Infection, Genetics and Evolution 28 (2014) 33–43

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Review

Molecular epidemiology, phylogeny and evolution of the filarial nematode *Wuchereria bancrofti*

Scott T. Small *, Daniel J. Tisch, Peter A. Zimmerman

The Center for Global Health and Diseases, Case Western Reserve University, School of Medicine, Cleveland, OH, United States

ARTICLE INFO

Article history:
Received 13 May 2014
Received in revised form 17 August 2014
Accepted 19 August 2014
Available online 29 August 2014

Keywords:
Filariasis
Wuchereria bancrofti
Genetics
Genomics
Transmission
Molecular epidemiology

ABSTRACT

Wuchereria bancrofti (Wb) is the most widely distributed of the three nematodes known to cause lymphatic filariasis (LF), the other two being *Brugia malayi* and *Brugia timori*. Current tools available to monitor LF are limited to diagnostic tests targeting DNA repeats, filarial antigens, and anti-filarial antibodies. While these tools are useful for detection and surveillance, elimination programs have yet to take full advantage of molecular typing for inferring infection history, strain fingerprinting, and evolution. To date, molecular typing approaches have included whole mitochondrial genomes, genotyping, targeted sequencing, and random amplified polymorphic DNA (RAPDs). These studies have revealed much about Wb biology. For example, in one study in Papua New Guinea researchers identified 5 major strains that were widespread and many minor strains some of which exhibit geographic stratification. Genome data, while rare, has been utilized to reconstruct evolutionary relationships among taxa of the *Onchocercidae* (the clade of filarial nematodes) and identify gene synteny. Their phylogeny reveals that speciation from the common ancestor of both *B. malayi* and Wb occurred around 5–6 millions years ago with shared ancestry to other filarial nematodes as recent as 15 million years ago. These discoveries hold promise for gene discovery and identifying drug targets in species that are more amenable to *in vivo* experiments. Continued technological developments in whole genome sequencing and data analysis will likely replace many other forms of molecular typing, multiplying the amount of data available on population structure, genetic diversity, and phylogenetics. Once widely available, the addition of population genetic data from genomic studies should hasten the elimination of LF parasites like Wb.

Infectious disease control programs have benefited greatly from population genetics data and recently from population genomics data. However, while there is currently a surplus of data for diseases like malaria and HIV, there is a scarcity of this data for filarial nematodes. With the falling cost of genome sequencing, research on filarial nematodes could benefit from the addition of population genetics statistics and phylogenetics especially in dealing with elimination programs. A comprehensive review focusing on population genetics of filarial nematode does not yet exist.

Here our goal is to provide a current overview of the molecular epidemiology of *W. bancrofti* (Wb) the primary causative agent of LF. We begin by reviewing studies utilizing molecular typing techniques with specific focus on genomic and population datasets. Next, we used whole mitochondrial genome data to construct a phylogeny and examine the evolutionary history of the *Onchocercidae*. Then, we provide a perspective to aid in understanding how population genetic techniques translate to modern epidemiology. Finally, we introduce the concept of genomic epidemiology and provide some examples that will aid in future studies of Wb.

Contents

* Corresponding author. Address: The Center for Global Health and Diseases, School of Medicine, Case Western Reserve University, 2109 Adelbert Road Biomedical Research Building, Room E426, Cleveland, OH 44106-4983, United States. Tel.: +1 216 368 6065.
E-mail address: sts45@case.edu (S.T. Small).

http://dx.doi.org/10.1016/j.meegid.2014.08.018
1567-1348/

1. Lymphatic filariasis and *Wuchereria bancrofti*

W. bancrofti (Wb) is the main causative agent (>90% of cases) of lymphatic filariasis (LF) (>90% of cases) (Farrar et al., 2013). LF is a debilitating disease affecting more than 120 million in 73 countries, with over 1.39 billion people living in endemic area for LF (World Health Organization Global Programme to eliminate Lymphatic Filariasis: Progress Report on Mass Drug Administration, 2012). Currently there is a massive undertaking by the World Health Organization (WHO), the global program to eliminate lymphatic filariasis (GPELF), to eliminate LF and its causative agents (Wb, *Brugia malayi*, *Brugia timori*), through a combination of mass drug administration (MDA) and integrated vector management (IVM) (Progress Report 2000–2009 and Strategic Plan 2010–2020 of the Global Programme to Eliminate Lymphatic Filariasis: Halfway Towards Eliminating Lymphatic Filariasis, 2010). The current goal of GPELF is to eliminate LF by the year 2020. Progress towards this goal is promising, with 3.9 billion doses of medicine distributed to people in 65 countries (12 in post-MDA surveillance) (Progress Report 2000–2009 and Strategic Plan 2010–2020 of the Global Programme to Eliminate Lymphatic Filariasis: Halfway Towards Eliminating Lymphatic Filariasis, 2010). However, while many countries have moved toward surveillance stages after >5 rounds of MDA, some countries remain stymied by infrastructure and geography (World Health Organization Global Programme to Eliminate Lymphatic Filariasis: Progress Report on Mass Drug Administration, 2012).

2. Detection and molecular typing of *W. bancrofti*

2.1. Detection of W. bancrofti

A variety of assays are available to diagnose and/or monitor LF infection in populations (see review in McNulty et al., 2013). Some of the assays are species specific, such that they can also identify the causative agent of the infection (Wb, *B. malayi*, or *B. timori*). Thus, accurate diagnosis has direct feedback on the ability to effectively treat LF and evaluate future risk. Here we provide a brief summary, in mechanism mostly, of common diagnostic tests. As we are mainly concerned with detection of Bancroftian filariasis, we focus on detection methods specific to Wb. We therefore point interested readers to McNulty et al., 2013 for methods covering both Bancroftian and Brugian Filariasis detection.

2.1.1. Microscopy

Microscopy is by far the most widely used of the assays to verify infection with Wb (Farrar et al., 2013). While useful alone, microscopy is also valuable as verification in combination with other methods (e.g., paired with antigen tests). Though there are many accepted methods to count Wb microfilaria (MF) in human blood smears (e.g., thick smear, thin smear etc...), the accuracy is dependent on the experience of the observer. We point the reader to Manson's Tropical Diseases 23rd Edition (Farrar et al., 2013) for further details.

2.1.2. Antigen tests

Adult worms of Wb cannot be directly observed while in humans (with the exception of ultrasonography of worm nests in the scrotum). However, we can detect the presence of Wb with semi-quantitative assays for circulating parasite antigen. These assays use monoclonal antibodies raised against related species of nematodes (e.g., *Dirofilaria immitis*) with antigenic targets that cross-react with human antibodies to Wb. Recent assays have vastly improved the sensitivity and specificity of antigen tests (Weil et al., 2013).

2.1.3. DNA based tests

Early efforts to improve diagnostic sensitivity and specificity of filarial parasites were first accomplished by identifying DNA sequences repeated in the parasite's genomes (Erttmann et al., 1987; McReynolds, 1986). Identification of these sequences and the technological advance of the polymerase chain reaction (PCR) enabled development of numerous different assays that made it possible to perform large-scale prevalence surveys on parasitic nematodes including *B. malayi* (Lizotte and Supali, 1994), Wb (Zhong et al., 1996), *Onchocerca* species and strains (Meredith et al., 1991; Zimmerman et al., 1992, 1993) and Wb (Fischer et al., 2003; Williams et al., 2002).

2.1.4. Antibody tests

The BmM14 antibody capture assay has recently become standardized and commercially available, enabling its consideration

for broader application to GPELF (Ramzy et al., 1995; Weil et al., 2011). In addition, an alternative recombinant protein (*Wb123*) has been developed that is highly immunogenic and specific to Wb L3 and made into a rapid, high-throughput point-of-care assay of LF exposure (Steel et al., 2012). Serologic markers of exposure, particularly in sentinel sub-populations with low LF exposure prior to MDA, are believed to represent the greatest statistical specificity for LF monitoring, as they have high sensitivity and specificity to capture LF exposure in sentinel sub-populations (Steel et al., 2012; Tisch et al., 2008).

2.2. Molecular typing of W. bancrofti

In addition to improving the detection of the parasites from biological samples, PCR-based assays make it possible to develop strategies to identify and study sequence variations to distinguish between related species and strains. For example, sequence variations in *Onchocerca volvulus* were first used to distinguish between forest and savannah strains, the former associated with mild onchocerciasis and the latter with severe blinding forms of the disease (Zimmerman et al., 1992).

A caveat of molecular typing methods is that we must account for the complexity of infection (COI). The complexity of infection refers to the condition where more than one genetically distinct individual is present in the infection (Volkman et al., 2012b). In Wb infection at least two individual adult worms must be present for MF production. COI inherently complicates population genetic analysis that assumes a single haploid or single diploid individual in the sample. COI has been handled with a multitude of pre- and post-sequencing methods. Pre-sequencing methods focus on diluting the sample before DNA sequencing (Peters et al., 2012) or isolating amplified fragments by vector-based cloning (Small et al., 2013). Post-sequencing methods focus on reconstructing parasite haplotypes (Beerenwinkel and Zagordi, 2011) or summarizing data based on only major alleles (Volkman et al., 2012a,b). We mention it here, as a brief aside, because it is an issue that must overcome before parasite sequencing can be deemed high-throughput.

2.2.1. Non-sequence based DNA variation

One of the simplest methods to query DNA variation is to construct a unique fingerprint of the individual. In regard to Wb infection this is to include the entire population of Wb worms in a single host, hereafter called an infrapopulation. The fingerprint can be constructed using restriction digests (RFLPs) or randomly amplified lengths of DNA (RAPDs). The benefit of these techniques is that they require little prior data for the organism of study. The drawback is that the data can be difficult to analyze and can only resolve the presence/absence of an allele at a site and not the genotype.

