X-treme loss of sequence diversity linked to neo-X chromosomes in filarial nematodes

Short Title: *Brugia* Genetic Diversity

John Mattick[1,*], Silvia Libro[2], Robin Bromley[1], Wanpen Chaicumpa[3], Matthew Chung[1], Darren Cook[4], Mohammad Behram Khan[5], Nikhil Kfumar[1], Yee-Ling Lau[5], Shailja Misra-Bhattacharya[6], Ramakrishna Rao[7], Lisa Sadzewicz[1], Atiporn Saeung[8], Mohd Shahab[6], Benjamin C. Sparklin[1], Andrew Steven[4], Joseph D. Turner[4], Luke J. Tallon[1], Mark J. Taylor[4], Andrew R. Moorhead[9], Michelle Michalski[10], Jeremy M. Foster[2], Julie C. Dunning Hotopp[1,11,12,*]

[1]Institute for Genome Science, University of Maryland, Baltimore, Maryland, USA.
[2]New England Biolabs, Ipswich, Massachusetts, USA.
[3]Department of Parasitology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand
[4]Centre for Neglected Tropical Diseases, Department of Tropical Disease Biology, Liverpool School of Tropical Medicine, Liverpool, UK
[5]Department of Parasitology, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia
[6]Division of Parasitology, CSIR-Central Drug Research Institute, Lucknow, India
[7]Division of Infectious Diseases, Washington University School of Medicine, St Louis, Missouri, USA
[8]Center of Insect Vector Study, Department of Parasitology, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand
[9]Department of Infectious Diseases, College of Veterinary Medicine, University of Georgia, Athens, Georgia, USA.
[10]University of Wisconsin Oshkosh, Oshkosh, Wisconsin, USA.
[11]Department of Microbiology and Immunology, University of Maryland, Baltimore, Maryland, USA.
[12]Greenebaum Cancer Center, University of Maryland, Baltimore, Maryland, USA.

* jdhotopp@som.umaryland.edu (JCDH); jmattick@som.umaryland.edu (JM)

Abstract

The sequence diversity of natural and laboratory populations of *Brugia pahangi* and *Brugia malayi* was assessed with Illumina resequencing followed by mapping to identify single nucleotide variants and insertions/deletions. In natural and laboratory *Brugia* populations, there is a lack of sequence diversity on chromosome X relative to the autosomes ($\pi_X/\pi_A = 0.2$), which is lower than the expected $\pi_X/\pi_A = 0.75$). A reduction in diversity is also observed in other filarial nematodes with neo-X chromosome fusions in the genera *Onchocerca* and *Wuchereria,* but not those without neo-X chromosome fusions in the genera *Loa* and *Dirofilaria.* In the species with neo-X chromosome fusions, chromosome X is abnormally large, containing a third of the genetic material such that a sizable portion of the genome is lacking sequence diversity. Such profound differences in genetic diversity can be consequential, having been associated with drug resistance and adaptability, with the potential to affect filarial eradication.

Author Summary

Almost a billion people receive >7.7 billion doses of treatment aimed at eliminating lymphatic filariasis, which is caused by three filarial nematodes: *Wuchereria bancrofti, Brugia malayi,* and *Brugia timori.* Drug resistance and adaptation are both associated with pathogen success as well as higher levels of genetic diversity. In an examination of genetic diversity in *Brugia malayi* and *Brugia pahangi,* we observed a lack of genetic diversity over a third of the genome that is found on chromosome X. These species have neo-X chromosomes where a chromosome X fused with an autosome. Using publicly-available published data, the other filarial nematodes of greatest human significance are also found to have a similar lack of genetic diversity on their

51 neo-X chromosomes. The two filarial nematodes with publicly-available data that lack neo-X
52 chromosomes did not have this lack of genetic diversity. This lack of sequence diversity in *B.*
53 *malayi, W. bancrofti,* and *O. volvulus* could have profound effects on all traits encoded on
54 chromosome X.

Introduction

Brugia malayi, Wuchereria bancrofti, and *Brugia timori* are filarial nematodes (roundworms) that are responsible for lymphatic filariasis in humans with almost a billion people receiving >7.7 billion doses of treatment through lymphatic filariasis elimination efforts [1]. All filarial nematodes undergo a complex reproductive cycle that includes multiple larval stages within an arthropod vector followed by more larval stages, sexual development, and reproduction in vertebrate hosts [2]. Of the three filarial species responsible for human lymphatic filariasis, only a subset of *B. malayi* strains can be maintained in small animals in the laboratory, a prerequisite for rigorous laboratory-based studies. These laboratory populations are critical to our understanding of filarial biology, and are commonly used for anti-filarial drug trials [3]. *Brugia pahangi* can also be maintained in a laboratory life cycle, infects cats and dogs, and is occasionally zoonotic. *B. pahangi* and *B. malayi* use mosquito insect vectors and can co-infect dogs and cats [4]. Male *B. malayi* and female *B. pahangi* can produce viable offspring following mating in laboratory conditions [5, 6], but the extent to which this happens successfully in nature is unknown. In addition to lymphatic filariasis, filarial nematodes are responsible for other diseases of medical and veterinary important, including human onchocerciasis [7] caused by the filarial nematode *Onchocerca volvulus*, human loiasis [8] caused by *Loa loa*, and dog and cat heartworm caused by *Dirofilaria immitis* [9].

Onchocerca volvulus [10], *Brugia malayi* [11-13], and *Brugia pahangi* [14] all have nearly complete genomes with chromosome-level assemblies of autosomes and chromosome X, while chromosome Y has yet to be resolved in any filarial nematode. Draft genomes are available for many other filarial nematodes [15], including *W. bancrofti* [16], *L. loa* [17], and *D. immitis* [18].

77 The genomes of all filarial nematodes are represented by six Nigon elements [12, 19, 20] that
78 reflect conserved chromosomal segments that likely reflect the ancestral chromosome state in
79 many nematodes, similar to Muller elements in *Drosophila* species [21]. In the case of filarial
80 nematodes, the composition of these elements was primarily determined through homology to
81 the completed genomes of *O. volvulus*, *Caenorhabditis elegans,* and/or *B. malayi* [12, 19, 20].

82 An important resource for filarial nematode research is the Filariasis Research Reagent
83 Resource Center, better known as FR3, which maintains both *B. malayi* and *B. pahangi* worms
84 across the life cycle in both Mongolian gerbils (jirds; *Meriones unguiculatus*) and cats [3]. At
85 FR3, *B. malayi* and *B. pahangi* are passaged in cats via a mosquito vector. First, blood containing
86 microfilariae is drawn from multiple cats, and pooled together. Then, this pooled blood is fed to
87 mosquitos to allow microfilariae to develop to infective third-stage larvae (L3) which are
88 extracted from mosquitos and introduced into an uninfected cat. Not all mosquitos survive
89 infection with microfilariae, and not all infective L3 worms that are introduced into cats mature
90 into viable adults. Infective L3s are also used to inoculate Mongolian gerbils that are used as a
91 source of much of the material that is distributed by FR3. There are several steps where
92 bottlenecks could occur, and different labs that maintain the life cycle have their own methods
93 to prevent bottlenecks.

94 Genetic diversity can be influenced by bottlenecks, polyandry, population size, sex-biased
95 population size, sex-biased or sex-exclusive inheritance, the rate of recombination, the
96 mutation rate, and selection [22, 23]. Bottlenecks occur when there is a rapid reduction in the
97 population size such that allele frequencies shift dramatically [24] and have been studied in
98 other parasite species [25-27]. These bottlenecks can significantly reduce genomic variation,

but the presence of alleles that confer survival advantages can also generate selective sweeps that produce similar reductions in genomic variation [28]. Sex chromosomes add additional complexity to genetic diversity. For instance, in heteromorphic sex chromosomes like those in X-Y sex determination systems (which includes some filarial nematodes), the X chromosome has reduced genetic diversity by virtue of reduced effective population size. In a population with random mating (e.g. one without polyandry), this results in ~0.75 variance on chromosome X and ~0.25 variance on chromosome Y relative to the autosomes, but in species with multiple mating, this variance can be reduced even further [29].

