International Journal for Parasitology 51 (2021) 137–147

Contents lists available at ScienceDirect

International Journal for Parasitology

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

Genetic epidemiology of lymphatic filariasis in American Samoa after mass drug administration

Shannon M. Hedtke [a,b,*,1], Patsy A. Zendejas-Heredia [a,1], Patricia M. Graves [c], Sarah Sheridan [d], Meru Sheel [e], Saipale D. Fuimaono [f], Colleen L. Lau [d], Warwick N. Grant [a,b]

[a] Department of Animal, Plant and Soil Sciences, La Trobe University, Bundoora, Victoria, Australia
[b] Department of Physiology, Anatomy and Microbiology, La Trobe University, Bundoora, Victoria, Australia
[c] College of Public Health, Medical and Veterinary Sciences, James Cook University, Cairns, Queensland, Australia
[d] Department of Global Health, Research School of Population Health, The Australian National University, Acton, Australian Capital Territory, Australia
[e] National Centre for Epidemiology and Population Health, Research School of Population Health, The Australian National University, Acton, Australian Capital Territory, Australia
[f] American Samoa Department of Health, Pago Pago, American Samoa

ARTICLE INFO

Article history:
Received 18 June 2020
Received in revised form 5 August 2020
Accepted 9 August 2020
Available online 7 November 2020

Keywords:
Lymphatic filariasis
American Samoa
Wuchereria bancrofti
Population genetics
Mass drug administration

ABSTRACT

Over 892 million people in 48 countries are at risk of infection by nematodes that cause lymphatic filariasis. As part of the Global Programme to Eliminate Lymphatic Filariasis, mass drug administration is distributed to communities until surveillance indicates infection rates are below target prevalence thresholds. In some countries, including American Samoa, lymphatic filariasis transmission persists despite years of mass drug administration and/or has resurged after cessation. Nothing is known about the population genetics of *Wuchereria bancrofti* worms in Polynesia, or whether local transmission is persisting and/or increasing due to inadequate mass drug administration coverage, expansion from residual hotspots, reintroduction from elsewhere, or a combination. We extracted DNA from microfilariae on blood slides collected during prevalence surveys in 2014 and 2016, comprising 31 pools of five microfilariae from 22 persons living in eight villages. We sequenced 1104 bp across three mitochondrial markers (ND4, COI, CYTB). We quantified parasite genetic differentiation using variant calls and estimated haplotypes using principal components analysis, F-statistics, and haplotype networks. Of the variants called, all but eight were shared across the main island of Tutuila, and three of those were from a previously described hotspot village, Fagali'i. Genotypic data did not support population genetic structure among regions or villages in 2016, although differences were observed between worms collected in Fagali'i in 2014 and those from 2016. Because estimated haplotype frequency varied between villages, these statistics suggested genetic differentiation, but were not consistent among villages. Finally, haplotype networks demonstrated American Samoan sequence clusters were related to previously published sequences from Papua New Guinea. These are, to our knowledge, the first reports of *W. bancrofti* genetic variation in Polynesia. The resurgent parasites circulating on the main island of American Samoa represent a single population. This study is the first step towards investigating how parasite population structure might inform strategies to manage resurgence and elimination of lymphatic filariasis.

1. Introduction

Lymphatic filariasis (LF) is caused by three species of nematodes transmitted by mosquitoes: *Wuchereria bancrofti*, *Brugia malayi*, and *Brugia timori*. In 2018, the World Health Organization (WHO) estimated that globally over 892 million people were infected or at risk of infection (World Health Organization, 2019). The Global Programme to Eliminate Lymphatic Filariasis aims to eliminate LF as a public health problem using two strategies: (i) interrupt transmission by conducting mass drug administration (MDA), and (ii) morbidity management and disability prevention for people with chronic complications such as lymphedema and scrotal hydroceles (World Health Organization, 2012). During MDA, at-risk populations are treated annually with a variety of two- or three-drug combinations of albendazole, ivermectin, and/or diethylcarbamazine until the prevalence of infection falls below target thresholds (World Health Organization, 2011a).

* Corresponding author at: Department of Physiology, Anatomy and Microbiology, La Trobe University, Bundoora, Victoria, Australia..
E-mail address: S.Hedtke@latrobe.edu.au (S.M. Hedtke).
[1] These authors contributed equally.

https://doi.org/10.1016/j.ijpara.2020.08.009
0020-7519/

Humans serve as the definitive host for adult nematodes, which can survive in the lymphatic system for many years. Females produce microfilariae (mf) which circulate in the blood. Transmission occurs after a mosquito ingests mf during a blood meal; the mf develop into L3s in the mosquito and migrate to its head, where they are transmitted to another human during subsequent mosquito feeding (Cross, 1996). During LF surveys, the prevalence of infection is measured using rapid diagnostic tests which detect circulating filarial antigen produced by adult worms. Antigen prevalence is used to assess whether parasite prevalence in the host community has dropped below the threshold considered sufficiently low for interruption of transmission (World Health Organization, 2011b), but filarial antigens circulate even after adult worms have died, and can persist for an unknown period of months or years after successful treatment. Only a proportion of antigen-positive individuals have active transmissible infection in which mf can be detected on blood slides or by DNA-based diagnostics (Bisht et al., 2006). Although LF elimination programs have successfully treated >910 million people and have delivered 7.7 billion treatments since 2000 (World Health Organization, 2018), global elimination goals continue to be challenged by ongoing transmission, resurgence, and the persistence of localised hotspots in some areas despite completing the recommended rounds of MDA (Lau et al., 2014, 2017; Rao et al., 2014; Kouassi et al., 2015). One critical question is whether, in a given community, resurgence is driven by migrant parasites introduced by travellers from neighbouring regions with ongoing transmission, or whether the circulation of local parasites had not been successfully interrupted by MDA.

In American Samoa, MDA for LF was conducted from 2000 to 2006 under the Global Programme to Eliminate LF. WHO-recommended transmission assessment surveys (TAS) in young children passed target thresholds in 2011 (TAS-1) and 2015 (TAS-2) (Coutts et al., 2017; Won et al., 2018). However, seroprevalence studies of LF post-MDA in American Samoa in 2010 and 2014 found two potential hotspots of ongoing transmission (Lau et al., 2014, 2017). In addition, a PCR-based molecular xenomonitoring study of mosquitoes confirmed widespread ongoing transmission across most of American Samoa (Schmaedick et al., 2014). In 2016, American Samoa failed TAS-3, and surveys of school children and community members identified antigen-positive people in all age groups, and confirmed widespread resurgence (Sheel et al., 2018).

Comparison of seroprevalence data from 2010, 2014, and 2016 suggests that the geographic distribution of antigen-positive people might be expanding over time (Lau et al., 2014, 2017; Sheel et al., 2018). Possible explanations for this continuing increase include: (i) expansion of hotspots through local movement of people and vectors within American Samoa, (ii) a time delay in establishment of the parasite from widespread, low-intensity infections across the island (i.e., current surveillance strategies might not be sufficiently sensitive at detecting low-level transmission), or (iii) introduction of migrant parasites from other countries by travellers. This last hypothesis of multiple re-introductions is plausible given the high volume of travellers between American Samoa and neighbouring Samoa (Xu et al., 2018b; Graves et al., 2020), where there is also ongoing LF transmission (Hapairai et al., 2015). The regrettable resurgence of LF transmission in a small, well-characterised population in the islands of American Samoa provides an ideal opportunity to investigate genetic diversity and population structure of filarial parasites as a means of determining possible sources of transmission.

Population genetics has been used to explore epidemiology and transmission dynamics of many human parasites, including those that cause malaria (Ariey et al., 2003; Fola et al., 2018), schistosomiasis (Prugnolle et al., 2005; Rudge et al., 2009), strongyloidiasis (Kikuchi et al., 2016; Nagayasu et al., 2017), and onchocerciasis (Choi et al., 2017; Doyle et al., 2017). In LF, studies on genetic diversity of parasites collected from India (Thangadurai et al., 2006; Hoti et al., 2008; Mahakalkar et al., 2017), Papua New Guinea (PNG) (Small et al., 2013, 2016), Ghana (de Souza et al., 2014), or sampled globally (Small et al., 2019) have suggested that (i) genetic differentiation arises in LF when there is no or reduced movement of parasites between regions (Thangadurai et al., 2006; Nuchprayoon et al., 2007; Dhamodharan et al., 2008; de Souza et al., 2014), (ii) shared genetic variants indicate historic or continuing migration of parasites between two areas (Small et al., 2019), and (iii) genetic diversity is higher in areas where transmission is high (Hoti et al., 2008) and is reduced in areas where transmission is low, such as after MDA (Small et al., 2013, 2014). Small et al. (2013) found that mitochondrial sequence data from the cytochrome oxidase I locus (COI) collected from parasites from several communities in PNG indicated high levels of genetic diversity. This research suggests that genetic diversity of mitochondrial sequences could, in turn, be used to identify transmission zones and other parameters of interest to control programs (Hedtke et al., 2020), provided the analytical methods are robust (Small et al., 2014).