One of the first studies to study genetic variation within populations of Wb infrapopulations took place in India in 2002 (Kumar et al., 2002). Kumar et al. used RAPDs to query the amount of genetic variation within and among three populations infected with Wb: one urban population and two rural populations. Their results demonstrated for the first time that genetic variability among infrapopulations from same location could vary greatly. The population of infected individuals within the rural village of Athipakkam was 6 times less variable than that in either the urban area of Pondicherry or rural village of Chinnanergunam. They explain this observation by noting that the village of Athipakkam was only recently infected with Wb and has yet to undergo chemotherapy, while the other locations have been under active treatment with diethylcarbamazine (DEC). Hoti et al. returned in 2008 to resample the same three populations, with the addition of the rural village of Alampoondi (Hoti et al., 2008). The results in 2008 largely mirrored those found earlier. Hoti et al. recovered higher genetic variation in the urban area of Pondicherry than the surrounding rural villages. They also determined that genetic similarity was greater between urban and geographically proximate rural villages. They conclude that high migration into Pondicherry from the outlying rural villages has increased genetic variation of the Wb.

A second study by Thangadurai et al. used RAPDs to characterize geographic variation in genotype (phylogeography (Avise, 2000)) from 71 MF positive individuals residing in 8 geoclimatic regions of India (Thangadurai et al., 2006). Overall genetic variation was similar within regions with heterozygosity ranging from 0.29 to 0.39. Average genetic differentiation was high among all regions with an average G_{st} of 0.5859. An estimation of gene flow indicated the presence of 2 isolated Wb strains: strain 1 on the eastern side and stain 2 on the western side of the Western Ghats mountain range. A regional test of genetic differentiation between east and west strains was highly significant with an F_{st} of 0.7978. The authors determine that high differentiation is due to very little gene flow, possibly because travel is hampered by the thickly forested mountain ranges. The eastern side strain was also recognized as the major parasite in the country spanning from the eastern peninsular region to the northern plains.

A final study in India utilized RAPDs to examine the genetic variability between diurnally sub-periodic Wb infections in the Nicobar group of Islands (Dhamodharan et al., 2008). Dhamodharan et al. collected blood samples from 40 MF positive individuals residing in eight locations on five different islands. They conclude that the biting cycle of the two primary vectors, *Culex quinquefasciatus* (nocturnal) and *Ochlerotatus niveus* (crepuscular), is driving incipient speciation between Wb strains exhibiting different periodicities. Noting that Wb utilizing the crepuscular vector has a bimodal peak in capillary densities.

2.2.2. DNA sequencing

Sequencing has replaced most other forms of data involving DNA, surpassing methods using restriction sites or randomly amplified fragments. DNA sequence allows the direct imputation of genotype as well as a multitude of statistical tests based on detecting natural selection via amino acid changes. Here we review the handful of studies that have used DNA sequence to infer infection dynamics of Wb. One useful distinction is the difference between putatively neutral loci—not under the influence of selection (typically not protein coding) and protein coding loci that may or may not behave in a neutral manner. Most population genetic models assume that natural selection is not acting to change allele frequencies and therefore assumptions based on selected loci may be incorrect.

One of the first studies to report DNA sequence data was by (Schwab et al., 2005). The main goal of the study was to detect the presence of albendazole resistance in the β-tubulin gene. Albendazole is a widely used anti-parasitic drug that targets the parasites ability to construct microtubules (Lubega and Prichard, 1990). There are two main positions noted for high-affinity binding of Albendazole: codon position 167 and 200. A mutation altering the amino acid from phenylalanine to tyrosine is known to confer resistance. A previous study in India in 2003 had failed to find the codon 200 substitution using infrapopulations from 14 patients (Hoti et al., 2003). The authors remark that pooling samples reduces the ability to detect rare mutation and therefore the 2003 study may not have had sufficient power to reject the presence of drug resistance. Similarly a study in India in 2009, focusing on exons 5 and 6 of the β-tubulin gene, also failed to detect the codon 167 or 200 substitution in 55 Wb isolates (Hoti et al., 2009). Schwab et al. developed an assay to query codon position 200 of the β-tubulin gene, specifically searching for the phenylalanine to tyrosine substitution (F → Y) at position 200. Schwab et al. utilized populations of Wb infected individuals from treated (400 mg albendazole in combination with 200 μg/kg ivermectin)

and untreated areas in Burkina Faso and Ghana. For all assays they tested individual MF separated from the infected blood sample. In the four populations from Burkina Faso, Schwab et al. observed the resistance allele at 26.2% frequency in untreated populations, 60.2% in populations treated once and 86.2% in populations treated twice with the drug combination. In the two Ghana populations, neither of which has ever received treatment, Schwab et al. observed the resistant allele at a frequency of 2.7% and 0.33%. Concluding that these untreated regions may provide a challenge for future elimination efforts.

Churcher et al. in 2008 made additional inferences using the data of Schwab et al. Citing an excess of homozygosity caused by non-random mating, they determined that the alleles at codon 200 were not within Hardy–Weinberg equilibrium (HWE) among host populations. The insight provided by Churcher et al. was that the inherent inbreeding would speed up the fixation of a resistance allele in a population.

Besides searching for the presence of known resistant variants, DNA sequencing has also been used to assess the suitability of new drug targets. The *glutathione S-transferase* (Wb-GST) gene is a promising drug target for control of Wb (Sakthidevi et al., 2013). *Wb-GST* represents a major class of genes responsible for production of detoxifying enzymes involved in lipid peroxidation, a defense against the host oxidative stress environment. Sakthidevi et al. characterized the genetic diversity in *Wb-GST* in four endemic regions of India. They discovered three variant *Wb-GST* alleles, confirming that there is natural variation in the parasite population. Subsequent kinetic experiments demonstrate that allelic variants posses different affinities to well-known *GST* inhibitors such as S-hexyl glutathione, ethacrynic acid, plumbagin, and curcumin (Srinivasan et al., 2009).

All of the above studies have dealt specifically with nuclear DNA but mitochondrial DNA (mtDNA) has unique advantages to the study of Wb. First, mtDNA is organelle based making it higher copy number than nuclear DNA. Second, mtDNA is haploid, eliminating the confounding problems of recombination in analysis. Third, mtDNA is maternally inherited, which makes the effective population size—geometric population size over time, 4 times smaller than nuclear DNA. Since the speed of genetic drift is proportional to the inverse of effective population size, this means that genetic differentiation is detected earlier in mtDNA. Also, typically, but not always, mutation rate is higher in mtDNA, creating a better molecular clock for recent evolutionary events (Avise et al., 1987). Because of these reasons, the mitochondrial genome has been used regularly for molecular epidemiology studies to characterize population diversity, track migration, and infer relationships between geographically separated population clusters. For example mtDNA polymorphisms have been used in nematode parasites to show contrasting levels of diversity, such as low-level variation in *Strongyloides ratti* (Fisher and Viney, 1996) and high-level variation among *Ascaris* spp. (Anderson et al., 1995).

Small et al. in 2013 used the Wb mitochondrial gene *cytochrome oxidase 1* (*COX1*) to infer the genetic diversity and genetic differentiation between populations of Wb (Small et al., 2013). They sequenced and cloned the *COX1* gene from 16 people infected with Wb selected from each of 7 villages that encompassed both high and moderate annual transmission potentials. Out of 487 sequences, they observed 203 unique *COX1* haplotypes, with a single haplotype appearing in 93% of all infections. Similar to Kumar et al. they observed high genetic variation within infrapopulations with up to 4-fold differences in diversity between different individual infections. Unlike previous studies in India, they did not find significant genetic differentiation between populations but found most variation within infrapopulations. This suggested that either recent mixing or high migration among sampled localities had homogenized infections. A haplotype network confirmed this assumption showing that most haplotypes appeared in all populations. Further analysis using *K*-means clustering established a connection between geographic proximity and genetic difference as well as highlighting several outlying individuals as possible immigrants from outside the study area. Finally, Small et al. used the sequence data to test whether MDA had any significant effect on genetic diversity. While they found no significant deviation from the null model, they determined that more loci are needed before the genetic effects of MDA could be realized.

2.2.3. Genomic sequencing

Wb falls far behind the other filarial nematodes in available genomic data. To date only a single whole genome sequence has been published (Desjardins et al., 2013) with two more studies publishing only mitochondrial genomes (McNulty et al., 2013; Ramesh et al., 2012). The paucity of genomic data stems from the difficulty in obtaining sufficient genetic material. Typically only MF are easily obtainable, with adult worms only accessible through invasive surgeries (MF lack sufficient genetic material to be sequenced without PCR or whole genome amplification). Further complicating genome sequencing is the problem of complexity of infection (COI) (see Molecular Typing), which makes genomes from pooled samples problematic.

In 2013, Desjardins et al. published the first Wb genome in the context of discovering what role the obligate α-proteobacterial endosymbiont *Wolbachia* plays in filarial nematodes. The Wb genome was sequenced in 2010 to approximately 12× coverage with a genome size of 81.5 million bases and approximately 14,496–15,075 genes. The results were inconclusive with regards to *Wolbachia*. However, Desjardins et al. did discover that the Wb genome is enriched for genes related to cellular adhesion and the extra-cellular matrix (e.g., cadherins, laminins and fibronectins) (Desjardins et al., 2013). In contrast the *Loa loa* genome is enriched for hyaluronidases; believed useful for tissue penetration in the highly mobile adults life stage (Desjardins et al., 2013). The genomic difference, while slight, provides evidence that some biology is reconstructed from genome sequence.