Though multiple centers across the globe maintain *B. malayi* in laboratories, many of these laboratory populations are derived from the same initial population. Several cats were experimentally infected in the early 1960s with a sub-periodic zoophilic *B. malayi* strain that is reported to be derived from a human patient from Malaysia [30] and distributed to numerous places by Prof. Dr. C. P. Ramachandran [31, 32]. Recipients included the Central Drug Research Institute, Lucknow, India, and the University of California Los Angeles (UCLA), among others. Most modern *B. malayi* laboratory lines are descended from this latter line at UCLA [3], including populations maintained and distributed by TRS labs and the NIAID-funded Filariasis Research Reagent Resource Center (FR3). FR3 and TRS supply one another worms when either laboratory has issues with their populations. In addition, investigators acquire worms from FR3 and/or TRS to establish their own culture collections and replenish with worms as needed, including the laboratories of Prof. Mark Taylor and Dr. Joseph Turner in the Liverpool School of Tropical Medicine and Dr. Gary Weil and Dr. Ramakrishna Rao at Washington University in St. Louis. A further *B. malayi* line was established independently from an infected woman in

Narathiwat Province, southern Thailand, and has been maintained at The Faculty of Tropical Medicine, Mahidol University, Bangkok, then Chiang Mai University, Thailand, for ~40 years with no mixing with the other laboratory lineages [33].

The *B. pahangi* lineage at FR3 is thought to have been established in the 1970s [34] from a green leaf monkey. Because *B. pahangi* and *B. malayi* share very similar life cycles, the procedure for laboratory maintenance for both species at FR3 is similar.

Using samples of *B. malayi* and *B. pahangi* from multiple laboratory centers as well as natural samples of *B. pahangi* that were acquired from wild cats [35], we sought to investigate the genomic diversity within these *Brugia* populations. Given the potential for frequent bottlenecks both in nature and the laboratory, there is the repeated and significant risk of a founder effect that we sought to examine. To this end, we have employed public data from other filarial nematodes, including *W. bancrofti*, *L. loa*, *O. volvulus* and *D. immitis* in order to place this population diversity in the context of the broader filarial nematode family.

Materials and Methods

B. malayi Library Preparation and Sequencing

Adult male worms were provided from the following *B. malayi* centers: Washington University in St. Louis, MO, USA; Liverpool School of Tropical Medicine, UK; TRS Laboratories, Athens, GA, USA; FR3, Athens, GA, USA; Central Drug Research Institute, Lucknow, India; and Chiang Mai University, Chiang Mai, Thailand (**S1 Text**). Adult male worms were sequenced, since females are typically gravid precluding obtaining their individual genome. While virgin females would be a viable alternative, the difficulties in isolating them would have precluded us from obtaining

many of the samples used here. Frozen single adult males recovered from the host gerbil were homogenized separately in 50 µl Buffer G2 from the genomic DNA buffer set (Qiagen) supplemented with RNase A (Qiagen) to 200 µg/mL. Homogenization was performed in a 1.5 mL microcentrifuge tube using a disposable micro pestle (Kimble-Chase). The homogenate was removed to a fresh tube and then the pestle and original tube were washed with an extra 0.95 mL of G2 buffer with RNase which was then added to the sample. The homogenized sample was then processed according to the protocol for tissue samples described in the genomic DNA handbook (Qiagen) and using genomic-tip 20/G gravity flow columns (Qiagen) except 80 U proteinase K (New England Biolabs) were used. Elution buffer QF was prewarmed to 50 °C to increase DNA recovery. The DNA was precipitated by centrifugation as recommended, but in the presence of 20 µg glycogen (Invitrogen). Genomic DNA was sheared to ~380 bp with an ultrasonicator (Covaris) and used to construct indexed PE Illumina libraries using the NEBNext Ultra DNA kit (New England Biolabs). All samples were sequenced on the Illumina HiSeq 2500 with a read length of 100 bp, except for W_male_2 and W_male_6, which were sequenced on the Illumina HiSeq 4000 with a read length of 150 bp. While the data was generated specifically for this study, the data from a subset of samples were used in a previously published study to aid in identification of sex chromosomes and as such these methods are previously described for those samples [12].

B. pahangi Library Preparation and Sequencing

Adult *B. pahangi* male worms were provided from the following locations: FR3 laboratories, at both University of Georgia, Athens, GA, USA; University of Wisconsin, OshKosh, WI, USA (**S1 Text**) and University of Malaya, Kuala Lumpur, Malaysia [35]. Adult females were obtained from

FR3 laboratories and pooled for the purposes of this analysis. Pooled adult female samples were prepared as described in Mattick et al [14]. Endemic isolates from Malaysia were prepared in an identical fashion to the *Brugia malayi* samples described above. Frozen single adult males obtained from FR3 and recovered from the same host gerbil were separately homogenized under liquid nitrogen in 1.5 mL microcentrifuge tubes. The samples were processed according to the Qiagen DNeasy blood and tissue insect protocol using 180 µl buffer ATL and 20 µL proteinase K. The samples were processed according to the manufacturer's recommendations and eluted in 200 µL of buffer AE. After DNA isolation, the pooled adult female sample and the *B. pahangi* male FR3_UWO_Bp1AM_09 sample were sequenced on the Illumina HiSeq2500 from KAPA Hyper libraries with 150 bp paired-end reads. For all other *B. pahangi* samples, genomic DNA was sheared to ~380 bp with an ultrasonicator (Covaris) and prepared into an indexed, paired-end Illumina library using the NEBNext Ultra DNA kit. These samples were sequenced on the Illumina HiSeq 4000 with 150 bp paired end reads.

Sample Variant Calling and Processing for All Individual Nematode Species

Each individual *B. pahangi, B. malayi, O. volvulus, D. immitis, C. elegans* and *Drosophila melanogaster* sample was mapped against its respective genome(GCA_000002995.5, GCA_012070555.1, GCA_000002985.3, GCA_001077395.1, GCA_000499405.2, GCA_000001215.4) [14, 36-40] using BWA MEM [41] with the following settings: -M -a. The resulting BAM files were all sorted and de-duplicated using the Picard tools SortSam and MarkDuplicates, respectively [42] using default parameters for both. Single Nucleotide Variants (SNVs) were jointly called for each sample using Genomic Variant Call Format (GVCF) files generated using the Genome Analysis Tool kit (GATK) [43] with the HaplotypeCaller with the --

read-filter MappingQualityReadFilter setting. The resulting GVCF files were merged and jointly called for SNVs using the GATK GenomicsDBImport and GenotypeGVCFs functions, then filtered using a manual filter with the following settings: --filter-name "QD" --filter-expression "QD < 5.0" --filter-name "QUAL" --filter-expression "QUAL < 30.0" --filter-name "DP" --filter-expression "DP < 14.0" --filter-name "MQ" --filter-expression "MQ < 30.0" --filter-name "MQRankSum" --filter-expression "MQRankSum < -12.5" --filter-name "ReadPosRankSum" --filter-expression "ReadPosRankSum < -8.0" --filter-name "FS" --filter-expression "FS > 60.0". For male samples from species where chromosome structure was known (*B. malayi*, *B. pahangi*), the autosomes were called with a ploidy of 2, while the X chromosome was called at a ploidy of 1. For female samples from species where chromosome structure was known (*O. volvulus*), the autosomes and X chromosome were called with a ploidy of 2. Filtration in samples called with a ploidy of 1 were filtered with --filter-name "DP" --filter-expression "DP < 7.0" to reflect the reduced sequencing depth on those sequences. Putative known pseudoautosomal regions from *B. malayi*, *B. pahangi*, and *O. volvulus* were excluded from variant analysis.

Sample Variant Calling and Processing for Multi-Individual Samples

Each multi-individual *W. bancrofti* sample was mapped against its respective genome (GCA_000002995.5, GCA_012070555.1) [14, 37] using BWA MEM [44] with the following settings: -M -a. The resulting BAM files were all sorted and de-duplicated using the Picard tools SortSam and MarkDuplicates respectively [42] using default parameters for both. SNVs were called using the Freebayes software, specifically the freebayes-parallel feature using default parameters.

SNV Density and Pi Analysis

SNV density can allow for the identification of regions of the genome that are under- or over-represented in variants relative to the entire genomic sequence. SNV density across each of the chromosomes was calculated over 10-kbp sliding non-overlapping windows, considered as 20,000 possible variant sites with homozygous variants counting for 2 site changes and heterozygous variants counting as 1 site change. Pi was calculated using VCFtools over 10 kbp non-overlapping windows for all samples with a genomic coverage > 80% (**S1 Table**) for samples with a ploidy of 2. Because VCFtools requires diploid sites, the R package PopGenome [45] was used with default parameters to calculate Pi for *B. malayi*, *B. pahangi* and *O. volvulus* X chromosomes. Plots of SNV density and Pi were generated using the ggplots2 package in R [46], with the 10-kbp regions as the X-axis and Pi as the Y-axis. A density plot for Pi for each species was generated using the geom_density function of ggplots with default settings on the 10-kbp values of Pi across each chromosome. SNV density and Pi were assigned to Nigon elements, which were determined as previously described [12]. Briefly, contigs were mapped against *B. malayi*, *O. volvulus* and *C. elegans* using the NUCmer tool from the MUMmer package v. 3.23 [47], and contigs were assigned to a specific Nigon element based on the largest match against each specific Nigon element. Principal component analysis was conducted on all autosomal variants in *Brugia malayi* and *Brugia pahangi* individuals using PLINK v.1.9 [48] with the --pca parameter. The resulting primary two principal components for each species were plotted using the geom_point function of ggplots with default settings in R.