To date, there have been no published LF DNA sequences (mitochondrial or other) from Polynesia. DNA-based detection methods for LF have been widely used in French Polynesia (Plichart et al., 2006; Gass et al., 2012; Plichart and Lemoine, 2013), American Samoa (Schmaedick et al., 2014), Samoa (Hapairai et al., 2015), and Tuvalu (Gass et al., 2012), but these have relied on amplification of conserved repetitive sequences. The only population genetic analysis of LF parasites from the Pacific was performed on samples from PNG in Melanesia (Small et al., 2013, 2016), although this focused on a relatively small region of the country. Based on genomic diversity in a small sample of LF parasites from PNG, Asia, and Africa, Small et al. (2019) hypothesised that two waves of human migration imported LF to PNG: one from the South Asian ancestor of PNG and Australian indigenous populations, and the second from an Austronesian admixture originating around Taiwan. Current opinion about Polynesian peoples suggests initial migration from a Taiwanese source, with later admixture due to migration from Melanesia (Skoglund et al., 2016). However, how the history of human migration might have influenced genetic variation of LF worms across Polynesia is unknown, although it is reasonable to assume the worms were carried with migrating humans.

Our research goal is to utilise the rich epidemiological data on American Samoa, including serological prevalence (Sheel et al., 2018), high resolution data on residential and work locations (Xu et al., 2018a), population mobility (Xu et al., 2018b), and post-MDA surveillance surveys (Lau et al., 2014, 2017; Sheel et al., 2018), to determine whether parasite population genetic analysis can answer questions relevant for LF epidemiology. In this study, we tested the utility of mf on blood slides collected during surveys in American Samoa for sequencing genetic markers for parasite population genetic analysis. We hypothesise that if the spatial clustering of antigen-positive persons and infected mosquito pools was the result of multiple, recent independent introductions of migrant parasites, then the parasite population genetic diversity will be similarly clustered. Alternatively, parasite population genetic analyses could indicate that parasites came from a single source population, which would occur if (i) the resurgence was the result of widespread dissemination following a single introduction event from another LF endemic region, or if (ii) MDA did not successfully interrupt transmission and a panmictic local parasite population rebounded after cessation of MDA.

2. Materials and methods

2.1. Ethics approval

Ethics clearance for collection and analysis of parasite materials from humans used in this project were given by the Department of Health, American Samoa (IRB00001249, FWA0024252), by the Human Research Ethics Committee at The Australian National University (2016/482), and by the SHE College Human Ethics Sub-Committee at La Trobe University, Australia.

2.2. Parasite collection and extraction

Parasite samples were obtained from blood slides prepared from antigen-positive people identified during multiple surveys in American Samoa: (i) a 2014 study that included community members from suspected hotspots and a population-representative sample of adults from a workplace and a clinic (Lau et al., 2017), (ii) a 2016 survey of 6–7 year old children from 29 elementary schools across American Samoa (Sheel et al., 2018), and (iii) a 2016 community-based survey of people aged ≥8 years from 28 villages across American Samoa (Fig. 1; Sheel et al., 2018).

For DNA extraction, lysis solution was prepared fresh immediately prior to use by combining 98.5 μL Direct PCR lysis reagent and 1.5 μL of 20 mg/ml proteinase K (Viagen Biotech, Los Angeles, California, USA or Roche Holding AG, Basel, Switzerland). mf were floated from positive blood smears in 2 μL of HPLC-grade water and picked using an eyelash into a PCR tube containing 20 μL of the lysis solution. These were incubated at 55 °C for 16 h, followed by heat inactivation of the proteinase K at 85 °C for 1 h. The resulting worm lysate was diluted to 1:10 by adding 180 μL of HPLC-grade water, and 2 μL of the diluted worm lysate was used as a template in subsequent PCRs.

2.3. PCR protocol development

Our expectation was that DNA from slide-dried mf would be highly degraded, so our goal was to target relatively short fragments of DNA for amplification and sequencing. We designed three primer sets for *W. bancrofti* for markers commonly used in population genetics using Primer-BLAST (Ye et al., 2012) based on a mitochondrial genome sequence from PNG (GenBank accession no. HQ184469; (Ramesh et al., 2012)) and modified with Illumina adapters for library construction: cytochrome oxidase I (COI; 402 bp), ND4 (353 bp), and cytochrome b (CYTB; 349 bp) (Supplementary Table S1).

To explore the feasibility of amplifying DNA for sequencing, we compared the relative efficacy of quantitative PCR (qPCR) on DNA extractions from single mf and from pools in which 2, 5, 10, or 20 mf were picked into the same tube. We used two protocols: first, we used the LDR primers from Rao et al. (2006), which target a high-copy, nuclear repeat region designed for sensitive detection of *W. bancrofti* DNA; second, we used the ND4 primers designed here for conventional PCR and sequencing. All qPCR assays were run on a CFX96 Real-Time System (Bio-Rad Laboratories, Hercules, California, USA), with an initial denaturing step of 3 min at 98 °C followed by 45 cycles of (98 °C 10 s, 58.3 °C (LDR)/57 °C (ND4) 15 s, 72 °C 15 s, read plate). Reactions contained 5 μL of SsoAdvanced Universal SYBR Green master mix (Bio-Rad Laboratories), 5 μL each of 10 μM forward and reverse primers, and 2 μL of 1:10 DNA extract, brought to a final volume of 10 μL. For LDR primers, we ran two replicates of each DNA extraction, and for ND4 primers we ran three replicates. Relative amplification of products over the course of each qPCR run and overall statistics were assessed using the CFX Maestro software (Bio-Rad Laboratories).

For conventional PCR to prepare amplicons for sequencing, in each reaction, 2 μL of diluted crude extract were added to 2.5 μL of 10× Immobuffer (Bioline, Alexandria, New South Wales, Australia), 10 μM forward primer, 10 μM reverse primer, 2 mM dNTPs, 50 mM $MgCl_2$, 0.2 μL of Immolase DNA (Bioline), and HPLC water for a total volume of 20 μL. Reactions were run on a TaKaRa PCR thermocycler (TaKaRa Bio, Mountain View, California, USA), for 8 min at 95 °C, 35× cycles of (1 min 94 °C, 1 min 57 °C, 1 min 72 °C), and a final extension of 7 min at 72 °C. PCR products were run on a 1% agarose gel containing GelRed nucleic acid stain (Biotium, San Francisco, California, USA) and visualised on a Gel Doc EZ imager using Image Lab software (Bio-Rad Laboratories) to ensure sufficient amplification for downstream library construction and sequencing.

2.4. Next generation sequencing

A total of 155 mf in 31 pools of five mf each from 22 people across eight villages were selected for deep sequencing (Fig. 1; Table 1). PCR amplicons were purified with an ISOLATE II PCR and Gel kit (Bioline) following the manufacturer's instructions. Each amplicon pool was quantified on a Qubit™ fluorometer using a Qubit™ dsDNA HS assay kit (ThermoFisher Scientific, Massachusetts, USA) and the three amplicons for each 5 mf DNA pool were combined such that each amplicon was present in equal concentrations. This 5 mf, 3 amplicon pool was then combined in equal concentration with amplicons from several unrelated studies on other species and indexed using a second, short round of PCR, as described in McCann et al. (2019). Each DNA pool of five extracted mf thus had a unique sequencing index that was identical across amplicons. Because amplicon sequences are expected to have low sequence diversity, and because this unbalanced nucleotide composition can negatively impact clustering, the libraries were spiked with 20% PhiX (Illumina, Inc., San Diego, California, USA) prior to sequencing. Sequencing was performed on an Illumina MiSeq using v3.0 2 × 8300 bp sequencing chemistry at the La Trobe Institute for Molecular Science (LIMS), Australia.

Reads were trimmed to remove any adapter and primer sequences, using Trimmomatic v0.32 in the paired-end, palindrome mode, with a minimum length of 200 bp and a minimum quality score of 30 (Bolger et al., 2014). Reads were aligned against the *W. bancrofti* mitochondrial genome (GenBank accession number **HQ184469**) using BWA v0.7.12 (Li and Durbin, 2009) with the BWA-MEM algorithm. GATK v3.4_46 was used for local realignment around indels according to GATK Best Practices recommendations (van der Auwera et al., 2013).

2.5. Single nucleotide polymorphism (SNP) calling and analysis

Single nucleotide polymorphisms (SNPs) were called using freebayes v1.0.2 (https://github.com/ekg/freebayes), assuming five genomes per pool and a minimum mapping quality of 20. We used these variant calls to calculate allele frequencies per locus for each pool of five mf. We then performed a principal components analysis (PCA) on allele frequencies using the R package adegenet (Jombart, 2008; Jombart and Ahmed, 2011). PCA was used to identify whether there were multiple genetic clusters in the data, and to determine whether clusters, if they existed, contained only worms from the same geographic area. The advantage of using allele frequencies is that information about parasite population genetic diversity across amplicons can be analysed, increasing the signal per population. However, the disadvantage is that it ignores any information from haplotypes (i.e., which variants belong to the same individual worm).

