Genomes of the human filarial parasites *Mansonella perstans* and *Mansonella ozzardi*

Amit Sinha[1], Zhiru Li[1], Catherine B. Poole[1], Richard D. Morgan[1], Laurence Ettwiller[1], Nathália F. Lima[2], Marcelo U. Ferreira[2,3], Fanny F. Fombad[4], Samuel Wanji[4], Clotilde K. S. Carlow[1*]

[1]New England Biolabs, Ipswich Massachusetts, 01938 USA;

[2]Department of Parasitology, Institute of Biomedical Sciences, University of São Paulo, São Paulo, Brazil;

[3]Global Health and Tropical Medicine, Institute of Hygiene and Tropical Medicine, NOVA University of Lisbon, Lisbon, Portugal;

[4]Department of Microbiology and Parasitology, University of Buea, Buea, Cameroon

*** Correspondence:** Clotilde Carlow, Division of Genome Biology, New England Biolabs, Ipswich, MA 01938. Tel: 978 380 7263, Fax: 978 921 1350, e-mail: carlow@neb.com.

Keywords: *Mansonella*, Filaria, Neglected Tropical Disease, Genomics, Evolution.

Abstract

The filarial parasites *Mansonella ozzardi* and *Mansonella perstans*, causative agents of mansonellosis, infect hundreds of millions of people worldwide, yet remain among the most understudied of the human filarial pathogens. *M. ozzardi* is highly prevalent in Latin American countries and Caribbean Islands, while *M. perstans* is predominantly found in sub-Saharan Africa as well as in a few areas in South America. In addition to the differences in their geographical distribution, the two parasites are transmitted by different insect vectors, as well as exhibit differences in their responses to commonly used anthelminthic drugs. The lack of genome

information has hindered investigations into the biology and evolution of *Mansonella* parasites and understanding the molecular basis of the clinical differences between species. In the current study, high quality genomes of two independent clinical isolates of *M. perstans* from Cameroon and two *M. ozzardi* isolates one from Brazil and one from Venezuela are reported. The genomes are approximately 76 Mb in size, encode about 10,000 genes each, and are largely complete based on BUSCO scores of about 90%, similar to other completed filarial genomes. These sequences represent the first genomes from *Mansonella* parasites and enabled a comparative genomic analysis of the similarities and differences between *Mansonella* and other filarial parasites. Horizontal DNA transfers (HDT) from mitochondria (nuMTs) as well as transfers from genomes of endosymbiotic *Wolbachia* bacteria (nuWTs) to the host nuclear genome were identified and analyzed. Sequence comparisons and phylogenetic analysis of known targets of anti-filarial drugs diethylcarbamazine (DEC), ivermectin and mebendazole revealed that all known target genes were present in both species, except for the DEC target encoded by *gon-2* gene, which is fragmented in genome assemblies from both *M. ozzardi* isolates. These new reference genome sequences will provide a valuable resource for further studies on biology, symbiosis, evolution and drug discovery.

1 Introduction

A large part of the world's population is infected with one or more filarial nematode parasites resulting in several distinct diseases often referred to as a group under the general term filariasis. The filarial parasites *Mansonella ozzardi* and *Mansonella perstans* are the two main causes of mansonellosis in humans. Infections are transmitted by infected biting midges (mostly of the genus *Culicoides*) or blackflies (genus *Simulium*) that release infective stage larvae during a blood meal. Adult male and female worms have been observed rarely in humans since they are presumed to live in the serous body cavities and/or subcutaneous tissues. Viviparous females release unsheathed

microfilaria which circulate in the blood, available for ingestion by an appropriate vector during feeding. Unlike some of the other human filarial diseases, *Mansonella* infections do not present with a distinct pathology or specific clinical picture. Infected individuals are often asymptomatic or may have a variety of clinical features including subcutaneous swellings, aches, pain, fever, headache, pruritis, corneal lesions and eosinophilia (Simonsen et al., 2011; Lima et al., 2016; Mediannikov & Ranque 2018; Vianna et al., 2012; Ferreira et al 2021). Other manifestations include immunosuppression (Ritter et al., 2018) with important ramifications such as increased susceptibility to other infections and reduced efficacy of vaccine programs (Ta-Tang et al., 2018).

Even though mansonellosis ranks first in prevalence among the human filariases, affecting hundreds of millions of people in Africa and Central and South America, it remains one of the least studied and most neglected of all filarial diseases (Downes & Jacobsen, 2010; Simonsen et al., 2011; Lima et al., 2016; Mediannikov & Ranque, 2018; Ta-Tang et al., 2018; Ta-Tang et al., 2021; Sandri et al., 2021; Bélard & Gehringer 2021). Indeed, much of the somewhat limited information available has been obtained serendipitously during studies on other filarial parasites or malaria, in which individuals were co-infected with *Mansonella* parasites. *M. perstans* is considered the most common filariasis in Central Africa with infection prevalence often very high (up to 96%) in endemic areas, even among children (Downes and Jacobsen, 2010; Simonsen et al., 2011; TaTang et al., 2018). It is also found in northern regions of the Amazon rainforest and the Caribbean coast in South America (TaTang et al., 2018; Bélard and Gehringer, 2021), with molecular evidence supporting introduction of *M. perstans* to the American continent from Africa because of the slave trade (Tavares da Silva et al., 2017). *M. ozzardi* infections have been reported from many countries in South America, Central America, and some Caribbean Islands (Lima et al., 2018; TaTang et al.,2018; Tavares de Silva et al., 2017; Raccurt et al., 2018; Ferriera et al., 2021). Co-infections of *M. perstans* and *M. ozzardi*, together with other filarial parasites, have also been reported in South America (Kozek et al., 1983; Calvopina et al.,

2019), making their treatment challenging as the two species respond differently to the commonly used anti-filarial drugs such as ivermectin, diethylcarbamazine (DEC) or mebendazole (Chadee et al., 1995; TaTang et al., 2018). The recent development of highly specific, simple molecular tests which can distinguish *M. perstans* and *M. ozzardi*, and do not cross-react with other filarial parasites (Poole et al., 2019), will facilitate an effective treatment regimen. The antibiotic doxycycline, targeting the endosymbiotic *Wolbachia* bacteria of *M. perstans* and *M. ozzardi*, has been shown to clear the microfilarial stage from the blood (Coulibaly et al., 2009; Debrah et al., 2019; Crainey et al., 2020) highlighting the importance of *Wolbachia* in these species. The *Wolbachia* genomes, *w*Mpe from *M. perstans*, and *w*Moz from *M. ozzardi* have recently been published (Sinha et al., 2022), providing a valuable resource for the discovery of new antibiotics and studies on the symbiotic relationship and evolution.