The first mitochondrial genome (mtgenome) of Wb was published by Ramesh et al. in (2012). Ramesh et al. sequenced three mtgenomes, one each from Mali, PNG, and India. Other than reporting on the gene content and tRNA structure, which can be found in (Ramesh et al., 2012), they also reconstructed the evolutionary history between the three regions. Ramesh et al. concluded that Wb most likely originated in Africa and later migrated to India and PNG, however more sequence data is needed to reject a hypothesis of recent gene flow.

3. Evolution and phylogeny of *W. bancrofti*

One of our interests in examining the phylogeny of the filarial nematodes (*Animalia:Nematoda:Secernentea:Spiruria:Spirurida:Filarioidea:Onchocercidae*) is to determine the evolutionary relationship between extant species for the purpose of comparative analysis. For example, *B. malayi* is closely related to Wb and can be cultured in non-human species, whereas Wb and *L. loa* are dependent on primate hosts. If we are to leverage information collected from lab experiments on *B. malayi* to help us better understand Wb, then it is necessary that we undertake a detailed analysis of their evolutionary relationship.

3.1. Mitochondrial genome phylogeny

The most complete mitochondrial genome phylogeny to date included 65 species of nematodes (Liu et al., 2013). To utilize the continuous publications of mitochondrial data, we searched

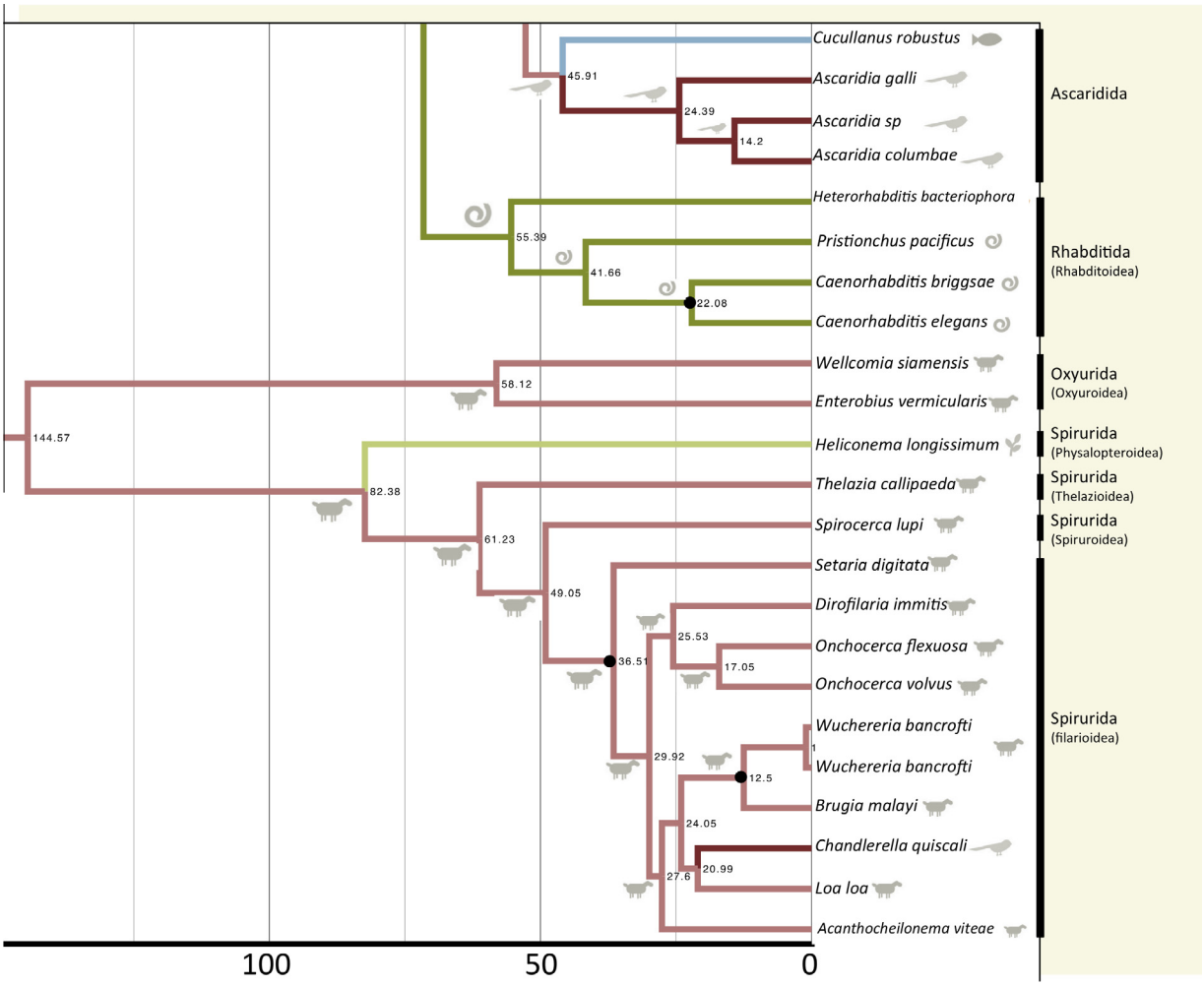

Fig. 1. Mitochondrial genome phylogeny of filarial nematodes. The phylogeny was truncated to emphasize relationships within the filarial worms (Onchocercidae) while providing out-group information from other well-known species. The phylogeny was constructed using the program BEAST 1.7.5 (Drummond et al., 2012) with divergence times calibrated using the mitochondrial mutation rate of *Pristionchus pacificus* (Molnar et al., 2011) and the estimated divergence time of 80–110 million years between *Caenorhabditis elegans* and *C. briggsae* (Stein et al., 2003). Divergence times are placed on each node as a decimal with a coarse scale on the bottom-horizontal of the phylogeny. All times are represented in millions of generations. We reconstructed host preferences by way of ancestral trait inference in BEAST 1.75, after obtaining a well-supported phylogeny (all nodes >95% unless otherwise noted), For each branch and node, colors denote the host preference as well as the inferred state of the most recent common ancestor. Colors correspond to: pink = mammals, red = birds, yellow = plants, green = free-living/bacterial feeders, blue = fish. We note that there is an over representation of mammalian parasitic nematodes on the phylogeny biasing deep node preferences. Colors of external branches are therefore extended only as deep as 95% confidence surrounding the ancestral node. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

GenBank for published mitochondrial genome sequences of nematode species (search terms: *("Nematoda"[Organism] OR Nematoda[All Fields]) AND complete[All Fields] AND mitochondrion [All Fields])*. We found complete mitochondrial genomes for 78 species of nematodes, an increase on the most recent mitochondrial phylogeny (Liu et al., 2013). A translational alignment of the 78 nematode species produced a dataset 12,160 nucleotides in length. The alignment was generated using only the 12 concatenated protein coding genes common to all nematodes (for an example using Wb see Ramesh et al., 2012), since an exploratory data analysis showed no improvement by including the transfer RNAs or ribosomal RNAs in the analysis.

We constructed a Bayesian phylogenetic tree using the aligned mitochondrial genomes and the program BEAST 1.7.5 (Drummond et al., 2012). The following settings in BEAST 1.7.5 were found to yield the highest tree likelihood: generalized time reversible model (GTR) with 4 rate categories, lognormal relaxed clock, Yule process speciation model, and a UPGMA starting tree. To estimate divergence times, we used the mitochondrial mutation rate of *Pristionchus pacificus* (Molnar et al., 2011) and the estimated divergence time of 100 million years between *Caenorhabditis elegans* and *Caenorhabditis briggsae* (Stein et al., 2003). The resulting model was run 3 independent times for 50 million generations each. Results were combined using LogCombiner 1.7.5.

The relationships among taxa were found to be in broad agreement with previously published datasets utilizing fewer samples (Liu et al., 2013). Therefore, we will forego discussion of the general topology and instead focus on the order Spirurida, which contains the family *Onchocercidae* (filarial nematodes). As noted in prior publications also utilizing the mitochondrial genome, there are some discrepancies between the rDNA and mitochondrial topology. For the filarial worms this concerns the asserted monophyly between the Spirurida, Ascaridida, and Oxyurida. In our phylogeny as well as others (Liu et al., 2013) there was no support for a common ancestor between the three orders. There was more support for the monophyly between the Spirurida and Oxyurida, with the Ascaridida nested within the Rhabditida (also found in Liu et al., 2013). The main result of the non-monophyletic relationship is then the independent origin of host mammalian and bird host preference.

3.2. Divergences and most recent common ancestor (MRCA)

Knowledge about when two species most recently shared a common ancestor (MRCA) provides a context to their evolutionary relationship. Wb and *B. malayi* are both causative agents of LF and share similarities in morphology, host choice, and geographic distribution (see McNulty et al., 2013 for a summary). Based on the mitochondrial genome, Wb and *B. malayi* are relatively young species with MRCA as recent as 12.5 (2.5–30) million generations ago (mga). Given estimates of generation time for Wb (8 months) and *B. malayi* (2–3 months) (Farrar et al., 2013), this was likely 4–6 million years ago (mya) (Fig. 1). Given this relatively recent divergence, it is possible that by studying *B. malayi*, we can gain insight into the life history of Wb (Aliota et al., 2011; Laney et al., 2010). Other interpretations of this divergence time are speculation, but future studies may want to examine the possibility that Wb and *B. malayi* began to diverge at the same time as the common ancestors of human and chimpanzees. The geographic distribution of Wb and evidence from the mitochondrial genomes of Africa and PNG (Ramesh et al., 2012), provide limited evidence to this end.