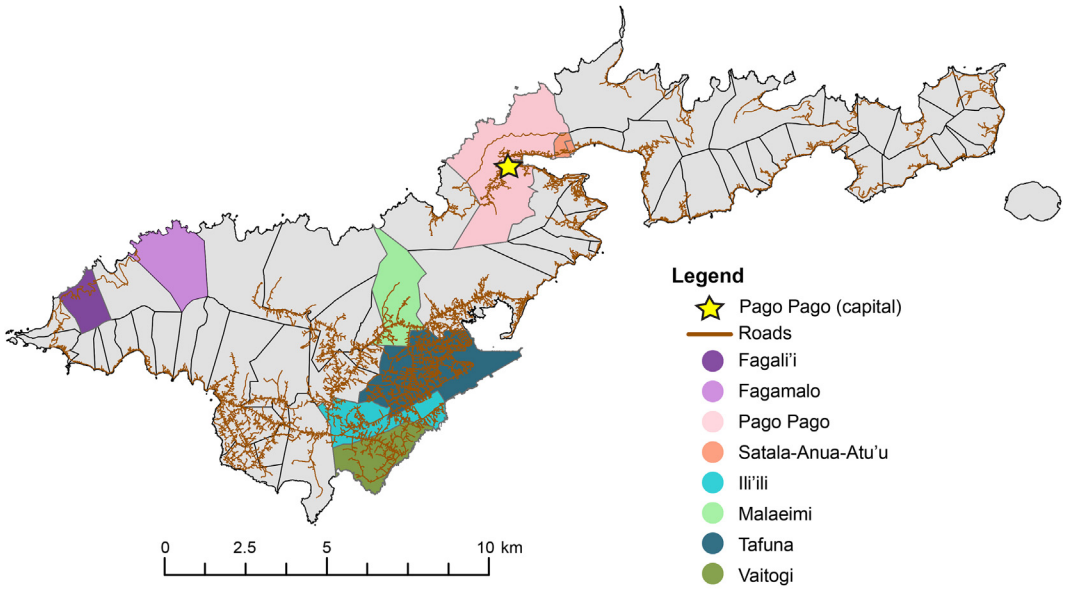

Fig. 1. Sampling locations across American Samoa.

Table 1
Sampling location and number of microfilariae extracted from blood slides collected in American Samoa from 22 people in eight villages. Microfilariae were picked from slides prepared in both 2014 and 2016 from Fagali'i, and from 2016 for the remaining villages. For downstream analyses, villages were classified as belonging either to the northwest, central, or south of Tutuila. Because the three villages of Satala, Anua, and Atu'u have small populations and are very close geographically, they were combined for analysis in Sheel et al. (2018) and in this study.

Region	Village	Person	Year of Collection	Number of mf Picked
North West	Fagali'i	1	2014	10
		2	2016	5
		3	2016	5
		4	2016	10
		5	2016	5
		6	2016	5
		7	2016	5
		8	2016	5
		9	2016	5
		10	2016	5
	Fagamalo	11	2016	15
		12	2016	5
Subtotal	2 villages	**12**		**80**
Central	Pago Pago	13	2016	5
	Satala/Anua/Atu'u	14	2016	5
Subtotal	2 villages	**2**		**10**
South	Ili'ili	15	2016	10
		16	2016	5
	Malaeimi	17	2016	10
	Tafuna	18	2016	15
	Vaitogi	19	2016	10
		20	2016	5
		21	2016	5
		22	2016	5
Subtotal	4 villages	**8**		**65**
Total	8 villages	**22**		**155**

2.6. Haplotype reconstruction and analysis

In addition to exploring genetic structure using variant calls from pools of worms, we reconstructed haplotypes for each worm from the pooled sequencing libraries bioinformatically. The challenge is to differentiate PCR/sequencing errors from biological diversity. To estimate haplotypes per worm we used SeekDeep (Hathaway et al., 2018), developed to infer haplotypes from pooled sequences of malaria. Since our expectation is that each mf should be represented equally, and thus be represented in relatively high frequency in the sequenced reads, we set a cut-off that each haplotype used in downstream analyses needed to be present at a frequency of at least 10% in the pool of five mf. We further used the frequency of each haplotype to estimate the number of worms in each pool of five that would have each sequence.

Haplotypes were aligned for each amplicon and examined visually using Mesquite v3.2 (https://www.mesquiteproject.org/). Arlequin v3.5.2.2 (Excoffier and Lischer, 2010) was used to estimate genetic diversity statistics per village, Tajima's D (Tajima, 1989), and Fu's Fs (Fu, 1997), which test for demographic changes and/or selection, assessing *P* values using the infinite site model for 10,000 permutations. Population genetic differentiation was estimated in Arlequin using both an analog of Wright's F_{ST} (Wright, 1943), ϕ_{ST} (Excoffier et al., 1992), and population pairwise difference (corrected using within-population variation), with significance based on 10,000 permutations. A three-way analysis of molecular variance (AMOVA) (Excoffier et al., 1992) among regions (Northwest, Central, South), villages, and individual people was estimated using Arlequin to explore whether genetic variation had hierarchical structure correlated with geography, with the probability that estimated variance differed significantly from zero assessed using 10,000 permutations. A two-way AMOVA among villages and people was calculated in R v2.6.1 (https://www.r-project.org/) using the package poppr (Kamvar et al., 2014) and implementations of AMOVA from ade4 (Dray and Dufour, 2007) and pegas (Paradis, 2010). To visualise relationships among haplotypes, median-joining haplotype networks (Bandelt et al., 1999) were constructed using PopART (Leigh and Bryant, 2015).

3. Results

3.1. qPCR quantification

Our first test was to determine the efficiency of DNA extraction from slide-dried mf. Because DNA concentrations were too low for standard methods of quantification such as Qubit® or nanodrop, we used the LDR primers developed for sensitive *W. bancrofti* detection (Rao et al., 2006). We confirmed that LF DNA was present in the extractions, and further that adding multiple mf to the DNA extraction buffer ("pooling") increased DNA quantity and thus amplification success (represented by fewer cycles required for detection of a real signal, i.e., lower C_q values; Fig. 2). There was no amplification in no-template (negative) controls.

Because the LDR marker is both short (90 bp) and is present in thousands of copies throughout the genome, it is well-suited for detection of the target species (Rao et al., 2006). Population genetics, however, requires single-copy sequences of sufficient length to be informative about demographic history and genetic structure. As a result, we also used qPCR to test the amplification success of one of our mitochondrial amplicons (ND4) with the goal of identifying the optimal strategy for amplification and sequencing of these larger products. As with the LDR primers, increasing the number of mf (and thus, the concentration of DNA) in the extraction pool decreased C_q (Fig. 2), although with greater variability between replicates. A target value of C_q of less than ~35 is considered optimal (i.e., fewer than 35 amplification cycles are required for the sample to pass the threshold of detection), while C_q values >40 can result in increased error rates and non-specific amplification, although these values should not be considered absolute cutoffs (Burns and Valdivia, 2008; Bustin et al., 2009). The results of the ND4 qPCR indicate that the concentration of DNA in extractions of only one mf is not suitable for amplification of targets of ~350 bp, while pools of >5 mf successfully amplify the target marker in fewer than 35 cycles (although with higher variation between replicates).

When products from conventional PCR of each of the three amplicons designed here were visualised on a 1% agarose gel, amplicons from one mf could not be detected, while those from five mf were detectable (data not shown). We decided that deep sequencing of pools of five mf would balance PCR efficiency with minimising the number of *W. bancrofti* per pool, with the understanding that genetic diversity in any given pool could be underestimated for each amplicon given the stochastic nature of the amplification process.

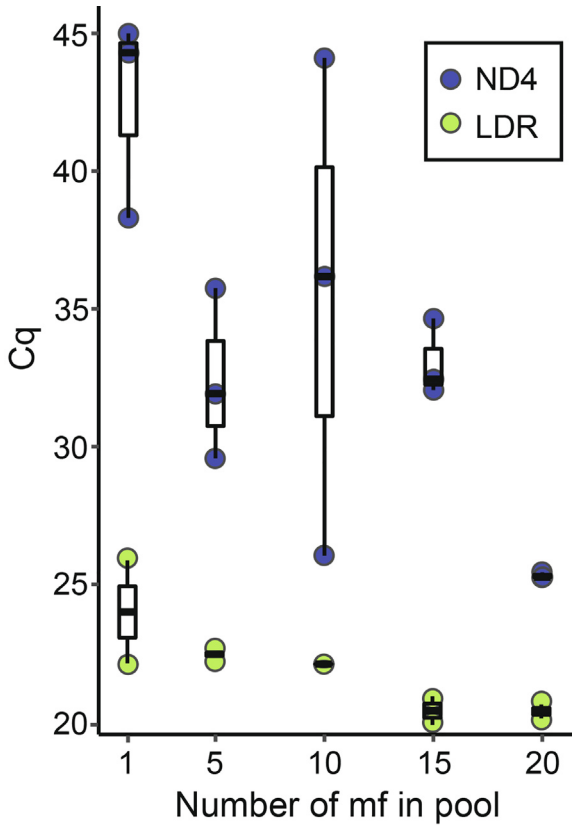

Fig. 2. Results of quantitative PCR for two primer sets using DNA extracted from pools of 1, 5, 10, 15, or 20 *Wuchereria bancrofti* microfilariae. Each point represents the cycle quantification, C_q, value, which is the cycle number at which a sample's fluorescence reaches the threshold for real signal. Smaller C_q values indicate earlier amplification, which can indicate higher concentrations of template DNA. Two runs are represented here: one targeting a nuclear repeat called LDR and the other targeting a portion of a mitochondrial gene ND4. Two replicates per sample were performed for LDR, while three were performed for ND4.

3.2. Sample DNA pooling and deep sequencing

A total of 31 pools of five mf, representing 155 mf from 22 people, were extracted and sequenced across all three amplicons with an average read depth of ~ 6,000, resulting in an average of 100% coverage at a minimum read depth of 20 for each amplicon (coverage for each pool is reported in Supplementary Table S2). The data sets supporting the conclusions of this study are available in the NCBI Sequence Read Archive, under BioProject **PRJNA610651** (https://www.ncbi.nlm.nih.gov/).

3.3. SNP calling and analysis

In any given amplicon from any given pool, there is no guarantee that amplification of a particular mf was represented equally or even at all, and thus we used multiple approaches to identify and analyse genetic variation. Using the SNP caller freebayes, 26 positions were identified as polymorphic. mf from the 2014 samples from Fagali'i had three polymorphisms not found in any of the 2016 Fagali'i samples. Samples from Fagamalo, Malaeimi, and Satala-Anua-Atu'u each had a SNP variant only found in that

community, and those from Vaitogi had two variants not found in other communities.