Here we provide the first *de novo* assemblies of the nuclear genomes from *M. perstans* and *M. ozzardi* using combined PacBio long-read and Illumina paired-end sequencing data. We used customized metagenomic assembly and binning pipelines, and extensive quality checks to ensure accurate assemblies from a mixture of DNA from the *Mansonella* parasites, their human host, and the associated microbiome. These methods were applied to generate genomes from two clinical isolates of each species, providing a unique opportunity for intra-species and inter-species comparisons and exploring features of sequence evolution of these parasites. We analyzed the genomic regions containing horizontally transferred DNA from mitochondria (nuMTs), as well horizontally transferred DNA from their endosymbiont *Wolbachia* (nuWTs). Phylogenetic analysis of orthologs of genes that encode drug targets of common anti-filarial drugs such as DEC, ivermectin and mebendazole revealed intriguing differences in the presence and absence of these targets between *M. perstans* and *M. ozzardi* that might have clinical implications for the treatment of mansonellosis.

2 Methods

2.1 Ethics statement

All research involving human subjects was approved by the appropriate committee and performed in accordance with all relevant guidelines and regulations. Informed consent was obtained from all participants or their legal guardians. These samples have also been described and utilized in previous studies (Poole et al., 2019; Sinha et al., 2022).

Ethical clearance for *M. perstans* samples was obtained from the National Institutional Review board, Yaoundé (REF: N° 2015/09/639/CE/CNERSH/SP) and administrative clearance from the Delegation of Public Health, South-West region of Cameroon (Re: R11/MINSANTE/SWR/ RDPH/PS/259/382). Approval for the study was granted by the "National Ethics Committee of Research for Human Health" in Cameroon. Special consideration was taken to minimize the health risks to which any participant in this study was exposed. The objectives of the study were explained to the consenting donor after which they signed an informed consent form. The participant's documents were given a code to protect the privacy of the study subject. At the end of the study, the donor received a cure of mebendazole (100 mg twice a day for 30 days).

The study protocols for *M. ozzardi* were approved by the Institutional Review Board of the Institute of Biomedical Sciences, University of São Paulo, Brazil (1133/CEP, 2013) as part of a previous study (Lima et al., 2018). Written informed consent was obtained from all patients, or their parents or guardians if participants were minors younger than 18 years. Diagnosed infections were treated with a single dose of 0.2 mg/kg of ivermectin after blood sampling (Lima et al., 2018).

2.2 Illumina library construction and sequencing

The parasite materials, DNA extractions, Illumina library preparation and sequencing have been described in detail previously (Poole et al., 2019; Sinha et al., 2022). Briefly, two clinical isolates of *M. perstans*, named Mpe-Cam-1 and Mpe-Cam-2, were obtained as microfilariae in a blood draw from consenting individuals in Cameroon. One clinical isolate of *M. ozzardi*, named Moz-Brazil-1, was obtained from a consenting individual as part of a previous study in Brazil, while another isolate from Venezuela, named Moz-Venz-1, was generously donated by Izaskun Petralanda in 1989. The DNA samples from Mpe-Cam-1, Mpe-Cam-2, and Moz-Brazil-1 were treated with the NEBNext Microbiome DNA enrichment to reduce human DNA contamination before being used for library preparation. Illumina libraries for all 4 isolates were constructed using the NEBNext Ultra II FS DNA Library Prep Kit (New England Biolabs Inc., USA) as per standard protocol and sequenced using the Illumina MiSeq and NextSeq500 platforms (paired end, 150 bps). Bioinformatic analysis programs were run with their default parameters unless otherwise stated. The bioinformatic analysis pipeline and scripts are available from https://github.com/aWormGuy/Mansonella-Genomes-Sinha-et-al.-2023.

2.3 PacBio library sequencing and *de novo* assembly

A total of 250 ng of genomic DNA from *M. perstans* isolate Mpe-Cam-1 was used for sequencing on PacBio RSII platform as per manufacturer's instructions. Consensus reads from raw PacBio reads were obtained using the protocol RS_PreAssembler.1 (parameters minLen=1000, minQual=0.80, genomeSize=100000000) on PacBio SMRT Portal (smrtanalysis_2.3.0.140936.p5.167094). Reads originating from human genomic DNA were removed by first by mapping the consensus reads to the human genome (grch38) using minimap2 v2.17-r941 (Li, 2018) followed by discarding all mapped reads (MAPQ score > 20). The remaining reads were assembled using canu v2.2 (Koren et al., 2017) (parameters genomeSize=90m correctedErrorRate=0.045). The assembly was iteratively polished

over 12 rounds using "Resequencing" protocol on PacBio SMRT Portal (v 6.0.0.47841) (parameters --minMatch 12 --bestn 10 --minPctSimilarity 70.0 --refineConcordantAlignments) till no new variants could be detected by the Resequencing pipeline. The assembly contiguity was further improved by using finisherSC software v2.1 (Lam et al., 2015), a scaffolder that uses raw PacBio reads while accounting for any genomic repeats in the assembly. The assembly was polished using the Illumina data for the same isolate via the polca software (Zimin and Salzberg, 2020). Finally, gap-filling and heterozygosity removal was performed on this assembly using Illumina reads as an input to the Redundans pipeline v0.14a (Pryszcz and Gabaldón, 2016). The assembly quality was evaluated using quast (Gurevich et al., 2013) at each step.

2.4 Metagenomic assembly using Illumina reads

Illumina raw reads were processed using the BBTools package (https://jgi.doe.gov/data-and-tools/bbtools/). Duplicate raw reads and bad tiles were removed using the clumpify.sh and filterbytile.sh utilities. Adapter trimming, removal of phiX reads, and reads with 13-nt or longer stretches of the same nucleotide (poly-G, poly-C, poly-A, poly-T) was performed using bbduk.sh. Human host reads were removed by mapping against the human genome (grch38) using bbmap.sh utility. The quality metrics of the processed reads at each step were assessed using FastqC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Reads from different runs of the libraries prepared from the same genomic DNA sample were combined and used as an input for the assembly of the metagenome using metaSpades (Nurk et al., 2017).