While the mitochondria provided robust support for topology, the divergence time estimates have large credibility intervals. Relying solely on estimates of mitochondrial mutation rates lead us to underestimate divergence time between *C. elegans* and *C. briggsae* (22.08 mya using mitochondrial mutation rate versus 80 mya in Stein et al.). The discrepancy is likely due to rate variation among the lineages as well as 3rd position mutational saturation of the mitochondrial genome. It is likely that the parasitic nematodes mutate faster than free-living nematodes due to host immune pressure and free-radical exposure *in vivo* (Holterman et al., 2006). We considered rate variation when estimating divergence times and are more confident in our estimates for younger species pairs than we are for deeper divergences. A solution would be to use multiple independent genes, since the mitochondria has only a single history, and more than a single calibration time. This is difficult on two accounts. First, given the species diversity in the nematodes it would be difficult to sample genes evenly across the taxa. Second, given the soft bodies of the nematodes, little fossil evidence exists for secondary calibration.

3.3. Ancestral state of host preference

Along with estimating a MRCA between species, we can also examine other traits with reference to the phylogeny, such as shared host preference. The known host specificity of extant filarial nematodes is primarily mammalian, with the single nested outlier species *Chandlerella quiscali*. *C. quiscali* parasitizes the common grackle (*Quiscalus quiscula*)—a widespread bird species in North America. Based on the ancestral reconstruction it is likely that *C. quiscali* evolved the ability to parasitize birds after diverging from *L. loa*. This acquisition of bird host preference is also independent of other parasitic nematodes such as *Ascaridia galli*—a well-known parasite of chickens. It is not clear how often along the phylogeny host switching among the filarial nematodes may have occurred or if the rate of trait loss and gain is of equal probability (Cunningham, 1999). A definite problem is the large amount of sampling bias, favoring nematode parasites of economic importance especially mammalian parasites. However, given estimates of the human-chimpanzee divergence—6 mya (Patterson et al., 2006)—it is likely that the MRCA of Wb and *B. malayi* was infecting non-human primates. This is also borne out of data showing that *B. malayi* can be made to parasitize non-human primates (Farrar et al., 2013). The specialization onto human hosts may have then taken place more recently in the evolutionary history of Wb as there is an extant species, *Wuchereria kalimantani*, that are presumed to parasitize non-human primates (Palmieri and Purnomo, 1980; Zhong et al., 1996), the necessary consideration in the evolution of host preference is that of exaptation. If *B. malayi*, which is closely related to Wb, retains functional adaptations to allow host generalization, then so too might Wb. Comparisons among filarial nematodes capable of infecting multiple hosts may lead to insight into functional gene classes that control host specialization. This is a necessary consideration during any treatment program to prevent reservoir populations of parasites from surviving.

4. Molecular epidemiology and population genetics

Population genetics have proved to be a valuable tool in deciphering the evolutionary history of parasites as well as improving knowledge of infectious disease epidemiology (Steinauer et al., 2010; Volkman et al., 2012a). Genetic analysis can provide a glimpse into the epidemic history of a parasite that can complement epidemiological analysis or even in some cases replacing missing surveillance data (Dearlove and Wilson, 2013; Didelot et al., 2014). With the advent of polymorphic molecular markers for Wb, it is now possible to apply population genetic analysis to LF. Thus, we now have the ability to answer questions like:

- Does strain distribution vary by infection or disease prevalence?
- Does strain distribution vary by location?
- Can strains be tracked to detect transmission patterns?
- Can strains identify sources of re-introduction?
- What are the dominant strains of Wb in specific locations?
- Does strain distribution change after MDA or vector control?
- Can we detect strains resistant to intervention via MDA or vector control?

In the following sections, we briefly introduce some common population genetic statistics and their utility to molecular epidemiology. We illustrate concepts by using the *COX1* sequences available from Small et al., 2013 on GenBank (KC558603–KC559091).

4.1. Genetic diversity and life history

In the above (Section 2.2), we have seen numerous examples of high genetic diversity in Wb populations. Yet, the measure of genetic diversity only provides us with a description of the molecular variation at the sequence level. What we are truly interested in is determining which mechanistic processes could have given rise to similar values of genetic diversity. One of the simplest ways for epidemiologists to think about genetic diversity is in the context of effective population size. In definition effective population size is a simplification of a complex biological population to one that fits a mathematical model. In practice it can be thought of as the geometric mean of a population over time (for an overview see Hamilton, 2010). Effective population size is useful because it can be mathematically modeled to test the influence of actual infection dynamics and recover information on life history such as reproductive variance, mating system (Tibayrenc and Ayala, 1991), and inbreeding. For example, under the coalescent theory models of effective population size have been built to explore population size fluctuations, population subdivision, and admixture.

Certain aspects of the Wb life cycle directly influence effective population size namely the long reproductive lifespan of adult worms (estimated to be approximately 5 years (Farrar et al., 2013)) and possible polyandry. Though polyandry has not been shown experimentally for Wb, other species of nematodes exhibit polyandrous reproduction where a single female reproduces with

many males (Redman et al., 2008). The long reproductive lifespan confounds generational divisions, typically increasing inbreeding (in small populations) and reducing overall effective population size. Polyandry, on the other hand, generally increases effective population size relative to monogamy. However, if there is a large amount of paternal skew (few males contributing to large portions of the offspring), then polyandry may actually lower effective population size (Holman and Kokko, 2013). Both these factors function independently of infection dynamics to influence effective population size.

4.2. Transmission as gene flow

Wb distribution can be extremely focal and heterogeneous across small geographic areas (World Health Organization Global Programme to Eliminate Lymphatic Filariasis: Progress Report on Mass Drug Administration, 2012). In PNG the disparities in distribution have also recently been affected by MDA that has varied in population coverage, individual participation, and distribution frequency (Bockarie et al., 2002). This has resulted in regions that have been effectively cleared of Wb transmission located adjacent to regions in which transmission continues. Traditionally, transmission is measured by the incidence of new cases in sentinel populations and detection of Wb in captured mosquitoes. Problematic is when prevalence is low, for example after an MDA, where detecting transmission is difficult. Here we will illustrate a method to utilize population genetic analysis on gene flow to estimate transmission. We believe this may prove useful in follow-up surveillance of MDA areas (Progress Report 2000–2009 and Strategic Plan 2010–2020 of the Global Progamme to Eliminate Lymphatic Filariasis: Halfway Towards Eliminating Lymphatic Filariasis, 2010).

We begin our explanation by defining transmission as the process by which Wb worms move between human hosts. Since hosts effectively act as populations of parasites (i.e., infrapopulations), transmission can be characterized by measuring the change of allele frequencies in infrapopulations across time. This is essentially analogous to gene flow—the movement of alleles—between individuals and therefore represented as genetic differentiation (Churcher et al., 2008; Gilabert and Wasmuth, 2013; Small et al., 2013). The difference between the infrapopulation genetic diversity to that in component populations will thus reflect the rate of transmission (Dearlove and Wilson, 2013 provide examples using meta-population models). For example, in areas where transmission rate is high, we would predict genetic differentiation between infrapopulations to be small (e.g., allele prevalence and frequencies are similar), owing to the homogenizing effect of gene flow. Furthermore, we predict that the ratio of infrapopulation genetic distance to that of component populations (within population to among population) should asymptotically approach unity as gene flow, and analogously transmission, increases.

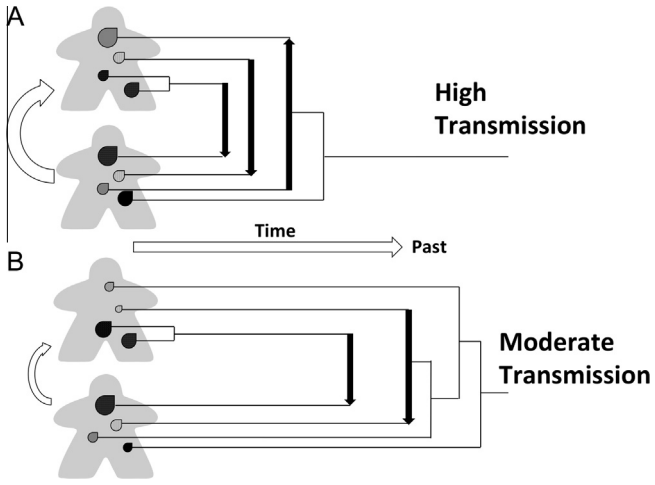

Fig. 2. Simple genetic model of parasite transmission. Two pieces of important information are gained by sequencing from an infrapopulation. First, we gain information on the number of strains in an infection (represented as different patterned circles). Second, we gain information from the frequency of any specific strain in an infrapopulation (represented by the size of the circles). With this data in hand we construct a simple model of Wb transmission. In this model thin horizontal lines connect similar strains and heavy vertical arrows represent transmission event. The total height of each genealogy (in time) reflects the genetic similarity between two infrapopulations. (A) In high transmission areas, the infrapopulations are more genetically similar and therefore the genealogy is shorter. (B) In a moderate transmission area, the populations are less similar and the genealogy will be longer.

To translate ratios of genetic differentiation into terms of transmission, we construct a simple model to estimate the amount of transmission needed to observe similar amounts of genetic differentiation (Fig. 2). Fig. 2 provides a visual illustration of this model for scenarios of high and moderate transmission, similar to PNG (Bockarie et al., 1996). Fig. 2A represents high transmission as a genetic genealogy. Thin horizontal lines connect similar shaded circles, where circles represent specific Wb strains and frequencies. The height of each genealogy (in time) reflects the genetic similarity between two infrapopulations, while the heavy vertical arrows represent transmission events. The model has two unknown parameters which we will estimate, (i) the height of the genealogy and (ii) the rate of transmission events. In reality, we would predict the high transmission model to have a shorter genealogy and more transmission events owing to greater observed genetic similarity, compared to the moderate transmission model (Fig. 2B).