Based on the allele frequencies of each pool, PCA did not identify clusters amongst mf collected in 2016. There was, however, significant differentiation by year: mf sequenced from Fagali'i collected in 2014 were discriminated by the first PC (Fig. 3), explaining nearly 29% of the genetic variation. However, the remaining PCs (not shown) did not discriminate among villages; i.e., the variation observed was not visibly correlated with geography.

3.4. Haplotype calling and analysis

Unique haplotypes were identified in Fagali'i 2014 (ND4, COI), Fagamalo (ND4), Malaeimi (COI), Satala-Anua-Atu'u (COI), and Vaitogi (COI, CYTB). A common haplotype was shared across all locations for each amplicon, and this common haplotype tended to be at high frequency or fixed (Fig. 4A).

The average pairwise difference, which represents counts of differences in genetic variants between pairs of sequences, was highest at the COI locus (Fig. 4B). Haplotype diversity, which can range from 0 for no diversity and 1 for every haplotype in the population as unique, was high (Fig. 4C) compared with the genetic or nucleotide diversity (Fig. 4D; Supplementary Table S3). Nucleotide diversity, which can range from 0 for no differentiation and 1 for polymorphisms at every position on the sequence, was estimated to be low (Fig. 4D). However, neither Tajima's D or Fu's Fs were significant (Supplementary Table S3), and thus there is no indication in the data of recent demographic changes in parasite population sizes or of recent strong selection affecting the genetic diversity of these markers.

Two-way and three-way hierarchical AMOVAs were performed on each mitochondrial marker to examine the spatial genetic partitioning across the island. The three-way AMOVA significantly attributed most of the parasite genetic variation to variation between people within villages, with an additional component of variation attributed to variation between villages within regions. Negative variance components were calculated among regions for each marker (Table 2), suggesting that grouping parasites by region was not supported by the data. The two-way AMOVA, which removes the geographic grouping by region, was consistent with the three-way, indicating that the majority of genetic variation was found within villages, with a smaller component of between-village variation.

Genetic differentiation between parasites collected from villages was estimated using F_{ST} and pairwise genetic differentiation (Table 3). No measures of differentiation were consistently significant across markers or approaches to measuring differentiation (Table 3), suggesting these data were insufficient to resolve population structure at the level of the island.

Median-joining haplotype networks suggest genetic admixture across all villages in American Samoa, because there was no clustering in the network by village or region (Fig. 5A–C). The network for each marker visualises a haplotype shared across villages as a large circle, with only a few base pair differences between different haplotypes. The greater haplotype diversity of COI was reflected in the larger number of nodes in the haplotype network (Fig. 5B).

An additional COI network was produced using sequences from both American Samoa and PNG (Small et al., 2013) (Fig. 5D). This visualisation suggests that parasites from each country were not well differentiated from each other, and that the diversity in COI captured in this study is part of a larger subset of *W. bancrofti* diversity within Melanesia and Polynesia.

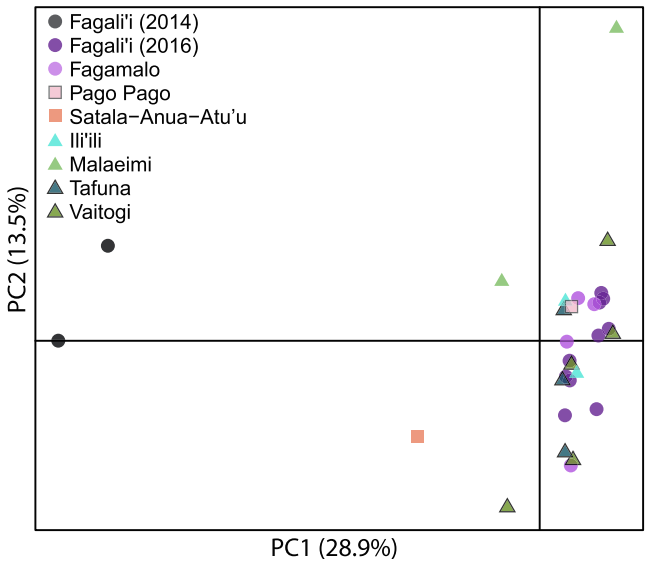

Fig. 3. First and second principal components (representing 42.4% of genetic variation) from a principal component analysis of three markers sequenced from *Wuchereria bancrofti* microfilariae (mf) collected from eight villages in American Samoa. Each circle represents a pool of five mf, with colours representing different villages, and some points have been jittered for ease of visualisation. Shapes indicate geographic regions: Northwest (circles), Fagali'i and Fagamalo; Central (squares), Pago Pago and Satala-Anua-Atu'u; South (triangles), Ili'ili, Malaeimi, Tafuna, and Vaitogi.

4. Discussion

The distribution of variation observed in our data suggest that *W. bancrofti* from American Samoa in 2016 form a single genetic population. Methods used to estimate geographic structure (haplotype networks, F_{ST}, and AMOVA) suggest that mitochondrial sequences were shared across the island with no to low correlations with geographic location (Tables 2 and 3, Figs. 3 and 5). The star-shaped haplotype networks reported here are consistent with an expansion of *W. bancrofti* across the Pacific islands concurrent with historic human migration (Skoglund et al., 2016).

MDA was distributed in American Samoa over the years 2000–2006 (King et al., 2011) and might have been expected to dramatically reduce the number of infections and thus parasite genetic diversity (i.e., force the population of parasites through a population bottleneck). Parasite population bottlenecks more severely reduce nucleotide diversity than haplotype diversity because single mutations add new haplotypes while contributing less to overall genetic diversity. Tajima's D and Fu's Fs are statistics which compare nucleotide and haplotype diversity to test for selection or changes in parasite population size, and neither were significant in our analyses. However, Small et al. (2013) calculated that populations undergoing MDA required >90% annual reduction in the number of adult worms for five consecutive years in order to observe a statistically significant positive Tajima's D value. In the case of American Samoa, MDA coverage ranged from 49% to 70% between 2001 and 2006 (Burkot et al., 2002; World Health Organization, 2006; Liang et al., 2008; King et al., 2011). Thus, the parasite population might not have sufficiently contracted during MDA for Tajima's D or Fu's Fs to reach statistically significant values, despite evidence from TAS that parasite reduction was substantial. In addition, small sample sizes of parasites per village could underestimate haplotype diversity, because rare variants would be less likely to be detected, reducing the power of test statistics such as Tajima's D. Finally, stochastic variation during PCR of pooled samples could cause failure to amplify haplotypes