2.5 Metagenomic binning to identify *Mansonella* contigs

Binning of metagenomic contigs to identify sequences originating from *Mansonella* genomes was carried out based on the BlobTools software (Laetsch and Blaxter, 2017) and additional curation, as

described previously (Sinha et al., 2022). Briefly, for each of the isolates Mpe-Cam-2, Moz-Brazil-1 and Moz-Venz-1, the Illumina reads were mapped back to the metagenomic contigs produced by metaSpades using bowtie2 (Langmead and Salzberg 2012). For the PacBio assembly obtained from the Mpe-Cam-1 sample, the corresponding Illumina reads were mapped to this assembly using bowtie2. Sequence similarities of assembled contigs to the NCBI nt database and Uniport database was carried out using blastn and diamond in BlobTools package. The bins annotated as "Nematoda" and "Proteobacteria" by BlobTools were analyzed further to retrieve sequences originating from *Mansonella*. Contigs that were marked as Proteobacteria but co-localized with the Nematoda blobs in the blobplots, were included in the *Mansonella* bins as they represented potential examples of horizontal gene transfer from *Wolbachia* to *Mansonella* host. These HGT candidates were validated by inspection of blastn and blastx hits of these contigs to the NCBI nt and nr databases respectively.

Contigs originating from mitochondrial DNA were identified by analysis of nucmer alignments between assembled sequences and the previously reported mitochondrial genomes, GenBank accessions MT361687.1 and KX822021.1 for *M. perstans* (Chung et al., 2020) and *M. ozzardi* (Crainey et al., 2018) respectively. The contigs with more than 98% sequence identity to published mitochondrial genomes were marked as mtDNA and removed from the nuclear assembly.

2.6 Assembly refinements and heterozygosity removal for Illumina assemblies

For each Illumina-only assembly, scaffolding, gap-filling and heterozygosity removal was performed using the corresponding Illumina reads in the Redundans pipeline v0.14a (Pryszcz and Gabaldón, 2016).

2.7 Genome annotations and comparative analysis

Identification of repetitive elements in the assembled genomes were performed using RepeatModeler version 2.0.3 (Flynn et al., 2020) and classified against the Dfam database (Storer et al., 2021). First

round of repeat annotations was performed on a fasta file combining all assembled genomes. The repeat sequence families identified from this run were used as a reference library in RepeatMasker 4.1.2-p1 (http://repeatmasker.org/) to mask repeats in each of the 4 genome assemblies independently. The outputs from RepeatMasker were processed with parseRM.pl utility (Kapusta et al. 2017) to generate summary statistics on repats classified as DNA transposons and retroelements.

Gene predictions were carried out using the braker2 pipeline v2.2.6 (Bruna et al., 2020), with a reference proteome comprised of all metazoan proteins in OrthoDB version 10.1 (Kriventseva et al., 2019) and all filarial proteomes downloaded from WormBase ParaSite (Howe et al., 2017). To ensure that the same parameters and hmm models are applied to all assemblies, the 4 assemblies were combined into a single file, and braker2 pipeline was run on this composite dataset. The genes predicted from this run were then assigned to respective isolates by parsing the gff3 output from braker2. The completeness of the protein-coding content of genomes was assessed using the BUSCO pipeline v5.0 beta using the “nematoda_odb10” reference dataset (Simão et al., 2015). OrthoFinder v2.4 (Emms and Kelly, 2019) was used to infer orthologs of protein coding genes across multiple nematode genomes. The set of proteomes included all 17 filarial species with publicly available genomes. The species *Setaria digitata* and *Thelazia callipaeda* from the same infraorder Spiruromorpha were included as the immediate outgroups, while *Caenorhabditis elegans* and *Pristionchus pacificus* were included as relatively more distant outgroups. The protein and transcript sequences were downloaded from WormBase parasite where available. Although the genome sequences for filarial nematodes *Cruorifilaria tuberocauda*, *Dipetalonema caudispina*, *Dirofilaria immitis*, *Litomosoides brasiliensis* and *Madathamugadia hiepei* (Lefoulon et al., 2020) were available from NCBI (GenBank accessions GCA_013365365.1, GCA_013365325.1, GCA_013365355.1, GCA_013365375.1 and GCA_013365335.1 respectively), the corresponding gene annotations for these assemblies were not available in NCBI. Hence, repeat-modeling, repeat-masking and gene

annotation using braker2 pipeline was performed on each of these species as described for *M. perstans* and *M. ozzardi* above, and have been made available at https://github.com/aWormGuy/Mansonella-Genomes-Sinha-et-al.-2023/tree/main/gene_predictions_on_Lefoulon2020_genomes.

available a. The source databases, the corresponding accession numbers, filenames, citations and BUSCO scores for all the genomes used in comparative analysis are listed in Supplementary Material Table 1.

Whole-genome alignments for global sequence similarity and synteny between various repeat-masked genomes were performed using minimap2 v2.17-r941 (Li, 2018) and visualized using the JupiterPlot tool (https://github.com/JustinChu/JupiterPlot). For global sequence comparisons across multiple filarial genomes, the average nucleotide identity (ANI) scores were calculated using the OrthoANIu tool (Yoon et al., 2017).

Comparison and analysis of sequence differences, including structural difference, between different pairs of genomes, namely (A) Mpe-Cam-1versus Mpe-Cam-2 (B) Moz-Brazil-1 versus Moz-Venz-1 (C) Mpe-Cam-1 versus Moz-Brazil-1 were performed using nucdiff version 2.0.3 (Khelik et al., 2017). The localization of the sequence variants within gene bodies and exons was calculated through the intersectBed tool in the bedtools package v2.29.2 (Quinlan and Hall, 2010). The effect of sequence variants on the coding potential of affected genes was determined using snpEff version 5.1d (Cingolani et al., 2012).