Phylogenetic Relationships

Phylogenetic relationships for chromosome X and the autosomes were developed by first obtaining current genomes for *B. timori*, *W. bancrofti* and *O. volvulus* from WormBase [49]. Conserved nematode genes from these genomes, in addition to *B. malayi* and *B. pahangi*, were predicted using BUSCO v. 4.06 package and its nematoda_odb10 database [50]. To ensure orthology, the genomes that were not in chromosome form (i.e. *B. timori* and *W. bancrofti*) were aligned against *B. malayi* using the NUCmer tool from the MUMmer package v. 3.23 [47]. Contigs were binned to a chromosome based on maximum match length, and genes were assigned to chromosome X or the autosomes based on their contig matches. Genes present in all 5 species were aligned using TranslatorX [51] and filtered to include only those that were <15% dissimilar (>85% similarity) at the amino acid level and had at most a difference of 10% in gene length amongst all 5 orthologues. This left a total of 38 genes on chromosome X, and 228 genes on the autosomes. Trees were generated for these sequences using IQ-TREE with default parameters [52], and plotted using iTOL [53]. Mitochondrial sequences (NC_004298.1, CM022469.1, NC_016186.1, AP017686.1) for each species were obtained from GenBank, and aligned at the nucleotide level using MAFFT v.7.427 [54]. The mitochondrial tree was generated and plotted in an identical manner to the autosome and chromosome X trees.

Ethics Statement

All animals in the US were handled in accordance with guidelines defined by the Animal Welfare Act (A3381-01), Association for Assessment and Accreditation of Laboratory Care International (AAAALAC), PHS Policy for the Humane Care and Use of Laboratory Animals, and the Guide for the Care and Use of Laboratory Animals. Animal work for FR3 was approved

249 under the University of Georgia Athens Institutional Animal Care and Use protocol A2010 12-
250 005 and A2013 11-009 or the University of Wisconsin Oshkosh under IACUC protocol number
251 0026-000246-R2-01-12-17. All animal research at TRS was approved under Institutional Animal
252 Care and Use Protocol 13-03 or 14-03. All animal work at WUSM was approved under WUSM
253 Institutional Animal Care and Use Protocol 20120025.

254 The study in Lucknow India bears IAEC approval number 129/08/Para/IAEC/renew (84/09)
255 dated April 27, 2009.

256 All experiments on animals at Liverpool School of Tropical Medicine were approved by the
257 ethical committees of Liverpool School of Tropical Medicine and the University of Liverpool and
258 were conducted according to Home Office Legislation, the revised Animals (Scientific
259 Procedures) Act of 1986 (project license numbers 3002974, P86866FD9).

260 Approval for using gerbils for sample work in Malaysia was granted by the University of Malaya
261 Animal Care and Use Committee (Ref. No. PAR/29/06/2012/RM [R]).

262 The protocol for samples obtained from Thailand was approved by the Institutional Animal Care
263 and Use Committee (Protocol Number 15/2562) of the Faculty of Medicine, Chiang Mai
264 University, Chiang Mai province, Thailand.

Results

Genomic Variation in *B. malayi* Laboratory Populations

Between 4-6 individual adult male *B. malayi* worms were sequenced from each of 6 laboratory populations, which are from three primary *B. malayi* population groups: (a) FR3 and FR3 derived lines, including the continually maintained FR3 line, the line maintained by TRS labs, and the lines at Washington University in St. Louis and the Liverpool School of Tropical Medicine; (b) those from a life cycle established at the same time as the FR3-derived lines, but maintained independently for decades in Lucknow, India; and (c) those from the life cycle in Chiang Mai, Thailand, established from a completely independent human infection and maintained in the laboratory independently for ~40 years. Paired-end Illumina sequencing reads were generated to an average of 85× sequencing depth from individual adult male *B. malayi* worms (**S1 Table**). These adult male worms from each site were collected from the same gerbil, with the exception of TRS, where half of the worms were obtained from a different host gerbil (**S2 Table**). All of the reads were mapped to the reference *B. malayi* genome [11-13] that was obtained with worms from FR3 and TRS. The *B. malayi* samples had an average of 105,264 SNVs per sample, and 21,227 insertions/deletions per sample identified with the GATK HaplotypeCaller called jointly on all samples. The *B. malayi* samples also had a transition/transversion ratio (ts/tv) ranging from 2.10-2.60 (**S2 Table**).

SNV Density and Pi Across the *B. malayi* Genome

The analysis of SNV distribution using Pi was calculated over the Nigon elements associated with each chromosome. Nigon elements are regions of nematode genomes that likely reflect

the ancestral five autosomes and a single sex chromosome. Nigon elements persist despite genome rearrangements because of the infrequency of recombination between chromosomes in nematodes [12, 19, 20]. These are similar to Muller elements in *Drosophila* [21] with Nigon elements being denoted as Nigon-A, Nigon-B, Nigon-C, Nigon-D, Nigon-E, and Nigon-X. The gene content on Nigon elements remains largely conserved even following neo-X chromosome evolution, like the fusion of Nigon-D and Nigon-X in *Brugia* spp. and Nigon-D and Nigon-E in *O. volvulus* [12].

The average SNV density across all samples **(S1 Fig)** and the amount of allelic diversity (Pi) for all 26 *B. malayi* samples **(Fig 1)** were similar when calculated in 10-kbp windows across each of the Nigon elements for each sample. For species where chromosome X and the pseudo-autosomal region were defined and the samples were known to be male (*B. malayi* and *B. pahangi*), Pi for this chromosome was calculated using a ploidy value of 1, while the remaining chromosomes were calculated using the standard ploidy of 2. In these cases, X-specific will refer to the region of chromosome X that is not shared with chromosome Y, while the pseudo-autosomal region will refer to the shared sequence between the X and Y chromosomes. After excluding the pseudo-autosomal region of chromosome X, the average Pi across the X specific Nigon-D and Nigon-X are 5-fold lower ($\pi_X/\pi_A = 0.19$) when compared to similar regions of the autosomes **(Fig 2**).

A principal component analysis identified that while populations recently supplemented from the FR3 lineage are very similar, the Thai samples and the Indian samples are significantly different, despite those from Lucknow, India, sharing a common background with the FR3 lines (**Fig 3A**).

Genomic Variation in *B. pahangi* Samples

Individual adult male *B. pahangi* worms were sequenced from endemic *B. pahangi* from a cat in Malaysia and from the *B. pahangi* FR3 laboratory population. For sequencing of endemic *B. pahangi*, *Aedes togoi* mosquitos were allowed to feed on a naturally-infected microfilaremic wild cat, L3s were recovered, and these L3s were used to infect gerbils as previously described by Lau et al. [35]; three of these adult worms from a single gerbil were individually sequenced and used for variant analysis. These three worms were compared to seven adult male *B. pahangi* worms from the FR3 laboratory population from two gerbils. All of these samples were sequenced on the Illumina HiSeq platform, resulting in an average 105× sequencing depth (range: 22×-217×) per individual across the genome (**S3 Table**). All samples were mapped to the *B. pahangi* FR3 genome [14]. On average there were 315,514 SNVs and 107,463 insertions/deletions identified with the GATK HaplotypeCaller in each *B. pahangi* sample with a consistent ts/tv of 2.67-2.95, which is higher than the ts/tv for *B. malayi* calculated above.

SNV Density and Pi Across the *B. pahangi* Genome

The average SNV density across all samples **(S2 Fig)** and the amount of allelic diversity (Pi) **(Fig 1)** for all 10 samples were calculated in 10-kbp windows across each of the Nigon elements for each sample. Based on both the sequencing depth (**S3 Fig**) difference between BP_ChrX_c and other contigs in the *B. pahangi* chromosome X and the decrease in apparent sequence diversity on chromosome X contigs in all but BP_ChrX_c (**S4 Fig**), BP_ChrX_c was determined to be the pseudo-autosomal region and analyses were adjusted accordingly. After excluding the pseudoautosomal region of the X chromosome, the average Pi across Nigon elements D and X is 5-fold lower ($\pi_X/\pi_A = 0.21$) when compared to Nigon elements in the autosomes **(Figs 1 and 2**).

