Chan, P.P. Trnascan-se on-line: Integrating search and context for analysis of transfer rna genes. *Nucleic Acids Res* **2016**, *44*, W54-57.
80. Alikhan, N.F.; Petty, N.K.; Ben Zakour, N.L.; Beatson, S.A. Blast ring image generator (brig): Simple prokaryote genome comparisons. *BMC Genomics* **2011**, *12*, 402.
81. Li, H.; Leavengood, J.M., Jr.; Chapman, E.G.; Burkhardt, D.; Song, F.; Jiang, P.; Liu, J.; Zhou, X.; Cai, W. Mitochondrial phylogenomics of hemiptera reveals adaptive innovations driving the diversification of true bugs. *Proc Biol Sci* **2017**, *284*.
82. Hua, J.; Li, M.; Dong, P.; Cui, Y.; Xie, Q.; Bu, W. Phylogenetic analysis of the true water bugs (insecta: Hemiptera: Heteroptera: Nepomorpha): Evidence from mitochondrial genomes. *BMC Evol Biol* **2009**, *9*, 134.

764 83. Sun, X.; Wang, Y.; Chen, P.P.; Wang, H.; Lixiang, L.; Ye, Z.; Wu, Y.; Li, T.; Bu, W.; Xie, Q. Biased
765 heteroplasmy within the mitogenomic sequences of gigantometra gigas revealed by sanger and high-
766 throughput methods. *Zoological Systematics* **2018**, *43*, 356–386.
767 84. Darling, A.C.; Mau, B.; Blattner, F.R.; Perna, N.T. Mauve: Multiple alignment of conserved genomic
768 sequence with rearrangements. *Genome Res* **2004**, *14*, 1394-1403.
769 85. Ronquist, F.; Teslenko, M.; van der Mark, P.; Ayres, D.L.; Darling, A.; Hohna, S.; Larget, B.; Liu, L.; Suchard,
770 M.A.; Huelsenbeck, J.P. Mrbayes 3.2: Efficient bayesian phylogenetic inference and model choice across a
771 large model space. *Syst Biol* **2012**, *61*, 539-542.
772 86. Dong, X.; Chaisiri, K.; Xia, D.; Armstrong, S.D.; Fang, Y.; Donnelly, M.J.; Kadowaki, T.; McGarry, J.W.;
773 Darby, A.C.; Makepeace, B.L. Genomes of trombidid mites reveal novel predicted allergens and laterally
774 transferred genes associated with secondary metabolism. *Gigascience* **2018**, *7*, giy127.
775 87. Huang, X.; Madan, A. Cap3: A DNA sequence assembly program. *Genome Res* **1999**, *9*, 868-877.
776 88. Ratnasingham, S.; Hebert, P.D. Bold: The barcode of life data system (http://www.Barcodinglife.Org). *Mol*
777 *Ecol Notes* **2007**, *7*, 355-364.
778 89. Katoh, K.; Standley, D.M. Mafft multiple sequence alignment software version 7: Improvements in
779 performance and usability. *Mol Biol Evol* **2013**, *30*, 772-780.
780 90. Talavera, G.; Castresana, J. Improvement of phylogenies after removing divergent and ambiguously
781 aligned blocks from protein sequence alignments. *Syst Biol* **2007**, *56*, 564-577.
782 91. Nguyen, L.T.; Schmidt, H.A.; von Haeseler, A.; Minh, B.Q. Iq-tree: A fast and effective stochastic algorithm
783 for estimating maximum-likelihood phylogenies. *Mol Biol Evol* **2015**, *32*, 268-274.
784 92. Bankevich, A.; Nurk, S.; Antipov, D.; Gurevich, A.A.; Dvorkin, M.; Kulikov, A.S.; Lesin, V.M.; Nikolenko,
785 S.I.; Pham, S.; Prjibelski, A.D., *et al.* Spades: A new genome assembly algorithm and its applications to
786 single-cell sequencing. *J Comput Biol* **2012**, *19*, 455-477.
787 93. Seemann, T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **2014**, *30*, 2068-2069.
788 94. Emms, D.M.; Kelly, S. Orthofinder: Solving fundamental biases in whole genome comparisons dramatically
789 improves orthogroup inference accuracy. *Genome Biol* **2015**, *16*, 157.
790 95. Bordenstein, S.R.; Bordenstein, S.R. Eukaryotic association module in phage wo genomes from wolbachia.
791 *Nat Commun* **2016**, *7*, 13155.
792 96. Gu, H.; Chen, Y.; Liu, X.; Wang, H.; Shen-Tu, J.; Wu, L.; Zeng, L.; Xu, J. The effective migration of massilia
793 sp. Wf1 by phanerochaete chrysosporium and its phenanthrene biodegradation in soil. *Sci Total Environ*
794 **2017**, *593-594*, 695-703.
795 97. Baldo, L.; Werren, J.H. Revisiting wolbachia supergroup typing based on wsp: Spurious lineages and
796 discordance with mlst. *Curr Microbiol* **2007**, *55*, 81-87.