2.8 Phylogenetic analysis

For genome-wide phylogenetic analysis, single copy orthologs across various nematodes were identified from the OrthoFinder analysis described above. The corresponding protein sequences were aligned using mafft v7.149b (Nakamura et al., 2018) and trimmed using Gblocks v0.91b (Talavera and Castresana, 2007). Phylogenetic analysis of the resulting supermatrix of 823,832 amino acids was carried out using iqTree v2.1.2 (Nguyen et al., 2015). The best fit substitution model was chosen automatically using ModelFinder implementation within iqTree (Kalyaanamoorthy et al., 2017). Bootstrap support values were calculated based on (i) the SH-like approximate likelihood ratio test (Guindon et al., 2010), and (ii) ultrafast bootstrap (Hoang et al., 2018) with 1000 replicates each (iqTree command-line options '-B 1000 -alrt 1000 -abayes -lbp 1000').

For phylogenetic analysis of specific genes and gene families, their homologs and the corresponding protein sequences were obtained from the OrthoFinder output. The protein domains for all the sequences were annotated using NCBI CD search and the NCBI CDD database (Lu et al., 2020). The protein sequences which lacked the domains present in the *C. elegans* homologs, or were too short in length (less than 60% of the length of the *C. elegans* homolog) were inspected and excluded from the final analysis. The multiple sequence alignments and tree inference was carried out using mafft and iqtree as described above. The consensus trees were edited and annotated using the iTOL webserver (Letunic and Bork, 2021) and Adobe Illustrator.

2.9 Identifying nuWTs and nuMTs

The assembled nuclear genomes were aligned to the respective *Wolbachia w*Mpe or *w*Moz (Sinha et al., 2022) genomes using nucmer v4.0.0beta2 (Marçais et al., 2018). Aligned fragments ("MuM"s) that were separated by less than 50 bp, were combined into a single nuWT locus. For manual

validation of HGT of *Wolbachia* sequences into filarial contigs, the *Wolbachia* regions were verified to be flanked by nematode-like sequences and protein coding genes from nematodes.

A similar strategy was employed for nuMT identification, with nucmer alignments against respective mitochondrial sequences, (Chung et al., 2020, Crainey et al., 2018), and blastx against nematode nuclear proteomes. The nuMT loci with more than 85% AT-content were removed to avoid spurious matches to low-complexity regions.

Comparison of nuMT or nuWT loci for conservation of their sequence and position across different genomes was based on their sequence comparisons using megablast coupled with the synteny analysis described above. The sequences for nuMT or nuWT loci were extracted for each genome and compared using megablast, and only hits with more than 90% identity over 90% of the sequences were analyzed further as candidates for conserved loci. In the next step, if these candidate loci were present on syntenic contigs, they were marked as conserved loci across strains or species. A blastx analysis of the nuWT or nuMT sequences against the respective mitochondrial or *Wolbachia* proteomes was performed to identify any intact protein coding genes.

3 Results

3.1 PacBio sequencing and *de novo* assembly generates a highly contiguous *M. perstans* genome

PacBio sequencing and *de novo* genome assembly of genomic DNA isolated from *M. perstans* isolate "Mpe-Cam-1", together with BlobTools analysis for metagenomic binning resulted in a high-quality *M. perstans* genome. The genome assembly is comprised of 581 scaffolds with a total size of 76 Mb, with the largest scaffold greater than 1.2 Mb and a N50 size greater than 285 kb (Table 1). Based on

these quality metrics, the Mpe-Cam-1 assembly ranks highly in comparison to all filarial genomes available so far (Supplementary Material Table 1). Repetitive elements account for 9.7% of the genome, of which about 27% of the repeat sequence length was comprised of low complexity and simple repeats. Long terminal repeat (LTR) retrotransposons comprised 17% of all the repeats, while 50% of the repeat content is comprised of unknown, novel repeat families (Supplementary Material Table 2). The genome has 9,460 protein coding genes, with a BUSCO score of 93.2 % (Supplementary Figure 1) when evaluated against a set of 3,131 core proteins conserved across the nematode phylum, further indicating the high quality of this genome assembly.

Another clinical isolate from Cameroon, "Mpe-Cam-2" was sequenced and assembled *de novo* based on Illumina sequencing. Although the quality metrics of this Illumina short-reads assembly are lower than the PacBio long-read Mpe-Cam-1 assembly (Table 1), the repeat content, (Supplementary Material Table 2), the number of predicted genes and their BUSCO scores are similar to Mpe-Cam-1 (Supplementary Figure 1).

3.2 Genome sequencing and assembly of *M. ozzardi* from South America

For *M. ozzardi*, two genome assemblies from Latin American isolates, "Moz-Brazil-1" from Brazil and "Moz-Venz-1" from Venezuela, were generated based on Illumina sequencing reads. These assembled genomes are about 76 Mb (Table 1), similar in size to the *M. perstans* genomes described above. The number of predicted genes, their BUSCO scores (Supplementary Figure 1) and repeat content (Supplementary Material Table 2) are also similar to those present in *M. perstans* genomes.

3.3 Analysis of genome variation across *M. perstans* and *M. ozzardi* and among multiple clinical isolates

The availability of genomes from two *Mansonella* species and multiple isolates provided a unique opportunity for analysis of whole genome sequence variation both between and within the two species. The genome-wide average nucleotide identity (gANI) scores for isolates within species are 99.5% (Table 2) in comparison to the score of 89.69% between the two *Mansonella* species. Visualizations of whole genome alignments between isolates as Jupiter plots (Figure 1) reveal high levels of synteny. The isolates Mpe-Cam-1and Mpe-Cam-2, collected from the same geographical region within a few years apart are most similar to each other (Figure 1A). Relatively more genomic re-arrangements (represented as crisscrossing lines in Jupiter plots) can be seen between *M. ozzardi* isolates Moz-Brazil-1 and Moz-Venz-1 (Figure 1B), which have been collected from different geographical regions (Brazil, Venezuela) and at about 30 years apart. This indicates that the evolution of *Mansonella* genomes is dependent on geographical and temporal separation. Whole-genome alignments comparing *M. perstans* and *M. ozzardi* genomes also demonstrate high levels of synteny (Figure 1C), although lower than the within-species levels, consistent with the trend in their gANI scores.