A principal component analysis using PLINK identified that the FR3 *B. pahangi* samples are distinct from the endemic samples, but that the FR3 samples are also much more closely related to each other than the endemic samples are to one another (**Fig 3B**). The second principal component primarily separates out each endemic sample, suggesting that these worms have significantly more diversity than those from the FR3 lineage.

Introgression

In each *Brugia* nematode, there are three genomes—the mitochondrial genome, the *Wolbachia* endosymbiont genome, and the nuclear genome. Because of the similarities in nucleotide identity, chromosome structure (including a largely shared X chromosome and similar pseudoautosomal region) and genome size between *B. pahangi* and *B. malayi*, as well as the documented ability for these species to successfully cross [5], we tested if there was introgression between *B. pahangi* and *B. malayi*. If an introgression occurred that resulted in the transfer of a chromosome X from one *Brugia* species to the other, one would expect that a phylogenetic tree drawn from chromosome X would look different than that of the autosomes. However, phylogenetic trees of a subset of conserved genes on the autosomes of these agents of lymphatic filariasis and a related filarial parasite, *Onchocerca volvulus,* are similar in topology and relative distance when compared to those on chromosome X and the mitochondria, while the rates of variation are different (**Fig 4**). These phylogenetic patterns between *B. malayi* and *B. pahangi* that are the same for chromosome X, the autosomes, and mitochondrial sequences suggest that the decreased variation on chromosome X did not result from introgression. The conserved phylogenetic topology suggests that this lack of sequence diversity predates the origins of *Brugia* spp.

Other Filarial Genomes

To examine the loss of sequence diversity on chromosome X more widely, particularly with respect to the two neo-X chromosomes, we compared the sequence diversity across exemplar filarial nematodes that have sequence data from multiple samples, including *B. malayi, B. pahangi*, *O. volvulus* [10], *W. bancrofti* [16]*, L. loa* [17]*,* and *D. immitis* [18]. These analyses capitalized on the organization of nematode genomes that allows for the attribution of contigs to Nigon elements even in the highly fragmented genomes like *W. bancrofti* [16]*, L. loa* [55] and *D. immitis* [56]. *W. bancrofti* is predicted to have a Nigon-D and Nigon-X fused neo-X chromosome like *Brugia* spp., *O. volvulus* has a Nigon-D and Nigon-E fused neo-X chromosome, and *D. immitis* and *L. loa* are predicted to have just Nigon-D as their X chromosome [12]. If the loss of sequence diversity in chromosome X of *Brugia* is associated with neo-X chromosome evolution, we would expect there to be a similar loss in the phylogenetically distinct *O. volvulus* that we do not see in *D. immitis* or *L. loa.* In addition, the results were compared to similar data [57, 58] for the model organisms *C. elegans* and *D. melanogaster* that have complete genomes [59, 60], and a large amount of available population data. *C. elegans* is a free-living nematode with an XO reproductive system, while *D. melanogaster* is an arthropod with an XY reproductive system.

Publicly-available WGS data from populations of *O. volvulus* (mixed sex individuals), *W. bancrofti* (mixed samples), *L. loa* (mixed samples), *D. immitis* (individual males), *C. elegans* (mixed sex individuals), and *D. melanogaster* (mixed sex individuals) were analyzed to ascertain whether the loss of diversity observed in *B. malayi* and *B. pahangi* was present in other filarial nematodes. Given the fragmented nature of some of the filarial nematode genomes and the

lack of Y chromosomes in some species, the pseudo-autosomal region could only be excluded from *O. volvulus* and *D. melanogaster*. Contigs from the nematode genomes were assigned to Nigon elements based on their homology to *B. malayi*, *C. elegans,* and *O. volvulus.* The distribution of Pi across Nigon elements was non-normal with a mean outside the interquartile range such that the data violates many of the assumptions of common statistical tests. However, visual inspection of the box plots reveals that in nematodes with neo-X chromosomes (i.e. *Brugia* spp., *W. bancrofti,* and *O. volvulus*) chromosome X can clearly be delineated with a lower Pi (**Fig 2**), despite the difference in the Nigon-composition of those neo-X chromosomes. In contrast, in nematodes without neo-X chromosomes (i.e. *D. immitis, L. loa,* and *C. elegans*) as well as in *D. melanogaster*, chromosome X cannot be clearly delineated (**Fig 2**), and Pi on chromosome X is in line with Pi on the autosomes. This indicates that this profound lack of sequence diversity on Pi is not due solely to the life cycle and lifestyle of filarial nematodes, but instead to creation of neo-X chromosomes through fusion with an autosome.

Discussion

B. malayi and *B. pahangi* filarial nematodes populations have genetic diversity that is consistent with the known separation over time of these populations (**Fig 3**). The greatest difference is seen between endemic nematodes and laboratory populations in the case of *B. pahangi,* or between independently derived laboratory populations in the case of *B. malayi*. To a lesser extent there are differences between nematodes that were derived from the same human sample but have been maintained separately for decades reflected in the differences between Lucknow and the FR3 samples.

Lack of access to clinical samples precluded their inclusion in this study. While the passage of laboratory populations through non-native hosts could impact the genetic diversity, introducing new bottlenecks and selective pressures, the lack of diversity on neo-X chromosomes was found in at least two populations for each of four species with known neo-X fusions (*B. malayi*, *B. pahangi*, *W. bancrofti*, and *O. volvulus*) and was absent from the two filarial nematodes that lack such fusions (*L. loa* and *D. immitis*). Further population level data and the completion of filarial nematode genomes will likely shed further light on the factors influencing genetic diversity in filarial nematodes as well as parasitic nematodes more broadly.

A significant difference in genetic diversity was observed between autosomes and chromosome X. Genetic diversity can be influenced by bottlenecks, polyandry, rate of recombination, mutation rate, selection, and effective population size [22, 23]. The loss of genetic diversity on chromosome X is not limited to just laboratory populations (and the bottlenecks associated with laboratory propagation) since natural populations of *W. bancrofti* and *B. pahangi* have the same loss of diversity. Although polyandry and population shrinkage may also contribute to loss of diversity in filarial nematodes, it is quite likely to be similar for all of the examined filarial nematodes given their life history.

The rate of recombination is expected to be suppressed in sex chromosomes relative to autosomes [61], which is supported by the significantly reduction in intrachromosomal inversions observed in the *Brugia* chromosome X relative to its autosomes [12]. In addition, chromosome Y has an abundance of repeats and transposable elements that prevented its assembly [12], and these repetitive elements are predicted to play a critical role in the further suppression of recombination [62].

In mammals and birds, the higher mutation rate in males over females leads to differences in the mutation rate between autosomes and sex chromosomes [63], while in at least one plant [64] the autosome and sex chromosome mutations are approximately equal. Differences in mutation rate on the sex chromosomes in mammals are associated with more rounds of replication in male gametes, which is likely also the case in filarial nematodes. However, we expect male gametogenesis to be similar between all examined filarial nematodes, such that the differences we observe are not likely attributed to the mutation rate.

Genetic diversity can also be influenced by sex-biased effective population size, sex-biased inheritance, and sex-exclusive inheritance [22, 23]. While we cannot rule out the effects of sex-biased inheritance or sex-exclusive inheritance, we suggest that they would likely be the same across all examined filarial nematodes.

Across nematodes and even filarial nematodes, there is a diversity of sex chromosomes, with XO sex determination being common, but XY being present, and even some nematodes having three sexes [65]. Among the filarial nematodes examined, *L. loa* and *D. immitis* are thought to be XO [66], with *Brugia* spp. and *Onchocerca* spp. being XY [66] resulting from different neo-X fusions [12]. In the absence of selection and no sex bias in reproduction, the expected population size for an organism with heteromorphic XY chromosomes, like *Brugia* and *Onchocerca* filarial nematodes, the autosome:(chromosome X):(chromosome Y) allelic frequency is 4:3:1. As a consequence, a reduction of nucleotide diversity is expected on heteromorphic sex chromosomes, with $\pi_X/\pi_A \sim 0.75$ [10, 22]. Similarly, nematodes with XO sex determination would have an expected autosome:(chromosome X):(chromosome Y) allelic frequency is 4:3:0 with $\pi_X/\pi_A \sim 0.75$. However, we observe $\pi_X/\pi_A \sim 0.2$ for both *Brugia* species.