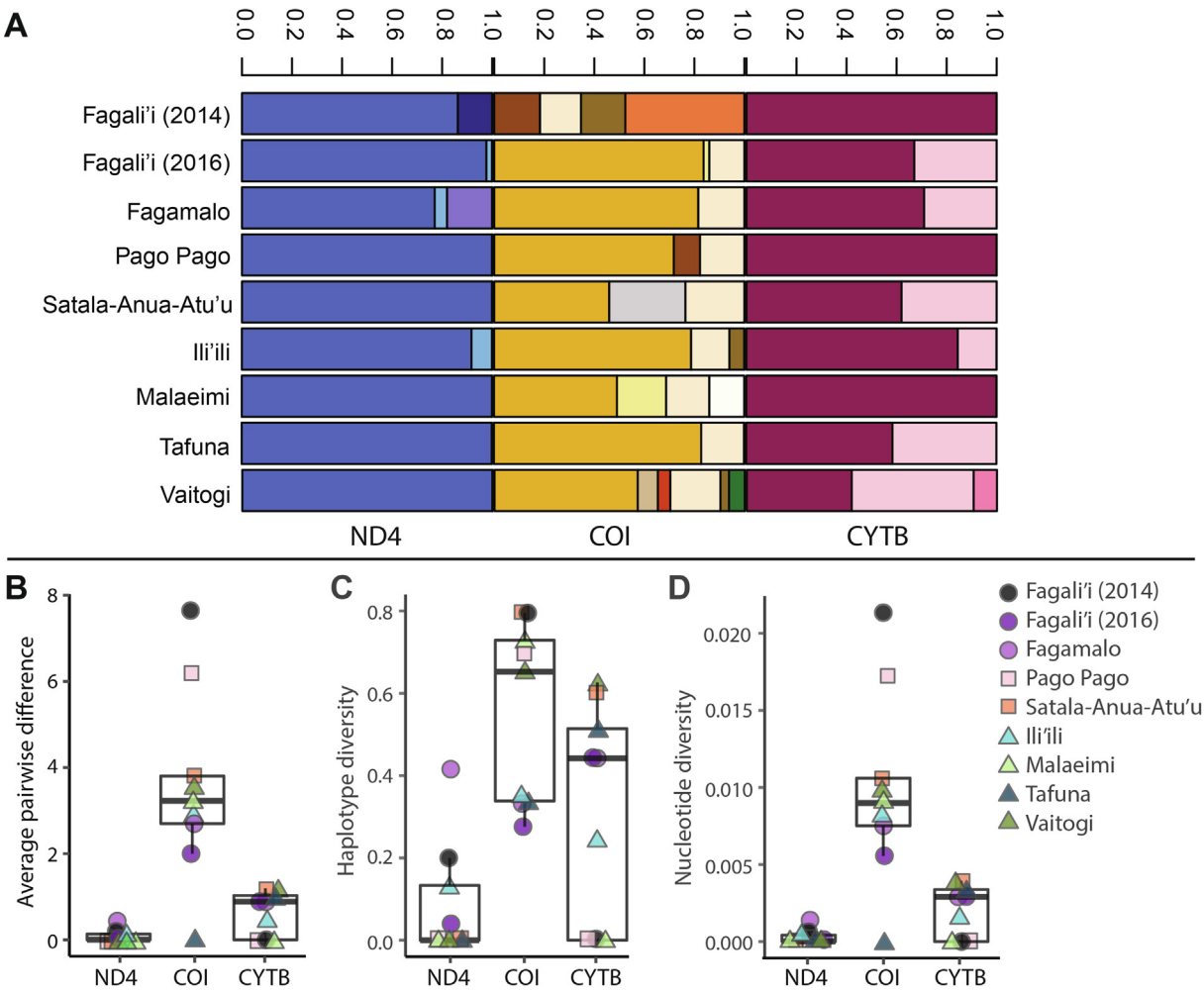

Fig. 4. Genetic diversity of *Wuchereria bancrofti* for eight villages in American Samoa. (A) Frequencies of haplotypes estimated by SeekDeep across three markers: ND4, COI, and CYTB. (B) Average pairwise number of nucleotide differences; (C) haplotype diversity (H_d); and (D) nucleotide/genetic diversity. Each point represents the value for a village; standard boxplots indicate the mean (solid line), 25–75% range (box: first through third quartiles), and the value within 1.5 times interquartile range (whisker). See Supplementary Table S3 for details.

Table 2
Analysis of molecular variance (AMOVA) for three mitochondrial markers sequenced from *Wuchereria bancrofti* collected across three geographic regions of American Samoa. Northwest region: villages Fagali'i and Fagamalo; Central: Pago Pago and Satala-Anua-Atu'u; South: Ili'ili, Malaeimi, Tafuna and Vaitogi.

Marker	Source of Variation	Percentage of Variation (3-way)	*P* (rand value > obs value)	Percentage of Variation (2-way)
ND4	Among regions	−2.96	0.737 ± 0.012	
	Among villages within regions	8.35	0.013 ± 0.003	2.61
	Among people within villages	94.61	0.042 ± 0.006	97.39
COI	Among regions	−4.51	0.685 ± 0.015	
	Among villages within regions	9.87	0.009 ± 0.003	1.14
	Among people within villages	94.63	0.013 ± 0.004	98.86
CYTB	Among regions	−6.81	0.767 ± 0.014	
	Among villages within regions	15.84	0.000 ± 0.000	9.29
	Among people within villages	90.97	0.006 ± 0.002	90.70

from one or more worms sampled from a person, further reducing the ability of these tests to infer changes in parasite population size.

The major axis of genetic differentiation in American Samoa was between samples taken in 2014 from Fagali'i and the samples collected across the main island in 2016 (Fig. 3). This observation could be driven in part by random sampling error causing failure to collect haplotypes that were actually present in the population in 2014, as the number of parasites sequenced from 2014 samples was relatively small (10) and were from only a single person. In comparison, in 2016, there were nine people and 50 mf sampled from Fagali'i, and the genetic variants estimated from those 2016 Fagali'i parasites were similar to those found elsewhere in the island. However, not all sequences found in the 2014 parasites were detected in 2016, despite the larger sample size. This finding could be explained by two hypotheses; the first being the rapid spread of LF from Fagali'i as an epidemiological hotspot between 2014 and 2016 (Lau et al., 2017). If Fagali'i were the major source

Table 3
Genetic differentiation using F_{ST} (below the diagonal) and pairwise differentiation (above the diagonal). Measures for each of three mitochondrial markers are given; bold indicates statistically significant differentiation ([a]$P < 0.05$; [b]$P < 0.01$ based on 10,000 permutations).

	Marker	Fagali'i 2014	Fagali'i 2016	Faga-malo	Pago Pago	Satala-Anua-Atu'u	Ili'ili	Malae-imi	Vaitogi	Tafuna
Fagali'i 2014	ND4		0.000	0.032	0.000	0.000	0.000	0.000	0.000	0.000
	COI		**1.377**[b]	**1.030**[a]	−0.322	0.878	**0.961**[a]	0.827	**0.844**[a]	**2.378**[b]
	CYTB		0.196	0.158	0.000	0.200	0.019	0.000	**0.613**†	0.286
Fagali'i 2016	ND4	0.091		**0.024**[b]	0.000	0.000	−0.003	0.000	0.000	0.000
	COI	**0.359**[b]		−0.061	−0.101	−0.213	−0.085	−0.088	0.023	0.143
	CYTB	**0.176**[a]		−0.030	0.196	−0.116	0.044	0.196	**0.092**[a]	−0.030
Fagamalo	ND4	0.066	**0.164**[b]		0.032	0.032	0.005	0.032	**0.032**[a]	0.032
	COI	**0.213**[a]	−0.023		−0.287	−0.287	−0.151	−0.178	−0.097	0.253
	CYTB	0.184	−0.035		0.158	−0.122	0.017	0.158	0.099	−0.036
Pago Pago	ND4	−0.084	−0.110	−0.007		0.000	0.000	0.000	0.000	0.000
	COI	−0.059	0.110	−0.004		−0.440	−0.304	−0.411	−0.330	0.500
	CYTB	0.000	0.122	0.108		0.200	0.019	0.000	**0.613**[a]	0.286
Satala- Anua- Atu'u	ND4	−0.084	−0.110	−0.007	0.000		0.000	0.000	0.000	0.000
	COI	0.094	−0.023	−0.075	−0.096		−0.304	−0.331	−0.250	0.100
	CYTB	0.417	−0.106	−0.116	0.250		0.006	0.200	−0.083	−0.154
Ili'ili	ND4	0.007	−0.008	0.007	−0.099	−0.099		0.000	0.000	0.000
	COI	**0.180**[a]	−0.030	−0.057	−0.020	−0.084		−0.195	−0.108	0.210
	CYTB	0.033	0.042	0.019	−0.043	0.042		0.019	**0.334**[b]	**0.091**[a]
Malaeimi	ND4	0.000	−0.051	0.064	0.000	0.000	−0.029		0.000	0.000
	COI	0.132	−0.018	−0.061	−0.065	−0.097	−0.067		−0.157	**0.289**[a]
	CYTB	0.000	**0.176**[a]	0.184	0.000	0.417	0.033		**0.613**[b]	0.286
Vaitogi	ND4	0.102	−0.015	**0.149**[a]	0.000	0.000	0.036	0.000		0.000
	COI	**0.171**[a]	0.015	−0.033	−0.029	−0.069	−0.037	−0.051		**0.450**[a]
	CYTB	**0.403**[b]	**0.089**[a]	0.085	0.339	−0.073	**0.260**[a]	**0.403**[b]		0.003
Tafuna	ND4	0.043	−0.031	0.100	0.000	0.000	0.000	0.000	0.000	
	COI	**0.456**[b]	0.058	0.129	0.405	**0.303**[b]	0.126	**0.215**[b]	**0.151**[a]	
	CYTB	0.300	−0.030	−0.039	0.211	−0.154	0.107	0.300	0.001	

for the resurgence in island-wide transmission, only a subset of Fagali'i diversity would have been transmitted, and genetic diversity would be lost as the population expanded. A second hypothesis is that Fagali'i had persistent transmission that maintained local diversity, but was subsequently invaded by parasites whose populations were rapidly expanding from elsewhere. Sequencing and analysis of *W. bancrofti* from additional time points—either from pre-MDA materials that have been retained, or from future collection efforts—would be valuable, both to increase the overall parasite sample size and to better quantify effects of interventions or of demographic changes on parasite genetic variation.

We observed common haplotypes in all regions sampled (Figs. 4 and 5). There are two possible explanations for this observation: (i) genetic similarities among villages were maintained by transmission across the island (i.e., gene flow) and/or (ii) genetic variation was shaped by previous parasite migration that occurred before MDA (i.e., retained ancestral polymorphisms).

Because the flight range of *Aedes polynesiensis,* the main mosquito vector species in American Samoa, is likely to be relatively short (~100 m) (Verdonschot and Besse-Lototskaya, 2014; Hapairai et al., 2015) compared with the distance across the island, people are most likely to be responsible for carrying parasites between regions and villages. Surveys have indicated that individual people move around the island for work, leisure, worship, and family, and this could allow transmission from one village to another (Xu et al., 2018b). Transmission through the movement of people throughout American Samoa would lead to genetic homogeneity of parasites across the island, as we have observed.

Alternatively, or in conjunction with parasite movement, low association of genetic variation with geography may be a characteristic of *W. bancrofti* across the Pacific as a whole, as the result of an evolutionarily recent expansion (Small et al., 2019). A widespread, rapid expansion of a parasite with large population sizes could result in ancestral haplotypes from the initial founding population being retained over time, and therefore, these similar haplotypes would be found throughout the island. Comparison of COI sequence data with those collected in PNG (Small et al., 2013) indicates that much of the variation detected in American Samoa in this locus is shared with parasites from elsewhere in the Pacific (Fig. 5D), suggesting that the potential for retention of ancestral polymorphisms is high. Data collected from additional worms across the Asia-Pacific region would help in determining which genetic variation is population-specific.