At the nucleotide level, the numbers of simple substitutions, insertions and deletions (Table 2) were 10 to 30-fold higher in across-species comparisons (6.22%), as compared to within-species (0.2% and 0.29%). In all the pairwise comparisons, most of the sequence variation (>52.8%) was localized to the intergenic portions of the genomes outside the protein-coding genes CDS regions. The proportion of variants within protein-coding genes ranged from about 5% to 10% (Table 2). Most of the nucleotide changes (54% to 60%) that occurred within coding regions, were mutationally silent as they resulted in synonymous amino-acid changes, 39 to 44% of coding variants did result in missense or non-synonymous changes in the predicted protein sequences and only a small fraction (0.21 % to

0.77%) of the coding variants resulted in a nonsense mutation. Overall, the intra-species differences between the two *M. ozzardi* isolates from different geographical areas were higher than the corresponding intra-species differences of two *M. perstans* isolates which were both collected in Cameroon. A similar trend was observed for structural differences as well, as the number of relocations, reshufflings and inversions of contig fragments was observed to be much higher in cross-species comparisons (Table 2).

3.4 Orthology and phylogenomic analysis of filarial nematodes

An Orthofinder analysis of proteomes from all available filarial genomes (Supplementary Material Table 1), with *S. digitata* and *T. callipaeda* as immediate outgroups from the sub-order Spiruromorpha, and *C. elegans* and *P. pacificus* as distant outgroups identified 3,829 orthogroups conserved across all these nematodes (Figure 2). This analysis also highlighted 110 orthogroups specific to *Mansonella* genus, 320 orthogroups specific to *Onchocerca* genus, and 130 orthogroups specific to *Brugia*. Most of these proteins had only weak similarities to other nematode proteins (median sequence identity and coverage both around 33%), making them good candidates for genus-specific biomarkers (Supplementary Material Table 3).

Among the orthogroups conserved across all nematode genomes analyzed here, 465 orthogroups comprised of single copy orthogroups (SCOs) and were used for analysis. The phylogenomic tree (Figure 3) generated using these 465 orthogroups recapitulated the phylogenetic relationships based on analysis of mitochondrial genomes (McCann et al., 2021) and the monophyletic feature of major clades ONC2, ONC3, ONC4 and ONC5 based on 7 genes (Lefoulon et al., 2015).

3.5 Analysis of nuMTs, horizontal transfers of mitochondrial DNA into nuclear genomes

Transfer of mitochondrial DNA fragments into the nuclear genome has been observed for multiple filarial nematodes (Grote et al., 2017) as well other eukaryotes (Hazkani-Covo et al., 2010). Since mitochondrial markers are often used as barcodes for species identification and phylogenetic analysis (Crainey et al., 2018), it is important to distinguish between genuine mitochondrial sequences versus nuMTs (Song et al., 2008, Crainey et al., 2018). In the newly sequenced *Mansonella* genomes, 39 to 94 nuMT loci with median sizes ranging around 250 bp were found across different isolates, with the largest loci ranging from 2.35 kb to 3.69 kb in size (Table 3). None of the nuMTs was found to carry an intact protein coding gene. The total number and the combined size of nuMTs in was observed to be higher in *M. ozzardi* as compared to *M. perstans* (Table 3). The locations and sequences of nuMTs were poorly conserved across *M. perstans* and *M. ozzardi*. For example, in a comparison of the 48 Mpe-Cam-1 nuMTs with the 82 Moz-Brazil-1 nuMTs, only 2 conserved loci were found (Supplementary Material Table 4). Even within the isolates of each species, approximately 35% and 60% of nuMT loci were conserved in *M. perstans* and *M. ozzardi* respectively (Supplementary Material Table 4), while the remainder were isolate-specific, indicating a fast turnover of these elements. The existence of multiple cox1 derived nuMT pseudogenes has been reported in *M. ozzardi* (Crainey et al., 2018). Similarly, multiple fragments of the *cox1* gene were also observed in both *M. ozzardi* isolates (Supplementary Figure 2). However, an exact match for these sequences could not be found in either of the Moz-Brazil-1 or Moz-Venz-1 assemblies, providing further evidence for poor conservation of such nuMTs across different isolates.

When the nuMT sequences from each isolate were mapped back to the mitochondrial genomes, they were found to cover almost the entire mitochondrial genome, without any preference for specific regions to be transferred into the nuclear genome (Supplementary Figures 2 and 3).

3.6 Analysis of nuWTs, horizontal transfers of *Wolbachia* endosymbiont DNA into host nuclear genomes

Instances of transfer of DNA sequences from *Wolbachia* endosymbionts into the nuclear genomes of their hosts, termed "nuWTs" for nuclear-*Wolbachia* transfers, have been observed frequently in filarial nematodes (McNulty et al., 2010; Cotton et al., 2016; Grote et al., 2017; Hotopp et al., 2017) as well as arthropod hosts (Klasson et al., 2014; Tvedte et al., 2022). The *Mansonella* isolates reported here harbor about 300 nuWT loci in each isolate, with median size about 180 bp, adding up to a total of 76 kb to 88 kb), with the largest nuWT (2.73 kb) observed in Mpe-Cam-1 (Table 4). None of these nuWTs were found to harbor any intact protein coding gene. Analysis of synteny and sequence identity of nuWTs across species identified only 34 conserved nuWTs between *M. perstans* and *M. ozzardi* (Supplementary Material Table 5). Interestingly, about 80% (260 from *M. perstans* and 280 from *M. ozzardi*) were found to be conserved in within-species comparisons of *M. perstans* and *M. ozzardi* respectively (Supplementary Material Table 5).