In the context of Fig. 2, we can use information on the observed genetic similarity between infrapopulations in the same locality to estimate transmission rates. Our observed value of genetic similarity from our published PNG dataset (genbank KC558603–KC559091) is 0.9173 for high transmission and 0.8341 for moderate transmission (Small et al., 2013), This observation supports our earlier premise of higher transmission increasing homogenization. Next we use the software *ms* (Hudson, 2002), to simulate different values genealogies under the model in Fig. 2. We analyze the resulting simulations with reference to our observed values of genetic differentiation using the R package *abc* (Csilléry et al., 2012) to predict the values of transmission best supported by the data.

From our simulations, we find that the highest posterior densities support transmission rates of 5–7.5 MF and 0.5–1.5 MF per generation for the high and moderate transmission communities, respectively. If we wish to compare our estimates to those from more traditional methodologies, such as annual transmission potential (ATP), we must extrapolate our estimates to the same units of MF/year. Thus, if we estimate generation time for Wb to be between 3 and 12 months (1–4 generations per year) we can extrapolate that on average 6–24 MF are being successfully transmitted per year in the high transmission community, and on average 1–4 MF in the moderate transmission communities. This provides a slightly different estimate of transmission in comparison to traditional estimates of ATP, since ATP does not take into account survival to reproductive adults. A study on transmission for similar communities in PNG, estimate transmission rates of 140 MF/year for high transmission, down to 7 MF/year for moderate transmission communities (Reimer et al., 2013).

Our example model of transmission (Fig. 2) is over-simplified by assuming (i) individual infection genetic diversity is the effect of multiple transmission events, (ii) all individuals are equally available to vectors, and (iii) vectors are equal in their ability to transmit parasites. Given the inherent flexibility of coalescent

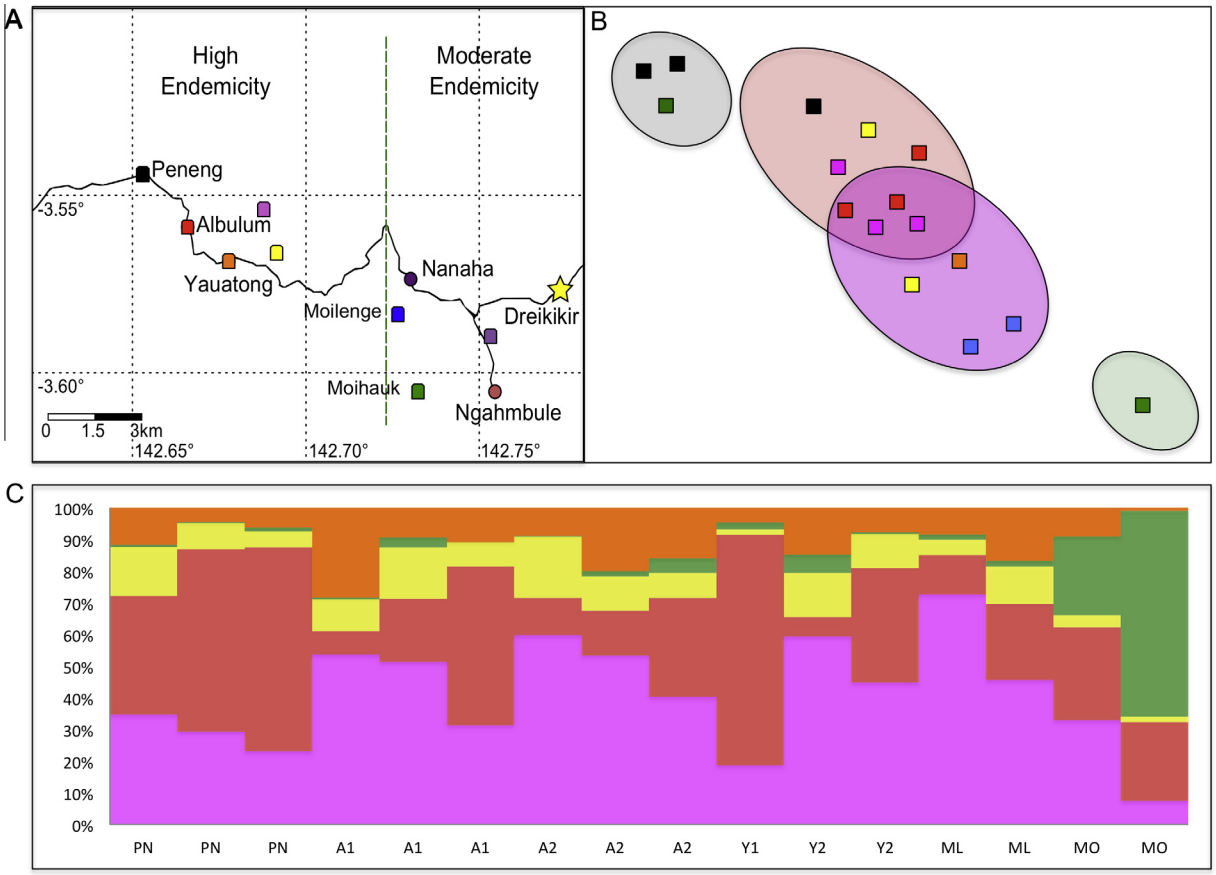

Fig. 3. Inferring parasite dispersal and human migration. To estimate human migration and subsequently parasite dispersal, we illustrate two different methods (B) *K*-means clustering and (C) admixture analysis (C). The map in (A) denotes the sampling localities as well as the color designated to each community. Each infrapopulation is then characterized by genotype and allele frequency. (B) The optimum number of clusters for the *K*-means analysis was four, which was found to minimize the genetic variance. Infrapopulations retain the color of their original communities, while background color clarifies clusters. (C) For the admixture analysis, individuals are displayed along the horizontal axis with the probability of assignment to a cluster, denoted as percent color. Here the best answer was $K = 5$ clusters.

models (see Section 4.4), layers of complexity can easily be added to our model in Fig. 2. Genetic estimates of transmission should not be a replacement for rigorous vector surveillance, especially with the recent tools to detect infective mosquitoes (Laney et al., 2010), but should be recognized as a supplemental tool that, with proper calibration, can even be used to estimate transmission in areas previously unexplored.

4.3. Admixture as migration

Human migration has a direct effect on the size of a treatment unit. If populations are highly mobile, then a multi-community area may in fact behave like a single community. Under this scenario, a previously treated community may experience a relapse when infected individuals from untreated communities immigrate. Furthermore, infected individuals may move between these regions such that an uninfected individual may acquire infection in a region with ongoing transmission. Traditionally, migration is recorded using surveys asking individuals where they previously resided. Yet we can also leverage information about migration directly from the Wb infection itself. We can detect the past movements of infected migrants and assign them to an origin community through the characterization of unique alleles segregating in each infrapopulation.

Let us begin by assuming that Wb worms can diffuse across a landscape in two ways (i) via dispersal of vectors and (ii) by migration of infected individuals. Human migration and vector dispersal will both contribute to genetic differentiation in the Wb population, but in different ways. For example, in PNG the vector of Wb, mosquitoes of the *Anopheles punctulatus* species complex, are estimated to have a flight distance of <2 Kilometers (Bockarie et al., 1996). Therefore, communities that are farther than the average flight distance should logically have little mixing of Wb by vector dispersal alone. Also only a small number of Wb worms are transported with each mosquito bite—on average 1–3 infective L3 larvae per mosquito (Erickson et al., 2013; Hairston and De Meillon, 1968). If we weigh this in contrast to human migration, whereas humans can easily move 2 Kilometers or more, and transport a whole population of Wb, we quickly determine that human migration will be more effective at dispersing Wb than vector dispersal.

If an individual were a migrant to a new community, we would expect their Wb infrapopulation to be more genetically similar to their origin community, although this would eventually be homogenized by transmission. In this way we could, hypothetically, detect not just recent migrants but also older migrants as well by taking into account transmission rate. To illustrate this concept, we use sequences from our published PNG dataset (GenBank KC558603–KC559091) to distinguish human migrants.

In Fig. 3A we recreate the map of our PNG study community (Small et al., 2013), where each sampled community is given a unique color. We present two different methods for detecting human migration, (i) AMOVA-based *K*-means clustering (Meirmans and Van Tienderen, 2004) which groups infrapopulations in a such a way as to minimize genetic differentiation within each cluster (Fig. 3B) and (ii) admixture analysis (Pritchard et al., 2000), which assigns individual Wb worms to predefined or

inferred communities, one or more, based on the genotype and allele frequency (Fig. 3C). As shown in Fig. 3, analyses return the highest likelihood (determined by Akaike's Information Criteria (AIC)), of there being four and five distinct populations of Wb in the study area. In Fig. 3B, representing the *K*-means analysis, we keep the colors of the original community (as shown in Fig. 3A), but have now shaded the background to represent clustering; clustering overlap represent duality in assignment. Given the small geographic scale of the study area (12 Kilometers), migration events are difficult to detect. Yet, one clear example of a migrant is seen in the Moihauk resident who most likely emigrated from Peneng. This is evident by referencing upper left corner of Fig. 3B where the single Moihauk individual (green symbol) clusters with the individuals from Peneng (black symbols). Other than this clear example, most of the resolution is lost by the high diversity in the Albulum regions. Small et al., 2013 reasoned that this was due to high migration into and out of Albulum, which was also supported by survey data (Krufinta Bun 2012 master's thesis).