797 98. LePage, D.P.; Metcalf, J.A.; Bordenstein, S.R.; On, J.; Perlmutter, J.I.; Shropshire, J.D.; Layton, E.M.;
798 Funkhouser-Jones, L.J.; Beckmann, J.F.; Bordenstein, S.R. Prophage wo genes recapitulate and enhance
799 wolbachia-induced cytoplasmic incompatibility. *Nature* **2017**, *543*, 243-247.
800 99. Bronstein, O.; Kroh, A.; Haring, E. Mind the gap! The mitochondrial control region and its power as a
801 phylogenetic marker in echinoids. *BMC Evol Biol* **2018**, *18*, 80.
802 100. Song, H.; Buhay, J.E.; Whiting, M.F.; Crandall, K.A. Many species in one: DNA barcoding overestimates
803 the number of species when nuclear mitochondrial pseudogenes are coamplified. *Proc Natl Acad Sci U S A*
804 **2008**, *105*, 13486-13491.
805 101. Marsollier, L.; Andre, J.P.; Frigui, W.; Reysset, G.; Milon, G.; Carbonnelle, B.; Aubry, J.; Cole, S.T. Early
806 trafficking events of mycobacterium ulcerans within naucoris cimicoides. *Cell Microbiol* **2007**, *9*, 347-355.
807 102. Wallace, J.R.; Mangas, K.M.; Porter, J.L.; Marcsisin, R.; Pidot, S.J.; Howden, B.; Omansen, T.F.; Zeng, W.;
808 Axford, J.K.; Johnson, P.D.R., *et al.* Mycobacterium ulcerans low infectious dose and mechanical
809 transmission support insect bites and puncturing injuries in the spread of buruli ulcer. *PLoS Negl Trop Dis*
810 **2017**, *11*, e0005553.
811 103. Djouaka, R.; Zeukeng, F.; Daiga Bigoga, J.; N'Golo Coulibaly, D.; Tchigossou, G.; Akoton, R.; Aboubacar,
812 S.; Tchebe, S.J.; Nantcho Nguepdjo, C.; Adeoti, R., *et al.* Evidences of the low implication of mosquitoes in
813 the transmission of mycobacterium ulcerans, the causative agent of buruli ulcer. *Can J Infect Dis Med*
814 *Microbiol* **2017**, *2017*, 1324310.
815 104. Garchitorena, A.; Roche, B.; Kamgang, R.; Ossomba, J.; Babonneau, J.; Landier, J.; Fontanet, A.; Flahault, A.;
816 Eyangoh, S.; Guegan, J.F., *et al.* Mycobacterium ulcerans ecological dynamics and its association with
817 freshwater ecosystems and aquatic communities: Results from a 12-month environmental survey in
818 cameroon. *PLoS Negl Trop Dis* **2014**, *8*, e2879.
819 105. Roche, B.; Benbow, M.E.; Merritt, R.; Kimbirauskas, R.; McIntosh, M.; Small, P.L.; Williamson, H.; Guegan,
820 J.F. Identifying the achilles' heel of multi-host pathogens: The concept of keystone "host" species illustrated
821 by mycobacterium ulcerans transmission. *Environ Res Lett* **2013**, *8*, 045009.

106. Morris, A.; Guegan, J.F.; Benbow, M.E.; Williamson, H.; Small, P.L.; Quaye, C.; Boakye, D.; Merritt, R.W.; Gozlan, R.E. Functional diversity as a new framework for understanding the ecology of an emerging generalist pathogen. *Ecohealth* **2016**, *13*, 570-581.
107. Iturbe-Ormaetxe, I.; Walker, T.; SL, O.N. Wolbachia and the biological control of mosquito-borne disease. *EMBO Rep* **2011**, *12*, 508-518.
108. Crainey, J.L.; Wilson, M.D.; Post, R.J. Phylogenetically distinct wolbachia gene and pseudogene sequences obtained from the african onchocerciasis vector simulium squamosum. *Int J Parasitol* **2010**, *40*, 569-578.
109. Woodford, L.; Bianco, G.; Ivanova, Y.; Dale, M.; Elmer, K.; Rae, F.; Larcombe, S.D.; Helm, B.; Ferguson, H.M.; Baldini, F. Vector species-specific association between natural wolbachia infections and avian malaria in black fly populations. *Sci Rep* **2018**, *8*, 4188.
110. Cordaux, R.; Pichon, S.; Hatira, H.B.; Doublet, V.; Greve, P.; Marcade, I.; Braquart-Varnier, C.; Souty-Grosset, C.; Charfi-Cheilchrouha, F.; Bouchon, D. Widespread wolbachia infection in terrestrial isopods and other crustaceans. *Zookeys* **2012**, 123-131.
111. Baltanas, A.; Zabal-Aguirre, M.; Pita, M.; Lopez-Fernandez, C. Wolbachia identified in a new crustacean host: An explanation of the prevalence of asexual reproduction in non-marine ostracods? *Fund Appl Limnol* **2007**, *169*, 217-221.