3.7 Analysis and comparison of genes encoding drug targets for the filaricides diethylcarbamazine, ivermectin and albendazole

The main preventive intervention for filarial infections is based on mass drug administration (MDA) using diethylcarbamazine (DEC), ivermectin (IVM) and albendazole (ALB) (Simonsen et al 2010). Although DEC has been successfully used to treat *M. perstans* infections, the drug has been found to be ineffective against *M. ozzardi* (Chadee et al., 1995). To investigate potential reasons for differences in drug susceptibility between *M. perstans* and *M. ozzardi*, the genes encoding the drug targets for DEC, ivermectin and albendazole, were identified and analyzed. A search for the DEC target *gon-2* (Verma et al., 2020), a gene that encodes a Transient Receptor Protein M (TRPM) channel subunit, identified two potential homologs in *M. perstans* but one in *M. ozzardi*. Phylogenetic analysis of the *gon-2* homologs in *Mansonella* and other nematode genomes revealed two distinct, paralogous clades, denoted as Clade 1 and Clade 2, which represent filarial orthologs of *Cel-gon-2* and *Cel-gtl-2* respectively (Figure 4). The nematodes known to be susceptible to DEC,

including *C. elegans* and filarial parasites *B. malayi L. loa*, and *W. bancrofti*, all have a *gon-2* ortholog in Clade 1. Interestingly, *M. perstans*, which is treatable with DEC, also has a corresponding ortholog in Clade 1 while in *M. ozzardi*, which is refractory to DEC, the orthologous gene was found in two fragments on the ends of two separate contigs in each of the two isolates (Supplementary Figure 4). Artificially combining the two gene fragments could reconstruct the full gene except for a few missing bases between exons 19 and 20 (Supplementary Figures 5 and 6). It is possible that the gene is intact in *M. ozzardi* but is split over two contigs due to incomplete genome assembly. To explore this further, the raw Illumina reads from each isolate were mapped to a synthetic contig constructed by joining the contigs carrying the *gon-2* fragments in appropriate orientations with a 50 bp spacer marked by Ns, and all read-pairs spanning the potentially un-assembled gap were investigated (Supplementary Figures 7 and 8). This analysis showed an abrupt drop in read coverage in the potential gap region, even though the contig terminals had sufficient coverage. Although a few read-pairs (less than 10 counts) were observed to span across the gap, the size of gap estimated via the insert size of the paired-end reads was around 2.5 kb, much larger than the insert size of the sequencing library. In contrast, the read-coverage across the corresponding region in *M. perstans* is uniform (Supplementary Figure 9). Together, these observations indicate that the synthetic contig is unlikely to represent the true genomic sequence, and therefore the *gon-2* gene is fragmented in both *M. ozzardi* isolates. Both *M. perstans* and *M. ozzardi* do possess orthologs of *Cel-gtl-2* in Clade 2, which share only 52 % sequence identity with the *C. elegans* GON-2 protein in Clade 1 and are unlikely to share their functions. Similar analysis for the other known DEC target genes, *trp-2*, *ced-11* and *slo*-1 identified the corresponding one-to-one orthologs for all three genes in all the filarial species including both isolates of *M. perstans* and *M. ozzardi* (Supplementary Figures 10, 11 and 12). Taken together, this suggests that the absence the *gon-2* ortholog in *M. ozzardi* is most likely responsible for the ineffectiveness of DEC against this parasite.

The targets of ivermectin as identified through genetic studies in *C. elegans* include various subunits of the Glutamate□gated Cl− channels (GluCl), encoded by genes *glc-1*, *glc-2*, *glc-3*, *glc-4*, *avr-14* and *avr-15* (Dent et al., 2000; Wolstenholme and Rogers, 2005). Ivermectin is effective in treating *M. ozzardi* infections (Basano et al., 2018), but is reported to be ineffective against *M. perstans* (Bregani et al., 2006). Phylogenetic analysis of these genes across filarial parasites shows that the homologs of *glc-1*, *glc-2*, *glc-3*, *avr-14* and *avr-15* form a family of paralogous genes with complex patterns of homology across filarial parasites (Figure 5), while *glc-4* has one-to-one orthologs across all genomes analyzed (Supplementary Figure 13). Orthologs for *Cle-glc-1*, *Cel-glc-2* and *Cel-avr-15* are not present in either of the *Mansonella* species (Figure 5). Both isolates of *M. perstans* and *M. ozzardi* have one ortholog most closely related to the *Cel-glc-3*, and one ortholog most closely related to to *Cel-avr-14*,

The anthelminthic drug mebendazole acts by binding to beta-tubulin protein and inhibiting microtubule formation (Lacey, 1990). Mebendazole is commonly used in combination DEC or ivermectin for treating *M. perstans* infections (Bregani et al., 2006) but its use for *M. ozzardi* has not been reported. The beta-tubulin gene family in *C. elegans* consists of 6 different members (*ben-1*, *tbb-1*, *tbb-2*, *tbb-4*, *tbb-6*, and *mec-7*) all of which have varying effects on drug sensitivity (Driscoll et al., 1989; Pallotto et al., 2022). A phylogenetic analysis of the beta-tubulin gene family in *Mansonella* and other nematodes reveals a complex pattern of expansion and contraction of this gene family across filarial species, with no direct orthologs for *ben-1* identifiable in any filarial nematodes (Supplementary Figure 14). The exact roles of different beta-tubulin homologs on drug susceptibility in *Mansonella* will require further study.

4 Discussion

To gain some insight into the neglected pathogens responsible for human mansonellosis, we report the first *de novo* genome assemblies from the filarial parasites *M. perstans* and *M. ozzardi*, providing reference genomes for each species. The genomes encode approximately 10,000 genes each and are high quality and largely complete based on universally conserved orthologs (BUSCO scores greater than 90%) relative to other filarial species. The availability and sequencing of two isolates for *M. perstans* and *M. ozzardi* also allowed higher confidence in calling presence or absence of genes in each species as it could be confirmed in two independent isolates. The *Mansonella* genomes average 76 Mb in length and are similar in size to the genomes described from other human and animal filarial parasites (Ghedin et al., 2007; Godel et al., 2012; Desjardins et al., 2013; Tallon et al., 2014; Lau et al., 2015; Foster et al., 2020, McCann et al., 2021). The sequencing of two clinical isolates of *M. perstans* and *M. ozzardi* also enabled a quantification of the genetic variability both between species as well in the same parasite from different regions. Most of the genetic differences observed between isolates and species are in the intergenic regions, and only a small fraction (10%) of variants affects coding exons. Relatively little is known about natural variation in the nuclear genomes of filarial parasites (McNulty et al 2013., Qing et al., 2021). In the heartworm *Dirofilaria immitis*, a closely related filarial parasite of dogs and cats, sequencing of two isolates from Europe and the United States revealed very little genetic variation in their nuclear or mitochondrial genomes (Godel et al., 2012). Other studies of genetic variation within isolates of a filarial species have used mitochondrial *cox1* gene (McNulty et al., 2013) or the mitochondrial genomes (Qing et al., 2021). Our study presents a more comprehensive analysis of variation in nuclear genomes within isolates. It will be interesting to investigate how these patterns evolve in other human filarial species.