Upon examination of other filarial nematodes, a reduction in π_X/π_A similar to that in *Brugia* spp. was observed for *W. bancrofti* and *O. volvulus*, all four of which have neo-X chromosomes that emerged after fusion of chromosome X with an autosome. In the case of filarial worms, different neo-X chromosomes were formed at least twice by the fusion of two Nigon elements [12, 19, 20]. The common Nigon element in these fusion events appears to be Nigon-D, which is likely the ancestral sex chromosome of filarial nematodes [12, 19, 20]. The chromosomal fusion event in the ONC3 clade, containing *Onchocerca* spp., joined Nigon-D and Nigon-E, while the chromosomal fusion in the ONC5 clade, containing *Brugia* spp. and *Wuchereria* sp., joined Nigon-D and Nigon-X (**Fig 5**). Both times that there is a loss in diversity on chromosome X in this study, there is a concomitant neo-X fusion. And conversely, where there is not a neo-X fusion, there is not the loss of diversity (i.e. *L. loa* and *D. immitis*). As such this lack of genetic diversity on chromosome X seems consistent with the formation of the neo-X chromosomes prior to several speciation events, like that of *Brugia* spp. and *W. bancrofti* (**Fig 5**). Chromosomal fusion events are known to reduce genomic diversity in species as the effective population size of the sex chromosome is reduced and novel genes and dosage mechanisms must be generated to compensate for the fusion [67, 68]. For example, in *Sylvoidea* bird species, a loss of diversity on chromosome Z (the equivalent of chromosome X in ZW systems) is attributed to a neo-sex chromosome fusion [69].

Chromosomal fusions may not be the only source of diversity loss on chromosome X. For example, *Haemonchus contortus*, a parasitic nematode, does not show evidence of a recent chromosomal fusion. Yet the *H. contortus* π_X/π_A is 0.36 [70], which is also lower than neutral expectation of $\pi_X/\pi_A \simeq 0.75$. This decrease in *H. contortus* was attributed to host sex biases due

461 to reproductive fitness being over-dispersed between males and females from polyandry and
462 high fecundity [70]. However, filarial nematodes only seem to have this lack of genetic diversity
463 on neo-X chromosomes despite likely polyandry and high fecundity across many or most filarial
464 nematodes.

465 In nematodes, there has also been a transition in the sex chromosomes. Nigon-D is likely the
466 ancestral chromosome for all Rhabditida nematodes, with a conversion of Nigon-X to
467 chromosome X in Rhabditina nematodes, which includes *C. elegans* [12]. This transition does
468 not appear to be associated with a difference in genetic diversity for chromosome X upon
469 comparisons of *C. elegans* and the filarial nematodes without neo-X fusions, like *D. immitis* and
470 *L. loa*. (**Fig 2**). It is possible that altering the sex determining Nigon element is not enough to
471 cause diversity loss, and that it is specifically associated with chromosomal fusion. Alternatively,
472 it is possible that enough time has elapsed to eliminate the signature associated with that
473 transition at least with the resolution with which it was examined here.

474 The same processes that subject chromosome X to decreased genetic diversity and Muller's
475 ratchet also affect chromosome Y to a much larger degree [63, 71]. In filarial nematodes, we do
476 not have an assembled chromosome Y, and are limited to male-specific contigs attributed to
477 chromosome Y. But the high repetitiveness of the sequences [12] suggests that filarial
478 nematode Y chromosomes are undergoing a degeneration consistent with neo-Y formation.

479 Although chromosomal fusions appear to be associated with diversity loss in filarial worms, it is
480 not yet clear if this will be found universally in other parasitic nematodes. This lack of
481 chromosome X genetic diversity is important since most medically important filarial nematodes

have neo-X fusions with a third of all genetic material being on chromosome X, representing a substantial loss of sequence diversity. Genetic material on chromosome X also undergoes recombination at a lower rate than the rest of the genome [61] . Thus the sex chromosome is more susceptible to Muller's Ratchet [72], which is a process whereby deleterious mutations accumulate in the absence of recombination. This loss of diversity on such a large portion of the genome could have significant consequences. In other parasites, drug resistance and adaptability are associated with a higher level of genetic diversity, and its absence can prevent an organism from developing strategies of coping with adverse events [73].

Conclusions

Populations were examined that were derived from two independent isolates of *B. malayi* and *B. pahangi.* For *B. malayi* this includes several populations derived from a human from Malaysia and a population from an infected woman in Thailand. For *B. pahangi* this includes the populations derived from a green leaf monkey from Malaysia and from naturally infected Malaysian cats. We observe a profound lack of sequence diversity on chromosome X in all independent populations of *B. malayi* and *B. pahangi* that is consistent with reduced chromosome X diversity in other sequenced filarial nematodes with neo-X chromosomes. Given the importance that sequence diversity has with respect to adaptability and the size of chromosome X, which is a third of the genome, this lack of sequence diversity in a third of the genome in medically important filarial nematodes is likely to have a large effect on the evolutionary trajectory of these species.

Supporting Information

AcknowledgmentsAdult *B. malayi* and *B. pahangi* males were obtained through the NIH Biodefense and Emerging Infections Research Resources Repository, NIAID, NIH, which procures material from the NIH/NIAID Filarial Research Reagent Resource Center (FR3) with morphological voucher specimens stored at the Harold W. Manter Museum at University of Nebraska, accession numbers P2021-2032. The NIAID, USA, and Central Drug Research Institute, India, played no part in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

1. Global programme to eliminate lymphatic filariasis: progress report. World Health Organization, 2018 Contract No.: WER No 41, 2019, 94, 457–472.

2. Center For Disease Control. Biology - Life Cycle of *Brugia malayi*. [Webpage]. 2011 [updated April 11]. Available from: https://www.cdc.gov/parasites/lymphaticfilariasis/biology_b_malayi.html.

3. Michalski ML, Griffiths KG, Williams SA, Kaplan RM, Moorhead AR. The NIH-NIAID Filariasis Research Reagent Resource Center. PLoS Negl Trop Dis. 2011;5(11):e1261-e. Epub 2011/11/29. doi: 10.1371/journal.pntd.0001261. PubMed PMID: 22140585.

4. Areekit S, Khuchareontaworn S, Kanjanavas P, Sriyapai T, Pakpitchareon A, Khawsak P, et al. Molecular genetics analysis for co-infection of *Brugia malayi* and *Brugia pahangi* in cat reservoirs based on internal transcribed spacer region 1. Southeast Asian J Trop Med Public Health. 2009;40(1):30-4. Epub 2009/03/28. PubMed PMID: 19323030.

5. Suswillo RR, Denham DA, McGreevy PB, Nelson GS. Hybridization between *Brugia patei*, *B. pahangi* and sub-periodic *B. malayi*. Parasitology. 1978;77(2):153-60. Epub 2009/04/06. doi: 10.1017/S0031182000049350.

6. Laing ABG, Edeson JFB, Wharton RH. Studies on Filariasis in Malaya: The Vertebrate Hosts of *Brugia Malayi* and *B. Pahangi*. Annals of Tropical Medicine & Parasitology. 1960;54(1):92-9. doi: 10.1080/00034983.1960.11685961.

7. Hoerauf A, Büttner DW, Adjei O, Pearlman E. Onchocerciasis. Bmj. 2003;326(7382):207-10. Epub 2003/01/25. doi: 10.1136/bmj.326.7382.207. PubMed PMID: 12543839; PubMed Central PMCID: PMCPMC1125065.

8. Boussinesq M. Loiasis. Ann Trop Med Parasitol. 2006;100(8):715-31. Epub 2007/01/18. doi: 10.1179/136485906x112194. PubMed PMID: 17227650.

9. Morchón R, Carretón E, González Miguel J, Mellado Hernández I. Heartworm Disease (Dirofilaria immitis) and Their Vectors in Europe – New Distribution Trends. Frontiers in Physiology. 2012;3(196). doi: 10.3389/fphys.2012.00196.

10. Choi Y-J, Tyagi R, McNulty SN, Rosa BA, Ozersky P, Martin J, et al. Genomic diversity in Onchocerca volvulus and its Wolbachia endosymbiont. Nature microbiology. 2016;2:16207-. doi: 10.1038/nmicrobiol.2016.207. PubMed PMID: 27869792.