We were not able to distinguish whether the resurgence of transmission was driven by local parasites within American Samoa expanding from residual infections, or by migrant parasites from neighbouring LF-endemic countries. With the first scenario, if the threshold for stopping MDA was too high for this intervention area, or if surveys used to estimate prevalence of infection were not sensitive enough to detect low levels of transmission, then post-MDA transmission could increase over time. In the second scenario, people regularly travel between American Samoa and Samoa (Xu et al., 2018b; Graves et al., 2020), another LF-endemic country with a much larger geographic area and population (~196,100 in Samoa compared with ~55,500 in American Samoa); the high intensity of travel could cause sufficient movement of parasites that the two countries might represent a single parasite population. If resurgence from local parasites alone were responsible for sustaining transmission hotspots then the existing criteria for passing TAS might need to be carefully reconsidered, particularly given the results of research indicating that surveys that include older children and adults might be more sensitive at detecting residual infection hotspots than those that only include 6–7 year old children (Sheel et al., 2018; Lau et al., 2020). However, if hotspots were the result of migrant parasites, then local surveys might need to be conducted more frequently to detect parasite reintroductions which could re-start the transmission cycle, or MDA needs to be continued until elimination thresholds have been reached in both

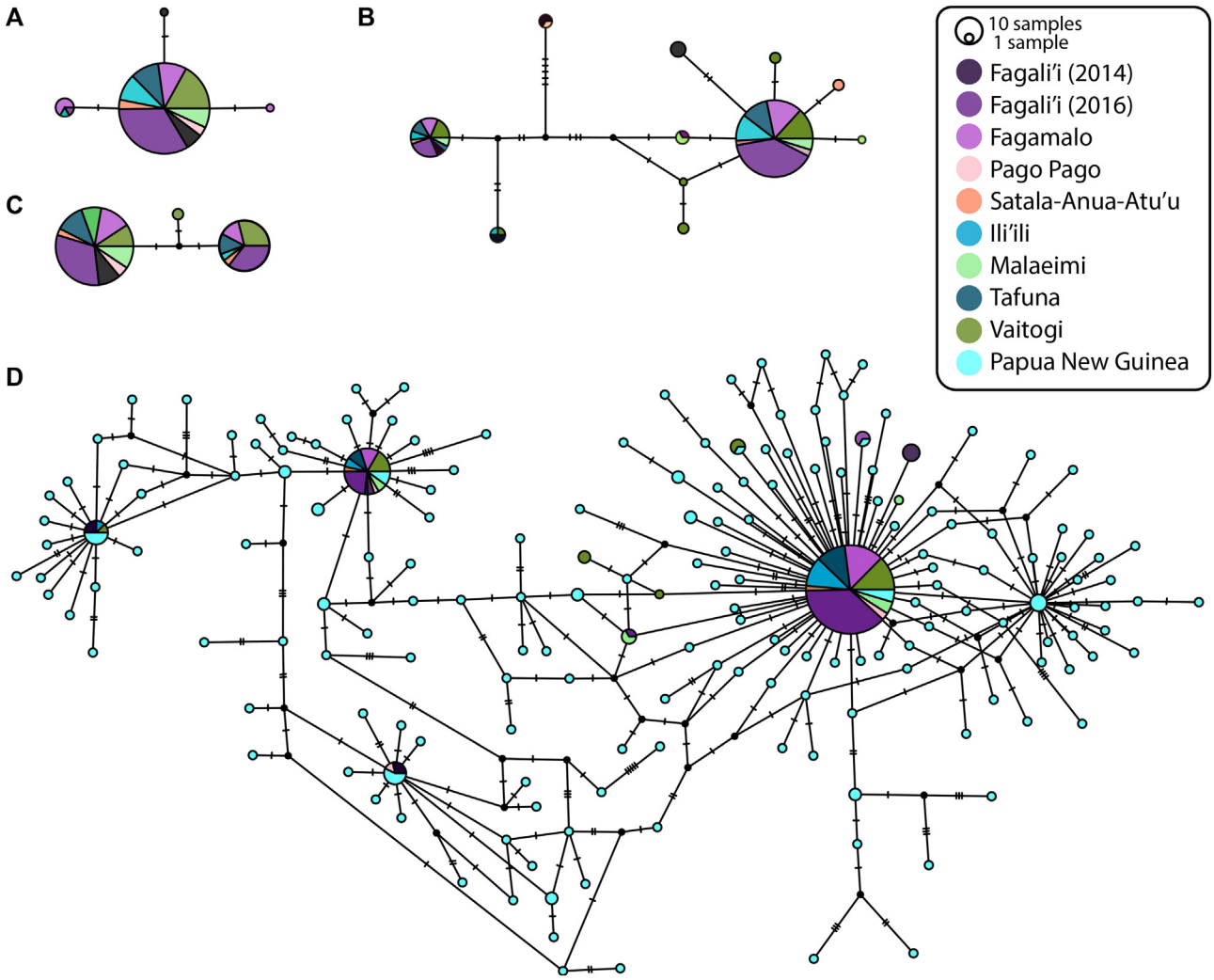

Fig. 5. Haplotype networks estimated from *Wuchereria bancrofti* microfilariae sequences from eight villages in American Samoa. Each circle represents a haplotype. The size of the circle is proportional to the number of mf with that haplotype, with the proportion found in each population indicated by colour. A hatch mark across a line represents a single bp difference between haplotypes. (A) Haplotype network based on sequence from the mitochondrial ND4 marker; (B) haplotype network based on sequence from the COI mitochondrial marker; (C) haplotype network based on sequence from the CYTB mitochondrial marker; (D) haplotype network based on haplotypes estimated from COI sequence data from American Samoa mf and sequences from *W. bancrofti* collected in Papua New Guinea (Small et al., 2013).

the target population (American Samoa) and any likely source of migrant parasites (i.e., Samoa in this case). Finally, it is of course possible for both scenarios to occur concurrently.

Assessing the risks of parasite reintroduction to elimination goals based on genetic analyses would require additional sequencing from nearby LF-endemic countries in the Pacific, particularly Samoa, French Polynesia, Tonga, and Fiji. However, the informativeness of the markers used here was relatively limited for distinguishing among worms from American Samoa and those from PNG. This suggests that more sequence data per marker may be necessary to identify genetic variation unique to particular geographic areas, which could then be used to identify any migrant worms. The advantage of using blood slides as a source for *W. bancrofti* DNA is that blood collection and slide preparation are already incorporated into program surveillance and monitoring activities for LF. The disadvantage is that mf from blood slides have limited DNA quality and quantity, which introduces challenges in amplification of longer sequences. One alternative would be to include collection and adequate preservation of immature worms during surveillance activities, e.g., by taking additional blood samples from antigen-positive people (Small et al., 2019) or infected mosquitoes from xenomonitoring. Extractions from these samples would return higher quality DNA from which sequences of more informative markers (e.g., whole mitochondrial genomes or highly variable nuclear markers) could be used in population genetic analyses.

This study provides the first known data on LF parasite genetic diversity in Polynesia. The aim of this research was to develop methods to investigate the likely origins of parasites responsible for post-MDA resurgence. We have demonstrated that sequencing from microfilarial DNA extracted from blood slides is achievable.

Population genetic analyses support the hypothesis that American Samoa likely represents a single transmission zone, through which parasites potentially move and interbreed in the long term. Our study was not able to determine the source of these parasites, but results are consistent with the hypothesis that movement of humans within the country could homogenise parasite populations and re-initiate transmission (from localised hotspots) in areas where it had stopped. Additional sampling will be necessary to draw conclusions on the effects of migration between villages and potentially among neighbouring countries. In moving towards this goal, we recommend a focus on improving methods to increase the success rate of single-worm sequencing as an alternative to the pooled-worm sequencing used here.

Because the ultimate goal is elimination of LF, epidemiological models that explicitly consider the extent of parasite mixing within countries and the potential for re-introduction from neighboring countries might provide valuable information for consideration of alternative treatment and assessment strategies (Xu et al., 2018a). Integration of parasite genetic variation with demographic and epidemiological data can improve models of disease dynamics to assist with decisions about when to stop MDA and how to best assess transmission post-MDA.

Acknowledgements

For assistance with sample collection and analysis, the authors wish to thank Motusa Tuliema Nua, Joseph Tufa, Margaret Sesepasara, Mary Matai'a, and staff (American Samoa Department of Health); Amor Gonzales and Ivy Clemente (LBJ Tropical Medical Center, American Samoa); Jeff Warner (James Cook University, Australia). For local support and technical advice, we would like to thank Mark Schmaedick (American Samoa Community College); Katie Gass, Patrick Lammie (Task Force for Global Health, USA); Kimberly Won (Division of Parasitic Diseases and Malaria, U.S. Centers for Disease Control and Prevention, USA); the American Samoa Departments of Samoan Affairs and of Education; the village chiefs and mayors; Starkist tuna cannery, American Samoa; and the numerous field and data assistants in American Samoa (particularly Paeae Sakalaia and Tapakea Tufono) and at James Cook University. For advice and assistance with genetic work and analyses, the authors additionally wish to thank Kirsty McCann and Katie Crawford at La Trobe University, Australia and Karen McCulloch at the University of Melbourne, Australia. We are also grateful for the support received from the National Institutes of Health/National Institute of Allergy and Infectious Diseases, USA, Centers for Disease Control and Prevention, World Health Organization, Switzerland, James Cook University, and La Trobe University. Funding for sequencing was provided by a grant to WNG from the UNICEF/UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases, Switzerland (TDR; B80153 and B80149), a La Trobe Asia Research Grant to SMH, a La Trobe University Understanding Diseases Research Focus Area Start-Up Grant to SMH, and from PRISM^2, Australia, to SMH and CLL. CLL was supported by an Australian National Health and Medical Research Council (NHMRC) Fellowship (1109035). The 2014 survey was funded by grants from the Australian Institute of Tropical Health and Medicine (AITHM; 13122014) and the Faculty of Medicine and Biomedical Sciences at the University of Queensland, Australia (2127835). The 2016 survey received financial support from the Coalition for Operational Research on Neglected Tropical Diseases (COR-NTD) which is funded at the Task Force for Global Health primarily by the Bill & Melinda Gates Foundation, USA, the United Kingdom Department for International Development, and by the United States Agency for International Development through its Neglected Tropical Diseases Program.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijpara.2020.08.009.