Using the admixture analysis, we see a very different picture (Fig. 3c). Whereas *K*-means always assigns an infrapopulation to a cluster, maximizing the likelihood of *K* clusters, admixture analysis gives the proportion of each infrapopulation, based on Wb worm assignment to a cluster. Thus, except in extreme cases, admixture analysis will always have greater uncertainty in assignment. In Fig. 3C, the admixture analysis represents each of 16 infrapopulations (horizontal axis) and the proportion of assignment (vertical axis). Solid colors represent different groupings, with no relation to colors in Fig. 3A. It is evident from the admixture analysis that the Moihauk individuals have a unique genetic component (solid green color), which diffuses into the western communities eventually being almost completely absent in the western most community of Peneng. By comparing Fig. 3B and C we see only slight evidence of the migration event from Peneng to Moihauk. Yet, we do recover the isolated individual from Moihauk represented in both analyses. The admixture analysis provides a strain specific illustration of migration by tracking specific alleles. Whereas the *K*-means can only detect a recent migration event that is unidirectional, thus the admixture analysis has more power to differentiate multiple migration events between adjacent communities.

It is clear from this example that detecting migration events using genetic data alone is difficult. However, the admixture analysis can utilize information on known residents available from survey data, which can then be used to infer the migration history of unsurveyed individuals. In general, resolution on migration could be improved through more complete sampling because individuals can only be assigned to a community if the community is characterized in the study. Uneven distribution of genetic diversity may also force less diverse communities to nest within more diverse community, obfuscating the true assignment. Finally, we may find that high rates of human migration, in combination with high transmission, would quickly erase genetic differentiation between communities.

4.4. Coalescent models

Most of the above analyses rely heavily on population genetic models that may not be specifically tailored to Wb. Yet, time and again the standard coalescent model has shown to be robust to variation observed in real biological organisms (for one example see Wakeley et al., 2012). Given the novelty of coalescent and population genetic analysis in Wb epidemiology, it is uncertain whether the standard coalescent model will correctly reconstruct infection dynamics. However, the coalescent model, in its reduced simplicity, is inherently flexible. This is especially necessary in the study of rapidly evolving RNA viruses like Dengue, hepatitis C, and HIV (Lemey et al., 2004; Pybus et al., 2001; Rasmussen et al., 2014), where the coalescent model can be made to easily integrate vector transmission and seasonal fluctuations (Dearlove and Wilson, 2013; Rasmussen et al., 2014).

Future epidemiological studies should consider comparing prevalence data, collected at various time points, with genetic genealogy data collected at the same time. The advantage of this approach is that we can discover which mechanistic factors are important in shaping genetic variation. Then by comparing the ratio of the observed number of infections to the estimated effective population size, we can infer epidemiological factors like the extent of transmission heterogeneity (Magiorkinis et al., 2013).

5. Genomic epidemiology and population genomics

Molecular epidemiology has recently undergone a revolution with the advent of next-generation sequencing. The cost of sequencing is such that now it is comparable to traditional methods of investigation (Didelot et al., 2012; Köser et al., 2012). Genomic data brings new relevance to our previous discussion about population genetics and epidemiology. In the above sections we provided examples on how single or multiple loci can be used to obtain information on Wb infections, but it is now possible to sequence an entire infrapopulation from a single infection or sequence multiple infrapopulations from a community. The data from these experiments can be used to reconstruct past prevalence and understand how strains have been transmitted through time. Here we discuss how the accessibility of next-generation sequencing is creating previously unheralded advances in our understanding of infectious diseases.

5.1. Historical prevalence

One of the most important parameters in epidemiology is R_0, the reproductive number of the disease. Multiple researchers have used genomic data to estimate changes in prevalence over time and thus estimate R_0, even when surveillance data was unavailable (Biek et al., 2007, 2012; Lemey et al., 2004; Magiorkinis et al., 2013; Pybus et al., 2001). Advances in genomic analysis now allow historic population sizes to be inferred using just a single genome (Li and Durbin, 2011). The measured change in effective population size through time is then an indication of how the prevalence of the disease has changed through time (Dearlove and Wilson, 2013). In addition to inferring historic parasite population sizes, we can also use genomic data to reconstruct a spatial expansion (Biek and Real, 2010). This allows us to predict infection rate, the likely epicenter of an outbreak, and determine any future areas at risk. These parameters are only some of the examples of what is possible with genomic data. As more genomic data is gathered, for a larger variety of parasites, new insights and analysis will be born out.

5.2. Transmission and phylodynamics

Multi-locus genotyping only examines a small part of the parasite genome while whole-genome sequencing (WGS) can capture small numbers of single nucleotide changes, typical of even short time scales. This allows the reconstruction of transmission events, as each generation of parasites undergoes meiotic mutations exclusive from the previous generation. As successful parasites propagate these mutations are passed down through a genomic genealogy. If we collect and sequence parasites over time we can then reconstruct a transmission tree that follows the transmission of individual strains and their descendants (Didelot et al., 2014; Rasmussen et al., 2014; Ypma et al., 2012, 2013). Transmission trees

are genealogies substantially differing from phylogenies in their flexibility to within-host infection dynamics (Ypma et al., 2012, 2013). Rasmussen et al. (2014) used this framework to test different hypotheses about dengue serotype 1 (DENV-1) population dynamics in southern Vietnam, successfully reconstructing seasonal variation in prevalence using genomic and epidemiological data. Ypma et al. (2012) used both epidemiological and genetic data in a combined approach to estimate the transmission tree of the avian influenza A (H7N7) virus epidemic that also provided inference on the relative infectiousness of different sizes and types of bird farms. Finally, Didelot et al. (2014) used a Bayesian method to reconstruct a tuberculosis outbreak and constructing a model of transmission delineating possible sources of infection. All the referenced cases provide evidence of the utility of genomic data, conditioning that genomics cannot wholly replace traditional epidemiology, but may form a useful starting point for investigation.

6. Conclusion

As epidemiology enters the genomic age, we gain greater power to understand infection dynamics using both population genetic analysis and molecular diagnostics. There is a rich, well-established theory of Wb infection dynamics independent of genetic data but not a comprehensive model integrating both types of data (Michael and Malecela-Lazaro, 2004; Michael et al., 2006; Slater et al., 2013). Without models relating genetic data to Wb infection, our interpretation and application of genetic data is limited. New models are needed that integrate population genetic inference into current models of Wb transmission and infection. New models will benefit from estimates of population growth, number and frequency of strains, population connectivity across a landscape, and estimates of population sizes. Advanced models have already been used in diseases such as rabies and HCV, where researchers were able to link population size fluctuations with causal events in the history of the disease treatments (Biek et al., 2007; Opgen-Rhein et al., 2005; Stadler and Kühnert, 2013). For parasitic nematodes, genomics is at its infancy. The future will see further studies utilizing genomics and population genetics coupled with epidemiology that will provide new data and new understanding of infection dynamics, giving insight into new strategies of elimination.

Acknowledgements

We thank 2 anonymous reviewers who provided feedback on earlier versions of this manuscript. Research was supported by a T32 Postdoctoral Fellowship in Geographic Medicine and Infectious Disease to S.T. Small (AI007024) and R21 AI103263.