Transfer of mitochondrial DNA fragments into the nuclear genome has been observed for multiple filarial nematodes (Grote et al., 2017). These nuMT loci are annotated for all *Mansonella* isolates using their reported mitochondrial genomes (Crainey et al., 2018; Chung et al., 2020). Most filarial

parasites possess the intracellular bacterial symbiont, *Wolbachia*, capable of transferring genetic material to the nuclear genome of its host. The genomes of *Wolbachia* endosymbionts *w*Mpe and *w*Moz from *M. perstans* and *M. ozzardi* have also been reported (Sinha et al., 2022), enabling an analysis of nuWTs, regions of DNA transfers from *Wolbachia* to its host nuclear genomes. While reported in other filarial species, the biological function and significance of such nuclear transfer events are not well understood. Moreover, the levels of conservation of such loci across multiple isolates of a species, or closely related species has not been previously studied. Our analysis combining synteny and sequence identity shows that the nuMT loci are poorly conserved in *Mansonella* even within the same species, and rarely harbor intact coding genes, suggesting that most of these transfers are not evolutionarily selected, and non-functional. For the nuWTs, relatively more loci were found to be conserved across the two species, but isolate specific acquisition of these nuWT fragments was also observed. However, none of these loci harbored any intact protein-coding genes, suggesting a lack of functional constraints on their evolution. This observation is consistent with the detection of mostly non-functional nuWTs in other filarial parasites (Grote et al., 2017) and suggests that these loci most likely arise randomly and degrade via genetic drift over shorter time scales, as seen in other organisms (Tvedte et al., 2022; Hazkani-Covo et al., 2010; Bensasson et al. 2003).

Despite their genetic relatedness, *M. perstans* and *M. ozzardi* differ in their geographical distribution, adaptation to insect vectors, and in their response to drug treatments. We have used these new genomes to explore the genetic basis for the reported differences in their response to commonly used anti-filarial drugs. Our analysis of the sequences and phylogeny of known gene products targeted by DEC revealed that *M. ozzardi* does not have an intact *gon-2* gene, one of the targets of this drug. In contrast, *M. perstans* as well as other filarial parasites that are susceptible to DEC have a *gon-2* ortholog. It can therefore be hypothesized that the absence of *gon-2* ortholog in *M. ozzardi* has

rendered it unsusceptible to DEC. Future research focusing on this gene sequence and its variation between different isolates of the two species can shed further light on its function.

All the genes known to be targets of ivermectin were found in both *M. perstans* and *M. ozzardi*. The observed differences in ivermectin susceptibility of the two parasites could not be readily explained by the subtle differences in sequences of these targets. There could be additional ivermectin targets that are not discovered yet but could play a role in the differential response of *M. perstans* and *M. ozzardi*. Although mebendazole is used for the treatment *M. perstans*, it was found to lack an ortholog of one of the targets encoded by the beta-tubulin gene *tbb-4*, while the corresponding ortholog is present in *M. ozzardi*. Given the complex patterns of homology for this gene family, and the absence of any reported effects of mebendazole on *M. ozzardi*, the significance of this observation remains to be determined. In addition to these observations, there could be yet other undiscovered filarial genes that have also have effects on drug susceptibility and could be responsible for differences between *M. perstans* and *M. ozzardi*. A functional validation of these hypotheses regarding drug and their target genes is currently not feasible, but future advances may be able to test these hypotheses. At least one of the parasites, *M. perstans*, can now be successfully cultured *in vitro* enabling functional investigations into its biology (Njouendou et al., 2019). Advances in genome editing techniques for non-model organisms including helminth parasites open further possibilities of genetic and functional studies, and control of parasites (Lok et al., 2016; Zamanian and Andersen, 2016; Liu et al., 2020; Kwarteng et al., 2021; Sankaranarayanan et al., 2021; Quinzo et al., 2022). The availability of the new *Mansonella* genomes reported here will be a critical genomic resource for such endeavors and for further studies on drug targets, as well as enabling insights into the biology and evolution of these parasites.

5 Conflict of Interest

Authors AS, ZL, CBP, LE, RDM, and CKSC are employees at New England Biolabs Inc., Ipswich, MA 01938 (USA), which manufactures biotechnology reagents including those used in Next Generation Sequencing. The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

6 Author Contributions

NFL, MUF, FFF, SW collected the clinical samples. AS, ZL, CBP, LE, RDM performed the sequencing experiments and bioinformatics analysis. AS, ZL, CBP, LE, RDM, CKSC analysed the data. AS, ZL, CBP, CKSC wrote the initial manuscript. AS, ZL, CBP, LE, RDM, MUF, SW, CKSC and all authors contributed to the final manuscript.

7 Funding

Research funding was provided by New England Biolabs, Ipswich, MA 01938, USA to AS, ZL, LE, RDM, CBP, CKSC. FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) supported the field research in Brazil through a research grant to M.U.F (2013/12723-7) and a doctoral scholarship to N.F.L. (2013/ 26928-0).

8 Acknowledgments

The authors thank the DNA sequencing core at New England Biolabs. The authors are grateful to the late Don Comb for initiating and inspiring research into NTDs at NEB.