11. Tracey A, Foster JM, Paulini M, Grote A, Mattick J, Tsai YC, et al. Nearly Complete Genome Sequence of Brugia malayi Strain FR3. Microbiol Resour Announc. 2020;9(24). Epub 2020/06/13. doi: 10.1128/MRA.00154-20. PubMed PMID: 32527783; PubMed Central PMCID: PMCPMC7291094.

12. Foster JM, Grote A, Mattick J, Tracey A, Tsai YC, Chung M, et al. Sex chromosome evolution in parasitic nematodes of humans. Nat Commun. 2020;11(1):1964. Epub 2020/04/25. doi: 10.1038/s41467-020-15654-6. PubMed PMID: 32327641; PubMed Central PMCID: PMCPMC7181701.

13. Ghedin E, Wang S, Spiro D, Caler E, Zhao Q, Crabtree J, et al. Draft genome of the filarial nematode parasite Brugia malayi. Science. 2007;317(5845):1756-60. Epub 2007/09/22. doi: 10.1126/science.1145406. PubMed PMID: 17885136; PubMed Central PMCID: PMCPMC2613796.

14. Mattick J, Libro S, Sparklin BC, Chung M, Bromley RE, Nadendla S, et al. Nearly Complete Genome Sequence of *Brugia pahangi* FR3. Microbiology Resource Announcements. 2020;9(27):e00479-20. doi: 10.1128/MRA.00479-20.

15. Coghlan A, Tyagi R, Cotton JA, Holroyd N, Rosa BA, Tsai IJ, et al. Comparative genomics of the major parasitic worms. Nature Genetics. 2019;51(1):163-74. doi: 10.1038/s41588-018-0262-1.

16. Small ST, Reimer LJ, Tisch DJ, King CL, Christensen BM, Siba PM, et al. Population genomics of the filarial nematode parasite Wuchereria bancrofti from mosquitoes. Molecular Ecology. 2016;25(7):1465-77. doi: https://doi.org/10.1111/mec.13574.

17. Whittaker C, Walker M, Pion SDS, Chesnais CB, Boussinesq M, Basáñez M-G. The Population Biology and Transmission Dynamics of Loa loa. Trends in Parasitology. 2018;34(4):335-50. doi: https://doi.org/10.1016/j.pt.2017.12.003.

18. Lau DC-W, McLeod S, Collaery S, Peou S, Tran AT, Liang M, et al. Whole-genome reference of Dirofilaria immitis from Australia to determine single nucleotide polymorphisms associated with macrocyclic lactone resistance in the USA. Current Research in Parasitology & Vector-Borne Diseases. 2021;1:100007. doi: https://doi.org/10.1016/j.crpvbd.2021.100007.

19. Gonzalez de la Rosa PM, Thomson M, Trivedi U, Tracey A, Tandonnet S, Blaxter M. A telomere-to-telomere assembly of Oscheius tipulae and the evolution of rhabditid nematode chromosomes. G3 (Bethesda). 2021;11(1). Epub 2021/02/10. doi: 10.1093/g3journal/jkaa020. PubMed PMID: 33561231; PubMed Central PMCID: PMCPMC8022731.

20. Tandonnet S, Koutsovoulos GD, Adams S, Cloarec D, Parihar M, Blaxter ML, et al. Chromosome-Wide Evolution and Sex Determination in the Three-Sexed Nematode Auanema

rhodensis. G3 (Bethesda). 2019;9(4):1211-30. doi: 10.1534/g3.119.0011. PubMed PMID: 30770412.

21. Graves JL, Jr., Hertweck KL, Phillips MA, Han MV, Cabral LG, Barter TT, et al. Genomics of Parallel Experimental Evolution in Drosophila. Molecular Biology and Evolution. 2017;34(4):831-42. doi: 10.1093/molbev/msw282.

22. Wilson Sayres MA. Genetic Diversity on the Sex Chromosomes. Genome Biol Evol. 2018;10(4):1064-78. Epub 2018/04/11. doi: 10.1093/gbe/evy039. PubMed PMID: 29635328; PubMed Central PMCID: PMCPMC5892150.

23. Webster TH, Wilson Sayres MA. Genomic signatures of sex-biased demography: progress and prospects. Curr Opin Genet Dev. 2016;41:62-71. Epub 2016/09/07. doi: 10.1016/j.gde.2016.08.002. PubMed PMID: 27599147.

24. Nei M, Maruyama T, Chakraborty R. The Bottleneck Effect and Genetic Variability in Populations. Evolution. 1975;29(1):1-10. doi: 10.2307/2407137.

25. Rich SM, Licht MC, Hudson RR, Ayala FJ. Malaria's Eve: Evidence of a recent population bottleneck throughout the world populations of *Plasmodium falciparum*. Proceedings of the National Academy of Sciences. 1998;95(8):4425. doi: 10.1073/pnas.95.8.4425.

26. Madubata C, Dunams-Morel DB, Elkin B, Oksanen A, Rosenthal BM. Evidence for a recent population bottleneck in an Apicomplexan parasite of caribou and reindeer, *Besnoitia tarandi*. Infection, Genetics and Evolution. 2012;12(8):1605-13. doi: https://doi.org/10.1016/j.meegid.2012.06.007.

27. Cole R, Viney M. The population genetics of parasitic nematodes of wild animals. Parasites & Vectors. 2018;11(1):590. doi: 10.1186/s13071-018-3137-5.

28. McVean G. The Structure of Linkage Disequilibrium Around a Selective Sweep. Genetics. 2007;175(3):1395. doi: 10.1534/genetics.106.062828.

29. Pool JE, Nielsen R. The impact of founder events on chromosomal variability in multiply mating species. Mol Biol Evol. 2008;25(8):1728-36. Epub 2008/05/31. doi: 10.1093/molbev/msn124. PubMed PMID: 18511425; PubMed Central PMCID: PMCPMC2734130.

30. Taylor MJ, Bilo K, Cross HF, Archer JP, Underwood AP. 16S rDNA Phylogeny and Ultrastructural Characterization of *Wolbachia* Intracellular Bacteria of the Filarial Nematodes *Brugia malayi*, *B. pahangi*, and *Wuchereria bancrofti*. Experimental Parasitology. 1999;91(4):356-61. doi: https://doi.org/10.1006/expr.1998.4383.

31. Ash LR, Riley JM. Development of Subperiodic *Brugia malayi* in the Jird, *Meriones unguiculatus*, with Notes on Infections in Other Rodents. The Journal of Parasitology. 1970;56(5):969-73. doi: 10.2307/3277515.

32. Narayanan B, Akademi Sains M. I am ... because of you : the biography of academician Professor Emeritus Dato' Dr. C.P. Ramachandran2014.

33. Saeung A, Hempolchom C, Baimai V, Thongsahuan S, Taai K, Jariyapan N, et al. Susceptibility of eight species members in the *Anopheles hyrcanus* group to nocturnally subperiodic *Brugia malayi*. Parasites & vectors. 2013;6:5-. doi: 10.1186/1756-3305-6-5. PubMed PMID: 23289957.

34. Ash LR. Chronic *Brugia pahangi* and *Brugia malayi* Infections in *Meriones unguiculatus*. The Journal of Parasitology. 1973;59(3):442-7. doi: 10.2307/3278769.

35. Lau Y-L, Lee W-C, Xia J, Zhang G, Razali R, Anwar A, et al. Draft genome of *Brugia pahangi*: high similarity between *B. pahangi* and *B. malayi*. Parasites & Vectors. 2015;8(1):451. doi: 10.1186/s13071-015-1064-2.

36. Genome sequence of the nematode C. elegans: a platform for investigating biology. Science. 1998;282(5396):2012-8. Epub 1998/12/16. doi: 10.1126/science.282.5396.2012. PubMed PMID: 9851916.

37. Tracey A, Foster JM, Paulini M, Grote A, Mattick J, Tsai Y-C, et al. Nearly Complete Genome Sequence of *Brugia malayi* Strain FR3. Microbiology Resource Announcements. 2020;9(24):e00154-20. doi: 10.1128/MRA.00154-20.

38. Godel C, Kumar S, Koutsovoulos G, Ludin P, Nilsson D, Comandatore F, et al. The genome of the heartworm, *Dirofilaria immitis*, reveals drug and vaccine targets. FASEB J. 2012;26(11):4650-61. doi: 10.1096/fj.12-205096. PubMed PMID: 22889830; PubMed Central PMCID: PMC3475251.