References

Ariey, F., Duchemin, J.-B., Robert, V., 2003. Metapopulation concepts applied to falciparum malaria and their impacts on the emergence and spread of chloroquine resistance. Infect. Genet. Evol. 2 (3), 185–192. https://doi.org/10.1016/S1567-1348(02)00099-0.

Bandelt, H.J., Forster, P., Rohl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. 16 (1), 37–48. https://doi.org/10.1093/oxfordjournals.molbev.a026036.

Bisht, R., Hoti, S.L., Thangadurai, R., Das, P.K., 2006. Isolation of *Wuchereria bancrofti* microfilariae from archived stained blood slides for use in genetic studies and amplification of parasite and endosymbiont genes. Acta Trop. 99 (1), 1–5. https://doi.org/10.1016/j.actatropica.2005.12.009.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120.

Burkot, T.R., Taleo, G., Toeaso, V., Ichimori, K., 2002. Progress towards, and challenges for, the elimination of filariasis from Pacific-island communities. Ann. Trop. Med. Parasitol. 96 (8), 61–69. https://doi.org/10.1179/000349802125002419.

Burns, M., Valdivia, H., 2008. Modelling the limit of detection in real-time quantitative PCR. Eur. Food Res. Technol. 226 (6), 1513–1524. https://doi.org/10.1007/s00217-007-0683-z.

Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., Vandesompele, J., Wittwer, C.T., 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin. Chem. 55, 611–622.

Choi, Y.-J., Tyagi, R., McNulty, S.N., Rosa, B.A., Ozersky, P., Martin, J., Hallsworth-Pepin, K., Unnasch, T.R., Norice, C.T., Nutman, T.B., Weil, G.J., Fischer, P.U., Mitreva, M., 2017. Genomic diversity in *Onchocerca volvulus* and its *Wolbachia* endosymbiont. Nat. Microbiol. 2 (2). https://doi.org/10.1038/nmicrobiol.2016.207.

Coutts, S.P., King, J.D., Pa'au, M., Fuimaono, S., Roth, J., King, M.R., Lammie, P.J., Lau, C. L., Graves, P.M., 2017. Prevalence and risk factors associated with lymphatic filariasis in American Samoa after mass drug administration. Trop. Med. Health 45 (1). https://doi.org/10.1186/s41182-017-0063-8.

Cross, J.H., 1996. Filarial nematodes. In: Baron, S. (Ed.), Medical Microbiology. University of Texas Medical Branch at Galveston, Galveston, TX.

de Souza, D.K., Osei-Poku, J., Blum, J., Baidoo, H., Brown, C.A., Lawson, B.W., Wilson, M.D., Bockarie, M.J., Boakye, D.A., 2014. The epidemiology of lymphatic filariasis in Ghana, explained by the possible existence of two strains of *Wuchereria bancrofti*. Pan Afr. Med. J. 17, 133.

Dhamodharan, R., Das, M.K., Hoti, S.L., Das, P.K., Dash, A.P., 2008. Genetic variability of diurnally sub-periodic *Wuchereria bancrofti* in Nicobarese tribe of Nicobar group of Islands, Andaman and Nicobar Islands, India. Parasitol. Res. 103 (1), 59–66. https://doi.org/10.1007/s00436-008-0927-2.

Doyle, S.R., Bourguinat, C., Nana-Djeunga, H.C., Kengne-Ouafo, J.A., Pion, S.D.S., Bopda, J., Kamgno, J., Wanji, S., Che, H., Kuesel, A.C., Walker, M., Basáñez, M.G., Boakye, D.A., Osei-Atweneboana, M.Y., Boussinesq, M., Prichard, R.K., Grant, W. N., 2017. Genome-wide analysis of ivermectin response by Onchocerca volvulus reveals that genetic drift and soft selective sweeps contribute to loss of drug sensitivity. PLoS Negl. Trop. Dis. 11, e0005816.

Dray, S., Dufour, A., 2007. The ade4 package: implementing the duality diagram for ecologists. J. Stat. Softw. 22, 1–20.

Excoffier, L., Lischer, H.E., 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Resour. 10, 564-567.

Excoffier, L., Smouse, P.E., Quattro, J.M., 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131, 479–491.

Fola, A.A., Nate, E., Abby Harrison, G.L., Barnadas, C., Hetzel, M.W., Iga, J., Siba, P., Mueller, I., Barry, A.E., 2018. Nationwide genetic surveillance of *Plasmodium vivax* in Papua New Guinea reveals heterogeneous transmission dynamics and routes of migration amongst subdivided populations. Infect. Genet. Evol. 58, 83–95. https://doi.org/10.1016/j.meegid.2017.11.028.

Fu, Y.X., 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147, 915–925.

Gass, K., Beau de Rochars, M.V., Boakye, D., Bradley, M., Fischer, P.U., Gyapong, J., Itoh, M., Ituaso-Conway, N., Joseph, H., Kyelem, D., Laney, S.J., Legrand, A.M., Liyanage, T.S., Melrose, W., Mohammed, K., Pilotte, N., Ottesen, E.A., Plichart, C., Ramaiah, K., Rao, R.U., Talbot, J., Weil, G.J., Williams, S.A., Won, K.Y., Lammie, P., 2012. A multicenter evaluation of diagnostic tools to define endpoints for programs to eliminate bancroftian filariasis. PLoS Negl. Trop. Dis. 6.

Graves, P.M., Sheridan, S., Fuimaono, S., Lau, C.L., 2020. Demographic, socioeconomic and disease knowledge factors, but not population mobility, associated with lymphatic filariasis infection in adult workers in American Samoa in 2014. Parasites Vectors 13 (1). https://doi.org/10.1186/s13071-020-3996-4.

Hapairai, L.K., Plichart, C., Naseri, T., Silva, U., Tesimale, L., Pemita, P., Bossin, H.C., Burkot, T.R., Ritchie, S.A., Graves, P.M., Melrose, W., Joseph, H., 2015. Evaluation of traps and lures for mosquito vectors and xenomonitoring of *Wuchereria bancrofti* infection in a high prevalence Samoan Village. Parasites Vectors 8 (1). https://doi.org/10.1186/s13071-015-0886-2.

Hathaway, N.J., Parobek, C.M., Juliano, J.J., Bailey, J.A., 2018. SeekDeep: single-base resolution de novo clustering for amplicon deep sequencing. Nucleic Acids Res 46, e21.

Hedtke, S.M., Kuesel, A.C., Crawford, K.E., Graves, P.M., Boussinesq, M., Lau, C.L., Boakye, D.A., Grant, W.N., 2020. Genomic epidemiology in filarial nematodes: transforming the basis for elimination program decisions. Front. Genet. 10, 1282.

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 (5), 644–652. https://doi.org/10.1016/j.meegid.2008.06.002.