References

Aliota, M.T., Chen, C.-C., Dagoro, H., Fuchs, J.F., Christensen, B.M., 2011. Filarial worms reduce Plasmodium infectivity in mosquitoes. PLoS Negl. Trop. Dis. 5, e963.
Anderson, T.J., Romero-Abal, M.E., Jaenike, J., 1995. Mitochondrial DNA and Ascaris microepidemiology: the composition of parasite populations from individual hosts, families and villages. Parasitology 110, 221–229.
Avise, J., 2000. Phylogeography: The History and Formation of Species. Harvard Press.
Avise, J., Arnold, J., Ball, R., Bermingham, E., Neigel, J.E., Reeb, C.A., Saunders, N.C., 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. Annu. Rev. Ecol. Syst., 489–522.
Beerenwinkel, N., Zagordi, O., 2011. Ultra-deep sequencing for the analysis of viral populations. Curr. Opin. Virol. 1, 413–418.
Biek, R., Real, L.A., 2010. The landscape genetics of infectious disease emergence and spread. Mol. Ecol. 19, 3515–3531.
Biek, R., Henderson, J.C., Waller, L.A., Rupprecht, C.E., Real, L.A., 2007. A high-resolution genetic signature of demographic and spatial expansion in epizootic rabies virus. Proc. Natl. Acad. Sci. U.S.A. 104, 7993–7998.
Biek, R., O'Hare, A., Wright, D., Mallon, T., McCormick, C., Orton, R.J., McDowell, S., Trewby, H., Skuce, R.A., Kao, R.R., 2012. Whole genome sequencing reveals local transmission patterns of *Mycobacterium bovis* in sympatric cattle and badger populations. PLoS Pathog. 8, e1003008.
Bockarie, M., Kazura, J., Alexander, N., Dagoro, H., Bockarie, F., Perry, R., Alpers, M., 1996. Transmission dynamics of *Wuchereria bancrofti* in East Sepik Province, Papua New Guinea. Am. J. Trop. Med. Hyg. 54, 577–581.
Bockarie, M.J., Tisch, D.J., Kastens, W., Alexander, N.D.E., Dimber, Z., Bockarie, F., Ibam, E., Alpers, M.P., Kazura, J.W., 2002. Mass treatment to eliminate filariasis in Papua New Guinea. N. Engl. J. Med. 347, 1841–1848.
Churcher, T.S., Schwab, A.E., Prichard, R.K., Basáñez, M.-G., 2008. An analysis of genetic diversity and inbreeding in *Wuchereria bancrofti*: implications for the spread and detection of drug resistance. PLoS Negl. Trop. Dis. 2, 1–9.
Csilléry, K., François, O., Blum, M.G.B., 2012. Abc: an R package for approximate Bayesian computation (ABC). Methods Ecol. Evol. 3, 475–479.
Cunningham, C.W., 1999. Some limitations of ancestral character-state reconstruction when testing evolutionary hypotheses. Syst. Biol. 48, 665–674.
Dearlove, B., Wilson, D., 2013. Coalescent inference for infectious disease: meta-analysis of hepatitis C. Philos. Trans. R. Soc. B, 1–8.
Desjardins, C.A., Cerqueira, G.C., Goldberg, J.M., Dunning Hotopp, J.C., Haas, B.J., Zucker, J., Ribeiro, J.M.C., Saif, S., Levin, J.Z., Fan, L., Zeng, Q., Russ, C., Wortman, J.R., Fink, D.L., Birren, B.W., Nutman, T.B., 2013. Genomics of *Loa loa*, a Wolbachia-free filarial parasite of humans. Nat. Genet. 45, 495–500.
Dhamodharan, R., Das, M., Hoti, S., 2008. Genetic variability of diurnally subperiodic *Wuchereria bancrofti* in Nicobarese tribe of Nicobar group of Islands, Andaman and Nicobar Islands, India. Parasitol. Res. 103, 59–66.
Didelot, X., Bowden, R., Wilson, D.J., Peto, T.E.A., Crook, D.W., 2012. Transforming clinical microbiology with bacterial genome sequencing. Nat. Rev. Genet. 13, 601–612.
Didelot, X., Gardy, J., Colijn, C., 2014. Bayesian inference of infectious disease transmission from whole-genome sequence data. Mol. Biol. Evol. 31, 1869–1879.
Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol., 3–6.
Erickson, S.M., Thomsen, E.K., Keven, J.B., Vincent, N., Koimbu, G., Siba, P.M., Christensen, B.M., Reimer, L.J., 2013. Mosquito-parasite interactions can shape filariasis transmission dynamics and impact elimination programs. PLoS Negl. Trop. Dis. 7, e2433.
Erttmann, K., Unnasch, T., Greene, B., 1987. A DNA sequence specific for forest form *Onchocerca volvulus*. Nature 327, 415–417.
Farrar, J., Hotez, P.J., Junghanss, T., Kang, G., Lalloo, D., White, N., 2013. Manson's Tropical Diseases, 23rd ed. Elsevier.
Fischer, P., Boakye, D., Hamburger, J., 2003. Polymerase chain reaction-based detection of lymphatic filariasis. Med. Microbiol. Immunol., 3–7.
Fisher, M.C., Viney, M.E., 1996. Microsatellites of the parasitic nematode *Strongyloides ratti*. Mol. Biochem. Parasitol. 80, 221–224.
Gilabert, A., Wasmuth, J.D., 2013. Unravelling parasitic nematode natural history using population genetics. Trends Parasitol. 29, 438–448.
Hairston, N.G., De Meillon, B., 1968. On the inefficiency of transmission of *Wuchereria bancrofti* from mosquito to human host. Bull. World Health Organ. 38, 935–941.
Hamilton, M.B., 2010. Population Genetics, first ed. Wiley-Blackwell.
Holman, L., Kokko, H., 2013. The consequences of polyandry for population viability, extinction risk and conservation. Philos. Trans. R. Soc. B 368, 20120053.
Holterman, M., van der Wurff, A., Van Den Elsen, S., van Megen, H., Bongers, T., Holovachov, O., Bakker, J., Helder, J., 2006. Phylum-wide analysis of SSU rDNA reveals deep phylogenetic relationships among nematodes and accelerated evolution toward crown Clades. Mol. Biol. Evol. 23, 1792–1800.
Hoti, S.L., Subramaniyan, K., Das, P.K., 2003. Detection of codon for amino acid 200 in isotype 1 β-tubulin gene of *Wuchereria bancrofti* isolates, implicated in resistance to benzimidazoles in other nematodes. Acta Trop. 88, 77–81.
Hoti, S.L., Thangadurai, R., Dhamodharan, R., Das, P.K., 2008. Genetic heterogeneity of *Wuchereria bancrofti* populations at spatially hierarchical levels in Pondicherry and surrounding areas, south India. Infect. Genet. Evol. 8, 644–652.
Hoti, S.L., Dhamodharan, R., Subramaniyan, K., Das, P.K., 2009. An allele specific PCR assay for screening for drug resistance among *Wuchereria bancrofti* populations in India. Indian J. Med. Res. 130, 193–199.
Hudson, R.R., 2002. Ms a program for generating samples under neutral models. Bioinformatics 18, 337–338.
Köser, C.U., Ellington, M.J., Cartwright, E.J.P., Gillespie, S.H., Brown, N.M., Farrington, M., Holden, M.T.G., Dougan, G., Bentley, S.D., Parkhill, J., Peacock, S.J., 2012. Routine use of microbial whole genome sequencing in diagnostic and public health microbiology. PLoS Pathog. 8, e1002824.
Kumar, N.P., Patra, K., Hoti, S., Das, P., 2002. Genetic variability of the human filarial parasite, *Wuchereria bancrofti* in South India. Acta Trop. 82, 67–76.
Laney, S.J., Ramzy, R.M.R., Helmy, H.H., Farid, H.A., Ashour, A.A., Weil, G.J., Williams, S.A., 2010. Detection of *Wuchereria bancrofti* L3 larvae in mosquitoes: a reverse transcriptase PCR assay evaluating infection and infectivity. PLoS Negl. Trop. Dis. 4, e602.
Lemey, P., Pybus, O.G., Rambaut, A., Drummond, A.J., Robertson, D.L., Roques, P., Worobey, M., Vandamme, A.-M., 2004. The molecular population genetics of HIV-1 group O. Genetics 167, 1059–1068.
Li, H., Durbin, R., 2011. Inference of human population history from individual whole-genome sequences. Nature 475, 493–496.
Liu, G.H., Shao, R., Li, J.-Y., Zhou, D.-H., Li, H., Zhu, X.-Q., 2013. The complete mitochondrial genomes of three parasitic nematodes of birds: a unique gene order and insights into nematode phylogeny. BMC Genomics 14, 414.