39. Cotton JA, Bennuru S, Grote A, Harsha B, Tracey A, Beech R, et al. The genome of Onchocerca volvulus, agent of river blindness. Nat Microbiol. 2016;2:16216. Epub 2016/11/22. doi: 10.1038/nmicrobiol.2016.216. PubMed PMID: 27869790; PubMed Central PMCID: PMCPMC5310847.

40. Hoskins RA, Carlson JW, Wan KH, Park S, Mendez I, Galle SE, et al. The Release 6 reference sequence of the Drosophila melanogaster genome. Genome Res. 2015;25(3):445-58. Epub 2015/01/16. doi: 10.1101/gr.185579.114. PubMed PMID: 25589440; PubMed Central PMCID: PMCPMC4352887.

641 41. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform.
642 Bioinformatics. 2009;25(14):1754-60. Epub 2009/05/20. doi: 10.1093/bioinformatics/btp324.
643 PubMed PMID: 19451168; PubMed Central PMCID: PMCPMC2705234.
644 42. Broad I. Picard tools. Broad Institute, GitHub repository; 2016.
645 43. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, del Angel G, Levy-Moonshine A, et
646 al. From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best
647 Practices Pipeline. Current Protocols in Bioinformatics. 2013;43(1):11.0.1-.0.33. doi:
648 10.1002/0471250953.bi1110s43.
649 44. Li H. Exploring single-sample SNP and INDEL calling with whole-genome de novo
650 assembly. Bioinformatics. 2012;28(14):1838-44. doi: 10.1093/bioinformatics/bts280.
651 45. Pfeifer B, Wittelsbürger U, Ramos-Onsins SE, Lercher MJ. PopGenome: An Efficient Swiss
652 Army Knife for Population Genomic Analyses in R. Molecular Biology and Evolution.
653 2014;31(7):1929-36. doi: 10.1093/molbev/msu136.
654 46. Ihaka R, Gentleman R. R: A Language for Data Analysis and Graphics. Journal of
655 Computational and Graphical Statistics. 1996;5(3):299-314. doi:
656 10.1080/10618600.1996.10474713.
657 47. Delcher AL, Salzberg SL, Phillippy AM. Using MUMmer to Identify Similar Regions in
658 Large Sequence Sets. Current Protocols in Bioinformatics. 2003;00(1):10.3.1-.3.8. doi:
659 10.1002/0471250953.bi1003s00.
660 48. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool
661 set for whole-genome association and population-based linkage analyses. Am J Hum Genet.
662 2007;81(3):559-75. Epub 2007/07/25. doi: 10.1086/519795. PubMed PMID: 17701901.

663 49. Harris TW, Antoshechkin I, Bieri T, Blasiar D, Chan J, Chen WJ, et al. WormBase: a
664 comprehensive resource for nematode research. Nucleic Acids Research.
665 2009;38(suppl_1):D463-D7. doi: 10.1093/nar/gkp952.
666 50. Waterhouse RM, Seppey M, Simão FA, Manni M, Ioannidis P, Klioutchnikov G, et al.
667 BUSCO Applications from Quality Assessments to Gene Prediction and Phylogenomics.
668 Molecular Biology and Evolution. 2017;35(3):543-8. doi: 10.1093/molbev/msx319.
669 51. Abascal F, Zardoya R, Telford MJ. TranslatorX: multiple alignment of nucleotide
670 sequences guided by amino acid translations. Nucleic Acids Res. 2010;38(Web Server
671 issue):W7-13. Epub 2010/05/04. doi: 10.1093/nar/gkq291. PubMed PMID: 20435676; PubMed
672 Central PMCID: PMCPMC2896173.
673 52. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: A Fast and Effective
674 Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. Molecular Biology and
675 Evolution. 2014;32(1):268-74. doi: 10.1093/molbev/msu300.
676 53. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new
677 developments. Nucleic Acids Research. 2019;47(W1):W256-W9. doi: 10.1093/nar/gkz239.
678 54. Katoh K. MAFFT version 7. 2018.
679 55. Desjardins CA, Cerqueira GC, Goldberg JM, Hotopp JCD, Haas BJ, Zucker J, et al.
680 Genomics of Loa loa, a Wolbachia-free filarial parasite of humans. Nature Genetics.
681 2013;45(5):495-500. doi: 10.1038/ng.2585.
682 56. Godel C, Kumar S, Koutsovoulos G, Ludin P, Nilsson D, Comandatore F, et al. The
683 genome of the heartworm, Dirofilaria immitis, reveals drug and vaccine targets. The FASEB
684 Journal. 2012;26(11):4650-61. doi: https://doi.org/10.1096/fj.12-205096.

685 57. Crombie TA, Zdraljevic S, Cook DE, Tanny RE, Brady SC, Wang Y, et al. Deep sampling of
686 Hawaiian Caenorhabditis elegans reveals high genetic diversity and admixture with global
687 populations. Elife. 2019;8:e50465. doi: 10.7554/eLife.50465. PubMed PMID: 31793880.
688 58. Begun DJ, Holloway AK, Stevens K, Hillier LW, Poh YP, Hahn MW, et al. Population
689 genomics: whole-genome analysis of polymorphism and divergence in Drosophila simulans.
690 PLoS Biol. 2007;5(11):e310. Epub 2007/11/09. doi: 10.1371/journal.pbio.0050310. PubMed
691 PMID: 17988176; PubMed Central PMCID: PMCPMC2062478.
692 59. Celniker SE, Rubin GM. The Drosophila melanogaster genome. Annu Rev Genomics Hum
693 Genet. 2003;4:89-117. Epub 2003/10/07. doi: 10.1146/annurev.genom.4.070802.110323.
694 PubMed PMID: 14527298.
695 60. Waterston R, Sulston J. The genome of Caenorhabditis elegans. Proc Natl Acad Sci U S A.
696 1995;92(24):10836-40. doi: 10.1073/pnas.92.24.10836. PubMed PMID: 7479894.
697 61. Furman BLS, Metzger DCH, Darolti I, Wright AE, Sandkam BA, Almeida P, et al. Sex
698 Chromosome Evolution: So Many Exceptions to the Rules. Genome Biology and Evolution.
699 2020;12(6):750-63. doi: 10.1093/gbe/evaa081.
700 62. Reichwald K, Petzold A, Koch P, Downie Bryan R, Hartmann N, Pietsch S, et al. Insights
701 into Sex Chromosome Evolution and Aging from the Genome of a Short-Lived Fish. Cell.
702 2015;163(6):1527-38. doi: https://doi.org/10.1016/j.cell.2015.10.071.
703 63. Johnson NA, Lachance J. The genetics of sex chromosomes: evolution and implications
704 for hybrid incompatibility. Ann N Y Acad Sci. 2012;1256:E1-22. Epub 2012/10/03. doi:
705 10.1111/j.1749-6632.2012.06748.x. PubMed PMID: 23025408; PubMed Central PMCID:
706 PMCPMC3509754.

64. Krasovec M, Chester M, Ridout K, Filatov DA. The Mutation Rate and the Age of the Sex Chromosomes in Silene latifolia. Curr Biol. 2018;28(11):1832-8.e4. Epub 2018/05/29. doi: 10.1016/j.cub.2018.04.069. PubMed PMID: 29804812.

65. Chaudhuri J, Bose N, Tandonnet S, Adams S, Zuco G, Kache V, et al. Mating dynamics in a nematode with three sexes and its evolutionary implications. Scientific Reports. 2015;5(1):17676. doi: 10.1038/srep17676.

66. Post R. The chromosomes of the Filariae. Filaria J. 2005;4:10-. doi: 10.1186/1475-2883-4-10. PubMed PMID: 16266430.

67. Cicconardi F, Lewis JJ, Martin SH, Reed RD, Danko CG, Montgomery SH. Chromosome fusion affects genetic diversity and evolutionary turnover of functional loci, but consistently depends on chromosome size. Molecular Biology and Evolution. 2021. doi: 10.1093/molbev/msab185.

68. Cheng Y, Shang D, Luo M, Huang C, Lai F, Wang X, et al. Whole genome-wide chromosome fusion and new gene birth in the Monopterus albus genome. Cell & Bioscience. 2020;10(1):67. doi: 10.1186/s13578-020-00432-0.

69. Pala I, Hasselquist D, Bensch S, Hansson B. Patterns of molecular evolution of an avian neo-sex chromosome. Mol Biol Evol. 2012;29(12):3741-54. Epub 2012/07/25. doi: 10.1093/molbev/mss177. PubMed PMID: 22826461.