112. Aikawa, T.; Anbutsu, H.; Nikoh, N.; Kikuchi, T.; Shibata, F.; Fukatsu, T. Longicorn beetle that vectors pinewood nematode carries many wolbachia genes on an autosome. *P Roy Soc B-Biol Sci* **2009**, *276*, 3791-3798.
113. McNulty, S.N.; Foster, J.M.; Mitreva, M.; Hotopp, J.C.D.; Martin, J.; Fischer, K.; Wu, B.; Davis, P.J.; Kumar, S.; Brattig, N.W., *et al.* Endosymbiont DNA in endobacteria-free filarial nematodes indicates ancient horizontal genetic transfer. *Plos One* **2010**, *5*.
114. Koutsovoulos, G.; Makepeace, B.; Tanya, V.N.; Blaxter, M. Palaeosymbiosis revealed by genomic fossils of wolbachia in a strongyloidean nematode. *Plos Genet* **2014**, *10*.
115. Funkhouser-Jones, L.J.; Sehnert, S.R.; Martinez-Rodriguez, P.; Toribio-Fernandez, R.; Pita, M.; Bella, J.L.; Bordenstein, S.R. Wolbachia co-infection in a hybrid zone: Discovery of horizontal gene transfers from two wolbachia supergroups into an animal genome. *Peerj* **2015**, *3*.
116. LePage, D.; Bordenstein, S.R. Wolbachia: Can we save lives with a great pandemic? *Trends in Parasitology* **2013**, *29*, 385-393.
117. Gerth, M.; Gansauge, M.T.; Weigert, A.; Bleidorn, C. Phylogenomic analyses uncover origin and spread of the wolbachia pandemic. *Nature Communications* **2014**, *5*.
118. Armbruster, P.; Damsky, W.E.; Giordano, R.; Birungi, J.; Munstermann, L.E.; Conn, J.E. Infection of new- and old-world aedes albopictus (diptera : Culicidae) by the intracellular parasite wolbachia: Implications for host mitochondrial DNA evolution. *J Med Entomol* **2003**, *40*, 356-360.
119. Ellegaard, K.M.; Klasson, L.; Naslund, K.; Bourtzis, K.; Andersson, S.G.E. Comparative genomics of wolbachia and the bacterial species concept. *Plos Genet* **2013**, *9*.
120. Casiraghi, M.; Bordenstein, S.R.; Baldo, L.; Lo, N.; Beninati, T.; Wernegreen, J.J.; Werren, J.H.; Bandi, C. Phylogeny of wolbachia pipientis based on glta, groel and ftsz gene sequences: Clustering of arthropod and nematode symbionts in the f supergroup, and evidence for further diversity in the wolbachia tree. *Microbiol-Sgm* **2005**, *151*, 4015-4022.
121. Hosokawa, T.; Koga, R.; Kikuchi, Y.; Meng, X.Y.; Fukatsu, T. Wolbachia as a bacteriocyte-associated nutritional mutualist. *P Natl Acad Sci USA* **2010**, *107*, 769-774.
122. Tijsse-Klasen, E.; Braks, M.; Scholte, E.J.; Sprong, H. Parasites of vectors--ixodiphagus hookeri and its wolbachia symbionts in ticks in the netherlands. *Parasit Vectors* **2011**, *4*, 228.
123. Chrostek, E.; Gerth, M. Is anopheles gambiae a natural host of wolbachia? *MBio* **2019**, *10*.
124. Cariou, M.; Duret, L.; Charlat, S. The global impact of wolbachia on mitochondrial diversity and evolution. *J Evol Biol* **2017**, *30*, 2204-2210.
125. Zabal-Aguirre, M.; Arroyo, F.; Garcia-Hurtado, J.; de la Torre, J.; Hewitt, G.M.; Bella, J.L. Wolbachia effects in natural populations of chorthippus parallelus from the pyrenean hybrid zone. *J Evol Biol* **2014**, *27*, 1136-1148.
126. Raychoudhury, R.; Baldo, L.; Oliveira, D.C.; Werren, J.H. Modes of acquisition of wolbachia: Horizontal transfer, hybrid introgression, and codivergence in the nasonia species complex. *Evolution* **2009**, *63*, 165-183.
127. Miyata, M.; Konagaya, T.; Yukuhiro, K.; Nomura, M.; Kageyama, D. Wolbachia-induced meiotic drive and feminization is associated with an independent occurrence of selective mitochondrial sweep in a butterfly. *Biol Lett* **2017**, *13*.

 doi:10.20944/preprints201908.0031.v1

878 128. Damgaard, J.; Buzzetti, F.M.; Mazzucconi, S.A.; Weir, T.A.; Zettel, H. A molecular phylogeny of the pan-
879 tropical pond skater genus limnogonus stal 1868 (hemiptera-heteroptera: Gerromorpha-gerridae). *Mol*
880 *Phylogenet Evol* **2010**, *57*, 669-677.
881