Lizotte, M., Supali, T., 1994. A polymerase chain reaction assay for the detection of *Brugia malayi* in blood. Am. J. Trop. Med. Hyg., 314–321.
Lubega, G., Prichard, R., 1990. Specific interaction of benzimidazole anthelmintics with tubulin: high-affinity binding and benzimidazole resistance in *Haemonchus contortus*. Mol. Biochem. Parasitol., 221–232.
Magiorkinis, G., Sypsa, V., Magiorkinis, E., Paraskevis, D., Katsoulidou, A., Belshaw, R., Fraser, C., Pybus, O.G., Hatzakis, A., 2013. Integrating phylodynamics and epidemiology to estimate transmission diversity in viral epidemics. PLoS Comput. Biol. 9, e1002876.
McNulty, S.N., Mitreva, M., Weil, G.J., Fischer, P.U., 2013. Inter and intra-specific diversity of parasites that cause lymphatic filariasis. Infect. Genet. Evol. 14, 137–146.
McReynolds, L., 1986. Cloning and comparison of repeated DNA sequences from the human filarial parasite *Brugia malayi* and the animal parasite *Brugia pahangi*. Proc. Natl. Acad. Sci. 83, 797–801.
Meirmans, P.G., Van Tienderen, P.H., 2004. GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. Mol. Ecol. Notes 4, 792–794.
Meredith, S.E., Lando, G., Gbakima, A.A., Zimmerman, P.A., Unnasch, T.R., 1991. Onchocerca volvulus: application of the polymerase chain reaction to identification and strain differentiation of the parasite. Exp. Parasitol. 73, 335–344.
Michael, E., Malecela-Lazaro, M., 2004. Mathematical modelling and the control of lymphatic filariasis. Lancet. Infect. Dis 44, 223–234.
Michael, E., Malecela-Lazaro, M.N., Kabali, C., Snow, L.C., Kazura, J.W., 2006. Mathematical models and lymphatic filariasis control: endpoints and optimal interventions. Trends Parasitol. 22, 226–233.
Molnar, R.I., Bartelmes, G., Dinkelacker, I., Witte, H., Sommer, R.J., 2011. Mutation rates and intraspecific divergence of the mitochondrial genome of *Pristionchus pacificus*. Mol. Biol. Evol. 28, 2317–2326.
Opgen-Rhein, R., Fahrmeir, L., Strimmer, K., 2005. Inference of demographic history from genealogical trees using reversible jump Markov chain Monte Carlo. BMC Evol. Biol. 5, 6.
Palmieri, J., Purnomo, 1980. Filarid parasites of South Kalimantan (Borneo) Indonesia. *Wuchereria kalimantani* sp. n. (Nematoda: Filarioidea) from the silvered leaf monkey, Presbytis cristatus Eschscholtz. J. Parasitol.
Patterson, N., Richter, D., Gnerre, S., Lander, E., Reich, D., 2006. Genetic evidence for complex speciation of humans and chimpanzees. Nature.
Peters, B.A., Kermani, B.G., Sparks, A.B., Alferov, O., Hong, P., Alexeev, A., Jiang, Y., Dahl, F., Tang, Y.T., Haas, J., Robasky, K., Zaranek, A.W., Lee, J.-H., Ball, M.P., Peterson, J.E., Perazich, H., Yeung, G., Liu, J., Chen, L., Kennemer, M.I., Pothuraju, K., Konvicka, K., Tsoupko-Sitnikov, M., Pant, K.P., Ebert, J.C., Nilsen, G.B., Baccash, J., Halpern, A.L., Church, G.M., Drmanac, R., 2012. Accurate whole-genome sequencing and haplotyping from 10 to 20 human cells. Nature 487, 190–195.
Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945–959.
Progress Report 2000–2009 and Strategic Plan 2010–2020 of the Global Programme to Eliminate Lymphatic Filariasis: halfway towards eliminating lymphatic filariasis., 2010. World Health Organization, Geneva.
Pybus, O.G., Charleston, M.A., Gupta, S., Rambaut, A., Holmes, E.C., Harvey, P.H., 2001. The epidemic behavior of the hepatitis C virus. Science 292, 2323.
Ramesh, A., Small, S.T., Kloos, Z.A., Kazura, J.W., Nutman, T.B., Serre, D., Zimmerman, P.A., 2012. The complete mitochondrial genome sequence of the filarial nematode *Wuchereria bancrofti* from three geographic isolates provides evidence of complex demographic history. Mol. Biochem. Parasitol. 183, 32–41.
Ramzy, R., Helmy, H., Faris, R., 1995. Evaluation of a recombinant antigen-based antibody assay for diagnosis of Bancroftian filariasis in Egypt. Ann. Trop. Med. Parasitol. 89, 443–446.
Rasmussen, D.A., Boni, M.F., Koelle, K., 2014. Reconciling phylodynamics with epidemiology: the case of dengue virus in southern Vietnam. Mol. Biol. Evol. 31, 258–271.
Redman, E., Grillo, V., Saunders, G., Packard, E., Jackson, F., Berriman, M., Gilleard, J.S., 2008. Genetics of mating and sex determination in the parasitic nematode *Haemonchus contortus*. Genetics 180, 1877–1887.
Reimer, L.J., Thomsen, E.K., Tisch, D.J., Henry-Halldin, C.N., Zimmerman, P.A., Baea, M.E., Dagoro, H., Susapu, M., Hetzel, M.W., Bockarie, M.J., Michael, E., Siba, P.M., Kazura, J.W., 2013. Insecticidal bed nets and filariasis transmission in Papua New Guinea. N. Engl. J. Med. 369, 745–753.
Sakthidevi, M., Prabhu, P.R., Chowdhary, S., Hoti, S.L., Kaliraj, P., 2013. Functional analysis of genetic polymorphism in *Wuchereria bancrofti* glutathione S-transferase antioxidant gene: impact on protein structure and enzyme catalysis. Mol. Biochem. Parasitol. 192, 10–20.
Schwab, A.E., Boakye, D.A., Kyelem, D., Prichard, R.K., 2005. Detection of benzimidazole resistance-associated mutations in the filarial nematode *Wuchereria bancrofti* and evidence for selection by albendazole and ivermectin combination treatment. Am. J. Trop. Med. Hyg. 73, 234–238.
Slater, H.C., Gambhir, M., Parham, P.E., Michael, E., 2013. Modelling co-infection with malaria and lymphatic filariasis. PLoS Comput. Biol. 9, e1003096.
Small, S.T., Ramesh, A., Bun, K., Reimer, L., Thomsen, E., Baea, M., Bockarie, M.J., Siba, P., Kazura, J.W., Tisch, D.J., Zimmerman, P.A., 2013. Population genetics of the filarial worm *wuchereria bancrofti* in a post-treatment region of Papua New Guinea: insights into diversity and life history. PLoS Negl. Trop. Dis. 7, e2308.
Srinivasan, L., Mathew, N., Muthuswamy, K., 2009. In vitro antifilarial activity of glutathione S-transferase inhibitors. Parasitol. Res., 1179–1182.
Stadler, T., Kühnert, D., 2013. Birth–death skyline plot reveals temporal changes of epidemic spread in HIV and hepatitis C virus (HCV). Proc. Natl. Acad. Sci. 110, 228–233.
Steel, C., Kubofcik, J., Ottesen, E.A., Nutman, T.B., 2012. Antibody to the filarial antigen Wb123 reflects reduced transmission and decreased exposure in children born following single mass drug administration (MDA). PLoS Negl. Trop. Dis. 6, e1940.
Stein, L.D., Bao, Z., Blasiar, D., Blumenthal, T., Brent, M.R., Chen, N., Chinwalla, A., Clarke, L., Clee, C., Coghlan, A., Coulson, A., D'Eustachio, P., Fitch, D.H.A., Fulton, L.A., Fulton, R.E., Griffiths-Jones, S., Harris, T.W., Hillier, L.W., Kamath, R., Kuwabara, P.E., Mardis, E.R., Marra, M.A., Miner, T.L., Minx, P., Mullikin, J.C., Plumb, R.W., Rogers, J., Schein, J.E., Sohrmann, M., Spieth, J., Stajich, J.E., Wei, C., Willey, D., Wilson, R.K., Durbin, R., Waterston, R.H., 2003. The genome sequence of *Caenorhabditis briggsae*: a platform for comparative genomics. PLoS Biol. 1, E45.
Steinauer, M.L., Blouin, M.S., Criscione, C.D., 2010. Applying evolutionary genetics to schistosome epidemiology. Infect. Genet. Evol. 10, 433–443.
Thangadurai, R., Hoti, S.L., Kumar, N.P., Das, P.K., 2006. Phylogeography of human lymphatic filarial parasite, *Wuchereria bancrofti* in India. Acta Trop. 98, 297–304.
Tibayrenc, M., Ayala, F., 1991. Towards a population genetics of microorganisms: the clonal theory of parasitic protozoa. Parasitol. Today 7, 228–232.
Tisch, D.J., Bockarie, M.J., Dimber, Z., Kiniboro, B., Tarongka, N., Hazlett, F.E., Kastens, W., Alpers, M.P., Kazura, J.W., 2008. Mass drug administration trial to eliminate lymphatic filariasis in Papua New Guinea: changes in microfilaremia, filarial antigen, and Bm14 antibody after cessation. Am. J. Trop. Med. Hyg. 78, 289–293.
Volkman, S.K., Ndiaye, D., Diakite, M., Koita, O.A., Nwakanma, D., Daniels, R.F., Park, D.J., Neafsey, D.E., Muskavitch, M.A.T., Krogstad, D.J., Sabeti, P.C., Hartl, D.L., Wirth, D.F., 2012a. Application of genomics to field investigations of malaria by the international centers of excellence for malaria research. Acta Trop. 121, 324–332.
Volkman, S.K., Neafsey, D.E., Schaffner, S.F., Park, D.J., Wirth, D.F., 2012b. Harnessing genomics and genome biology to understand malaria biology. Nat. Rev. Genet. 13, 315–328.
Wakeley, J., King, L., Low, B.S., Ramachandran, S., 2012. Gene genealogies within a fixed Pedigree, and the robustness of Kingman's Coalescent. Genetics 190, 1433–1445.
Weil, G., Curtis, K., Fischer, P., Won, K., 2011. A multicenter evaluation of a new antibody test kit for lymphatic filariasis employing recombinant *Brugia malayi* antigen Bm-14. Acta Trop. 120, 1–9.
Weil, G.J., Curtis, K.C., Fakoli, L., Fischer, K., Gankpala, L., Lammie, P.J., Majewski, A.C., Pelletreau, S., Won, K.Y., Bolay, F.K., Fischer, P.U., 2013. Laboratory and field evaluation of a new rapid test for detecting *Wuchereria bancrofti* antigen in human blood. Am. J. Trop. Med. Hyg. 89, 11–15.
Williams, S.A., Laney, S.J., Bierwert, L.A., Saunders, L.J., Boakye, D.A., Fischer, P., Goodman, D., Helmy, H., Hoti, S.L., Vasuki, V., Lammie, P.J., Plichart, C., Ramzy, R.M.R., Ottesen, E.A., 2002. Development and standardization of a rapid, PCR-based method for the detection of *Wuchereria bancrofti* in mosquitoes, for xenomonitoring the human prevalence of bancroftian filariasis. Ann. Trop. Med. Parasitol. 96 (Suppl. 2), S41–S46.
World Health Organization Global Programme to Eliminate Lymphatic Filariasis: Progress Report on Mass Drug Administration, 2012. In: Weekly Epidemiological Record 37, pp. 345–356.
Ypma, R.J.F., Bataille, A.M.A., Stegeman, A., Koch, G., Wallinga, J., van Ballegooijen, W.M., 2012. Unravelling transmission trees of infectious diseases by combining genetic and epidemiological data. Philos. Trans. R. Soc. B 279, 444–450.
Ypma, R.J.F., van Ballegooijen, W.M., Wallinga, J., 2013. Relating phylogenetic trees to transmission trees of infectious disease outbreaks. Genetics 195, 1055–1062.
Zhong, M.I.N., Mccarthy, J., Bierwert, L., Williams, S.A., 1996. A polymerase chain reaction assay for detection of the parasite *Wuchereria bancrofti* in human blood samples. Am. J. Trop. Med. Hyg. 54, 357–363.
Zimmerman, P.A., Dadzie, K.Y., De Sole, G., Remme, J., Alley, E.S., Unnasch, T.R., 1992. Onchocerca volvulus DNA probe classification correlates with epidemiologic patterns of blindness. J. Infect. Dis. 165, 964–968.
Zimmerman, P.A., Toe, L., Unnasch, T.R., 1993. Design of Onchocerca DNA probes based upon analysis of a repeated sequence family. Mol. Biochem. Parasitol. 58, 259–267.