70. Doyle SR, Tracey A, Laing R, Holroyd N, Bartley D, Bazant W, et al. Genomic and transcriptomic variation defines the chromosome-scale assembly of Haemonchus contortus, a model gastrointestinal worm. Commun Biol. 2020;3(1):656. Epub 2020/11/11. doi:

728 10.1038/s42003-020-01377-3. PubMed PMID: 33168940; PubMed Central PMCID:
729 PMCPMC7652881.
730 71. Qiu S, Bergero R, Forrest A, Kaiser VB, Charlesworth D. Nucleotide diversity in Silene
731 latifolia autosomal and sex-linked genes. Proc Biol Sci. 2010;277(1698):3283-90. Epub
732 2010/06/04. doi: 10.1098/rspb.2010.0606. PubMed PMID: 20519224; PubMed Central PMCID:
733 PMCPMC2981921.
734 72. Muller HJ. The relation of recombination to mutational advance. Mutation
735 Research/Fundamental and Molecular Mechanisms of Mutagenesis. 1964;1(1):2-9. doi:
736 https://doi.org/10.1016/0027-5107(64)90047-8.
737 73. Hughes D, Andersson DI. Evolutionary consequences of drug resistance: shared
738 principles across diverse targets and organisms. Nature Reviews Genetics. 2015;16(8):459-71.
739 doi: 10.1038/nrg3922.
740 74. Riddle NC, Elgin SCR. The Drosophila Dot Chromosome: Where Genes Flourish Amidst
741 Repeats. Genetics. 2018;210(3):757-72. Epub 2018/11/08. doi: 10.1534/genetics.118.301146.
742 PubMed PMID: 30401762; PubMed Central PMCID: PMCPMC6218221.
743 75. Lefoulon E, Bain O, Bourret J, Junker K, Guerrero R, Cañizales I, et al. Shaking the Tree:
744 Multi-locus Sequence Typing Usurps Current Onchocercid (Filarial Nematode) Phylogeny. PLoS
745 Negl Trop Dis. 2015;9(11):e0004233. doi: 10.1371/journal.pntd.0004233.

746

Figure Legends

Fig 1. Pi across *B. malayi* and *B. pahangi* samples from multiple laboratory backgrounds. Pi was calculated across each of the *B. malayi* and *B. pahangi* contigs/scaffolds using VCFTools on a combined VCF file containing all samples. The results are organized by chromosome and Nigon elements. Chromosome X shows a distinct lack of nucleotide diversity relative to the autosomes. The lack of diversity on chromosome X appears to be present in nematodes from all laboratory centers for *B. malayi* and in both endemic and laboratory populations for *B. pahangi*. The plots for chromosome X are larger reflecting the increased size of chromosome X which is approximately twice the size of the autosomes. Chromosome Y is not resolved in either organism, and as such Pi could not be calculated.

Fig 2. Pi across filarial nematode species and model organisms. Pi was calculated across *B. malayi, B. pahangi, W. bancrofti, O. volvulus, L. loa, D. immitis, D. melanogaster* and *C. elegans* using VCFTools on a combined VCF file containing all samples for each of those species. For all nematode species, contigs were assigned to a Nigon element based on their homology to *B. malayi*, *O. volvulus* and *C. elegans*. Values of Pi were $\log_{10}$-transformed to more readily visualize the distributions. Filarial nematodes with neo-X chromosomes (Nigon-D/Nigon-X in *Brugia* spp. and *W. bancrofti* and NigonD/Nigon-E in *O. volvulus*) have a significantly depressed Pi compared to autosomal Nigon elements or X chromosomes in other species (Nigon-D in *L. loa* and *D. immitis*, Nigon-X in *C. elegans*, and chromosome X in *D. melanogaster*). This suggests that the loss of diversity observed in *B. malayi* and *B. pahangi* are not limited to those species and related to the formation of the neo-X chromosome. Chromosome 4 in *D. melanogaster* also has

a decrease in Pi; it is a small chromosome sometimes referred to as the dot chromosome that is largely heterochromatic and may formerly have been a sex chromosome [74].

Fig 3. Principal component analysis of *B. malayi* and *B. pahangi* samples. Principal component analyses of the *B. malayi* **(A)** and *B. pahangi* **(B)** samples were conducted using PLINK with default parameters on each individual sample for each population, and the resulting outputs were imported into R and plotted using geom_point from ggplots. All of the FR3-derived *B. malayi* samples cluster very tightly together, except for those derived from the Lucknow strain, which are separated by principal component 2. Principal component 1 primarily divides the 4 samples from Thailand, which not only are distinct from FR3-derived worms, but are much more distinct from each other than FR3-derived worms are from each other. The FR3 single adult male *B. pahangi* all cluster together, while samples from wild infected cats from Malaysia appear to dominate the variation along both principal components.

Fig 4. Phylogenetic trees of conserved nematode BUSCO genes and mitochondria between filarial species. Conserved genes predicted by BUSCO in *B. malayi*, *B. pahangi*, *W. bancrofti*, *B. timori* and *O. volvulus* were separated out by their location and divided based on their presence on chromosome X of *B. malayi* and *B. pahangi* **(A)** or the autosomes of those species **(B)**. These gene sets were used to construct phylogenetic trees using IQ-TREE (bootstrap=1000) that were midpoint rooted in IQ-TREE (https://itol.embl.de/). **(C)** Mitochondrial genome sequences of these organisms were aligned via MAFFT, and trees were generated via IQ-TREE. The relationships between filarial species consistently show *B. malayi* and *B. timori* as more closely related to each other than to *B. pahangi* such that any loss of chromosome X diversity likely predates the divergence of the three organisms.

Fig 5. Phylogenetic relationships related to sex chromosome Nigon content. The phylogenetic relationship of filarial nematodes is shown as adapted from Lefoulon et al. [75]. Nigon element assignments for the sex chromosomes are shown when known or inferred previously [12]. The loss of diversity on the sex chromosome co-occurs with the instances of chromosomal fusions between Nigon-IV and Nigon-X in ONC5 and between Nigon-IV and Nigon-V in ONC3, but does not appear to be present in species that do not contain the chromosomal fusion in either the ONC3 or ONC5 clades.

Supporting Information

S1 Fig. *B. malayi* variant distribution across samples from multiple laboratory backgrounds. SNV density was calculated across each of the *B. malayi* chromosomes and averaged across all samples using 10 kbp windows across each contig and normalized to the total sample number (n=26). Heterozygous variants were considered as half of the value of homozygous variants for the purposes of density calculations. There is a significant loss of variants in chromosome X that is consistent across all individual samples and is displayed here in aggregate, and the pseudoautosomal region of chromosome X is indicated by a red bar.

S2 Fig. *B. pahangi* variant density. SNV density was plotted across each of the *B. pahangi* chromosomes averaged across all samples. Density was calculated using 10 kbp windows across each contig using R and normalized to the total sample number (n=10). Heterozygous variants were considered as half of the value of homozygous variants for the purposes of density calculations. Similar to *B. malayi*, chromosome X of all of the *B. pahangi* samples show a

significant lack of variation in the central region that is in contrast to the autosomes and the rest of chromosome X.

S3 Fig. *B. pahangi* sequencing depth across all samples. Depth plots were calculated over 10 kbp non overlapping regions across the *B. pahangi* chromosomes. The predicted pseudoautosomal region (**S4 Fig**) has depth that is consistent with autosomal depth while the rest of chromosome X appears to be at half depth. This is consistent with a pseudo-autosomal profile in this contig.

S4 Fig. Heterozygous *B. pahangi* SNV density across chromosome X. Density plots were generated for heterozygous SNV density values calculated over 10 kbp non-overlapping regions across the *B. pahangi* chromosome X. Pseudo-autosomal regions in chromosome X of *B. malayi* have been previously described [12], and an analysis of heterozygous SNVs in chromosome X of adult *B. pahangi* males (which should only be possible in pseudoautosomal regions) reveals that a similar region of the chromosome has an enriched value of Pi, indicating that *B. pahangi* has a similar pseudoautosomal region to *B. malayi*.

S1 Table. Coverage and Sequencing Depth Metrics for all Samples in Figure 2.

S2 Table. Metadata, Mapping statistics, and Variant Calls for *Brugia malayi.*

S3 Table. Metadata, Mapping statistics, and Variant Calls for *Brugia pahangi.*

S1 Text. Standard operating procedure for *Brugia* rearing in jirds across all Centers. This document includes all standard operating procedures for maintaining the laboratory life cycle

for *B. malayi* and/or *B. pahangi* for the different laboratories that provided samples from their collections.