Jombart, T., 2008. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24, 1403–1405.
Jombart, T., Ahmed, I., 2011. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics 27, 3070–3071.
Kamvar, Z.N., Tabima, J.F., Grünwald, N.J., 2014. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ e281.
Kikuchi, T., Hino, A., Tanaka, T., Aung, M.P., Afrin, T., Nagayasu, E., Tanaka, R., Higashiarakawa, M., Win, K.K., Hirata, T., Htike, W.W., Fujita, J., Maruyama, H., 2016. Genome-wide analyses of individual *Strongyloides stercoralis* (Nematoda: Rhabditoidea) provide insights into population structure and reproductive life cycles. PLoS Negl. Trop. Dis. 10, e0005253.
King, J.D., Zielinski-Gutierrez, E., Pa'au, M., Lammie, P., 2011. Improving community participation to eliminate lymphatic filariasis in American Samoa. Acta Trop. 120, S48–S54. https://doi.org/10.1016/j.actatropica.2010.08.021.
Kouassi, B.L., de Souza, D.K., Goepogui, A., Narh, C.A., King, S.A., Mamadou, B.S., Diakité, L., Dadzie, S.K., Boakye, D.A., Utzinger, J., Bockarie, M.J., Koudou, B.G., 2015. Assessing the presence of *Wuchereria bancrofti* in vector and human populations from urban communities in Conakry, Guinea. Parasites Vectors 8 (1). https://doi.org/10.1186/s13071-015-1077-x.
Lau, C.L., Sheel, M., Gass, K., Fuimaono, S., David, M., Won, K., Sheridan, S., Graves, P., 2020. Potential strategies for strengthening surveillance of lymphatic filariasis in American Samoa after mass drug administration: targeting older age groups, hotspots, and household members of infected person. PLoS. Negl. Trop. Dis. in press.
Lau, C.L., Sheridan, S., Ryan, S., Roineau, M., Andreosso, A., Fuimaono, S., Tufa, J., Graves, P.M., 2017. Detecting and confirming residual hotspots of lymphatic filariasis transmission in American Samoa 8 years after stopping mass drug administration. PLoS Negl. Trop. Dis. 11, e0005914.
Lau, C.L., Won, K.Y., Becker, L., Soares Magalhaes, R.J., Fuimaono, S., Melrose, W., Lammie, P.J., Graves, P.M., 2014. Seroprevalence and spatial epidemiology of lymphatic filariasis in American Samoa after successful mass drug administration. PLoS Negl. Trop. Dis. 8, e3297.
Leigh, J.W., Bryant, D., 2015. PopART: Full-feature software for haplotype network construction. Methods Ecol. Evol. 6, 1110–1116.
Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25 (14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324.
Liang, J.L., King, J.D., Ichimori, K., Handzel, T., Pa'au, M., Lammie, P.J., 2008. Impact of five annual rounds of mass drug administration with diethylcarbamazine and albendazole on *Wuchereria bancrofti* infection in American Samoa. Am. J. Trop. Med. Hyg. 78, 924–928.
Mahakalkar, A.L., Sapkal, H.P., Baig, M.M., 2017. Report of high genetic diversity of filarial worm, *Wuchereria bancrofti*, from endemic region of Eastern Maharashtra (India). Helminthologia 54, 292–297.
McCann, K.M., Grant, W.N., Spratt, D.M., Hedtke, S.M., 2019. Cryptic species diversity in ticks that transmit disease in Australia. Int. J. Parasitol.: Parasites Wildl. 10, 125–131. https://doi.org/10.1016/j.ijppaw.2019.08.002.
Nagayasu, E., Aung, M.P.P. Thet Hnin Htwe, Hortiwakul, T., Hino, A., Tanaka, T., Higashiarakawa, M., Olia, A., Taniguchi, T., Win, S.M.T., Ohashi, I., Odongo-Aginya, E.I., Aye, K.M., Mon, M., Win, K.K., Ota, K., Torisu, Y., Panthuwong, S., Kimura, E., Palacpac, N.M.Q., Kikuchi, T., Hirata, T., Torisu, S., Hisaeda, H., Horii, T., Fujita, J., Htike, W.W., Maruyama, H., 2017. A possible origin population of pathogenic intestinal nematodes, *Strongyloides stercoralis*, unveiled by molecular phylogeny. Sci. Rep. 7 (1). https://doi.org/10.1038/s41598-017-05049-x.
Nuchprayoon, S., Junpee, A., Poovorawan, Y., 2007. Random Amplified Polymorphic DNA (RAPD) for differentiation between Thai and Myanmar strains of *Wuchereria bancrofti*. Filaria J 6 (1). https://doi.org/10.1186/1475-2883-6-6.
Paradis, E., 2010. pegas: an R package for population genetics with an integrated-modular approach. Bioinformatics 26, 419–420.
Plichart, C., Lemoine, A., 2013. Monitoring and evaluation of lymphatic filariasis interventions: an improved PCR-based pool screening method for high throughput *Wuchereria bancrofti* detection using dried blood spots. Parasit Vectors 6, 110.
Plichart, C., Sechan, Y., Davies, N., Legrand, A.M., 2006. PCR and dissection as tools to monitor filarial infection of *Aedes polynesiensis* mosquitoes in French Polynesia. Filaria J. 5, 2.
Prugnolle, F., Roze, D., Theron, A., Meeûs, T.D.E., 2005. F-statistics under alternation of sexual and asexual reproduction: a model and data from schistosomes (platyhelminth parasites). Mol Ecol 14, 1355–1365.
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.
Rao, R.U., Atkinson, L.J., Ramzy, R.M., Helmy, H., Farid, H.A., Bockarie, M.J., Susapu, M., Laney, S.J., Williams, S.A., Weil, G.J., 2006. A real-time PCR-based assay for detection of *Wuchereria bancrofti* DNA in blood and mosquitoes. Am. J. Trop. Med. Hyg. 74, 826–832.
Rao, R.U., Nagodavithana, K.C., Samarasekera, S.D., Wijegunawardana, A.D., Premakumara, W.D., Perera, S.N., Settinayake, S., Miller, J.P., Weil, G.J., 2014. A comprehensive assessment of lymphatic filariasis in Sri Lanka six years after cessation of mass drug administration. PLoS Negl. Trop. Dis. 8, e3281.
Rudge, J.W., Lu, D.B., Fang, G.R., Wang, T.P., Basanez, M.G., Webster, J.P., 2009. Parasite genetic differentiation by habitat type and host species: molecular epidemiology of *Schistosoma japonicum* in hilly and marshland areas of Anhui Province, China. Mol. Ecol. 18, 2134–2147.
Schmaedick, M.A., Koppel, A.L., Pilotte, N., Torres, M., Williams, S.A., Dobson, S.L., Lammie, P.J., Won, K.Y., 2014. Molecular xenomonitoring using mosquitoes to map lymphatic filariasis after mass drug administration in American Samoa. PLoS Negl. Trop. Dis. 8, e3087.
Sheel, M., Sheridan, S., Gass, K., Won, K., Fuimaono, S., Kirk, M., Gonzales, A., Hedtke, S.M., Graves, P.M., Lau, C.L., 2018. Identifying residual transmission of lymphatic filariasis after mass drug administration: Comparing school-based versus community-based surveillance - American Samoa, 2016. PLoS Negl. Trop. Dis. 12, e0006583.
Skoglund, P., Posth, C., Sirak, K., Spriggs, M., Valentin, F., Bedford, S., Clark, G.R., Reepmeyer, C., Petchey, F., Fernandes, D., Fu, Q., Harney, E., Lipson, M., Mallick, S., Novak, M., Rohland, N., Stewardson, K., Abdullah, S., Cox, M.P., Friedlaender, F.R., Friedlaender, J.S., Kivisild, T., Koki, G., Kusuma, P., Merriwether, D.A., Ricaut, F.X., Wee, J.T., Patterson, N., Krause, J., Pinhasi, R., Reich, D., 2016. Genomic insights into the peopling of the Southwest Pacific. Nature 538, 510–513.
Small, S.T., Labbé, F., Coulibaly, Y.I., Nutman, T.B., King, C.L., Serre, D., Zimmerman, P. A., 2019. Human migration and the spread of the nematode parasite *Wuchereria bancrofti*. Mol. Biol. Evol. 36, 1931–1941.
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.
Small, S.T., Reimer, L.J., Tisch, D.J., King, C.L., Christensen, B.M., Siba, P.M., Kazura, J. W., Serre, D., Zimmerman, P.A., 2016. Population genomics of the filarial nematode parasite *Wuchereria bancrofti* from mosquitoes. Mol. Ecol. 25, 1465–1477.
Small, S.T., Tisch, D.J., Zimmerman, P.A., 2014. Molecular epidemiology, phylogeny and evolution of the filarial nematode *Wuchereria bancrofti*. Infect Genet Evol 28, 33–43.
Tajima, F., 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123, 585–595.
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.
van der Auwera, G.A., Arneiro, M., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., Banks, E., Garimella, K., Altshuler, D., Gabriel, S., DePristo, M., 2013. From FastQ data to high-confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr. Prot. Bioinform. 43, 11.10.11-11.10.33.
Verdonschot, P.F.M., Besse-Lototskaya, A.A., 2014. Flight distance of mosquitoes (Culicidae): A metadata analysis to support the management of barrier zones around rewetted and newly constructed wetlands. Limnologica 45, 69–79.
Won, K.Y., Robinson, K., Hamlin, K.L., Tufa, J., Seespesara, M., Wiegand, R.E., Gass, K., Kubofcik, J., Nutman, T.B., Lammie, P.J., Fuimaono, S., 2018. Comparison of antigen and antibody responses in repeat lymphatic filariasis transmission assessment surveys in American Samoa. PLoS Negl. Trop. Dis. 12, e0006347.
World Health Organization, 2006. The PacELF Way: Towards the Elimination of Lymphatic Filariasis from the Pacific, 1999–2005. WHO Regional Office for the Western Pacific, Manila.
World Health Organization, 2011a. Managing morbidity and preventing disability in the Global Programme to Eliminate Lymphatic Filariasis: WHO position Statement. Prise en charge de la morbidité et prévention des incapacités dans le Programme mondial pour l'élimination de la filariose lymphatique: declaration de l'OMS. Wkly Epidemiol. Rec. 86, 581–585.
World Health Organization, 2011b. Monitoring and Epidemiological Assessment of Mass Drug Administration: A Manual for National Elimination Programmes. World Health Organization, Geneva, Switzerland.
World Health Organization, 2012. Accelerating Work to Overcome the Global Impact of Neglected Tropical Diseases: a Roadmap for Implementation. World Health Organization, Geneva, Switzerland.
World Health Organization, 2018. Global programme to eliminate lymphatic filariasis: progress report, 2017. Wkly Epidemiol. Rec. 44, 589–604.
World Health Organization, 2019. Global programme to eliminate lymphatic filariasis: progress report, 2018. Wkly Epidemiol. Rec. 94, 457–472.
Wright, S., 1943. Isolation by distance. Genetics 28, 114–138.
Xu, Z., Graves, P.M., Lau, C.L., Clements, A., Geard, N., Glass, K., 2018a. GEOFIL: A spatially-explicit agent-based modelling framework for predicting the long-term transmission dynamics of lymphatic filariasis in American Samoa. Epidemics 27, 19–27.
Xu, Z., Lau, C.L., Zhou, X., Fuimaono, S., Soares Magalhaes, R.J., Graves, P.M., 2018b. The extensive networks of frequent population mobility in the Samoan Islands and their implications for infectious disease transmission. Sci. Rep. 8, 10136.
Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., Madden, T.L., 2012. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC Bioinform. 13, 134.