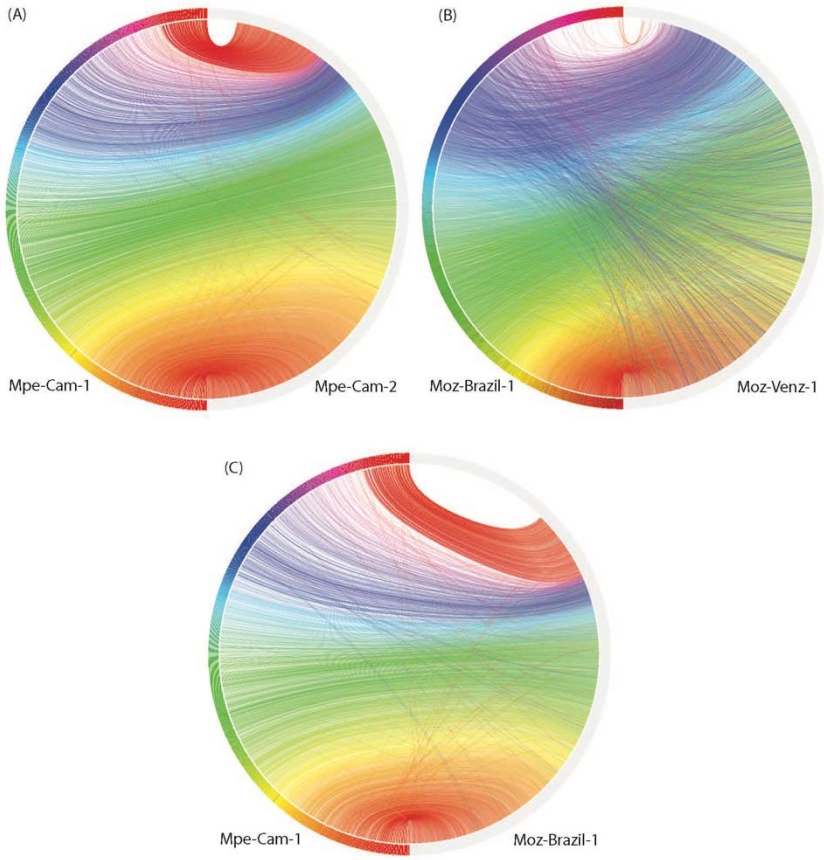
9 Data Availability Statement

Raw sequencing reads for the *M. per*stans isolates Mpe-Cam-1 and Mpe-Cam-2 have been deposited in the NCBI SRA database under BioProject accession numbers PRJNA917716 and PRJNA917721 respectively. Illumina sequencing data for *M. ozzardi* isolates Moz-Brazil-1 and Moz-Venz-1 have been deposited in the NCBI SRA database under BioProject accession numbers PRJNA917722 and PRJNA917766 respectively.

10 Figure Legends

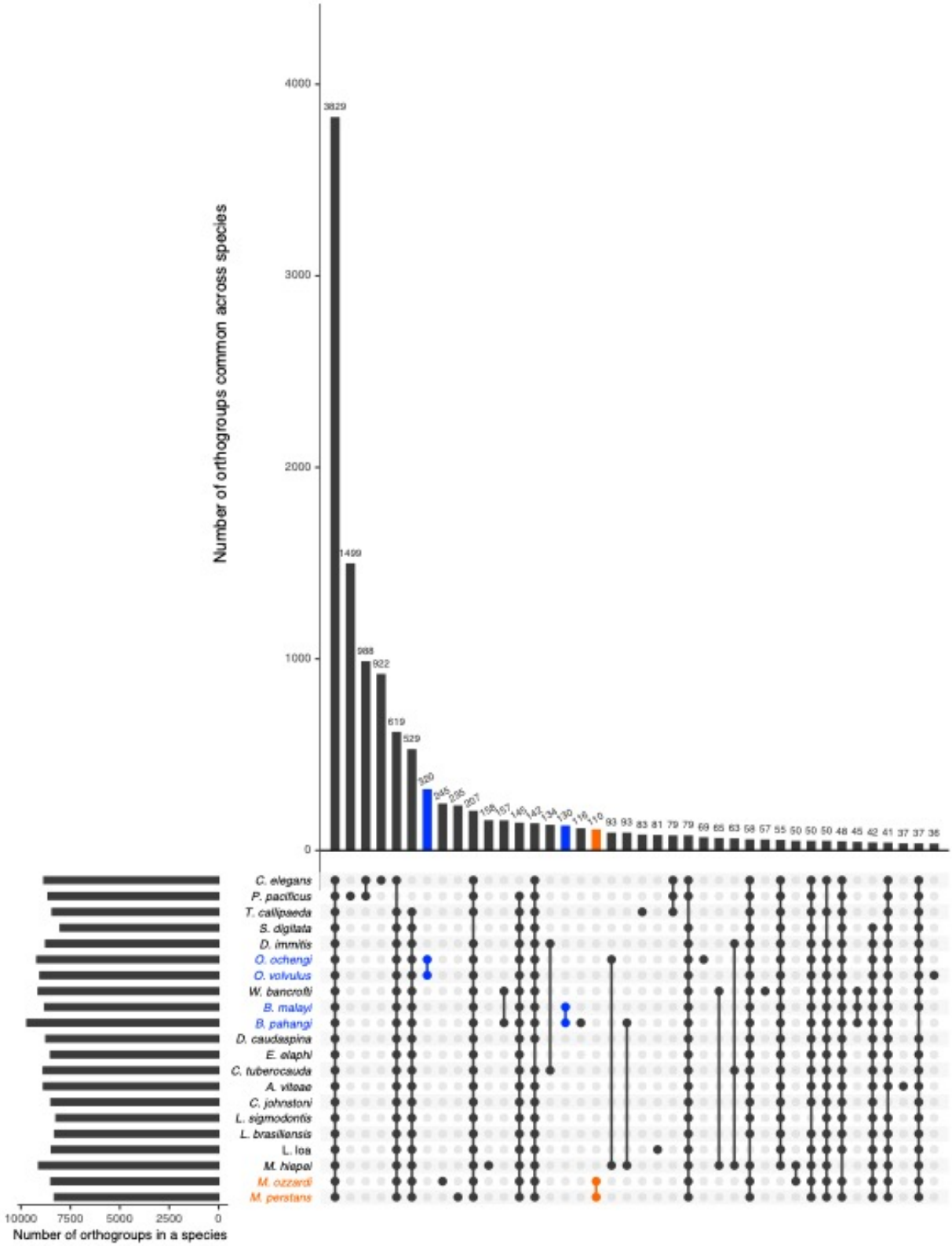
10.1 Figure 1. Comparisons of syntenic regions across the assembled *Mansonella* genomes.

Whole genome alignments were performed using minimap with default parameters. Syntenic regions were visualized using Jupiter plots, comparing (A) The two isolates of *M. perstans* (B) The two isolates of *M. ozzardi*, and (C) the cross-species comparison between Mpe-Cam-1 and Moz-Brazil-1.



10.2 Figure 2. Orthology analysis across filarial parasites.

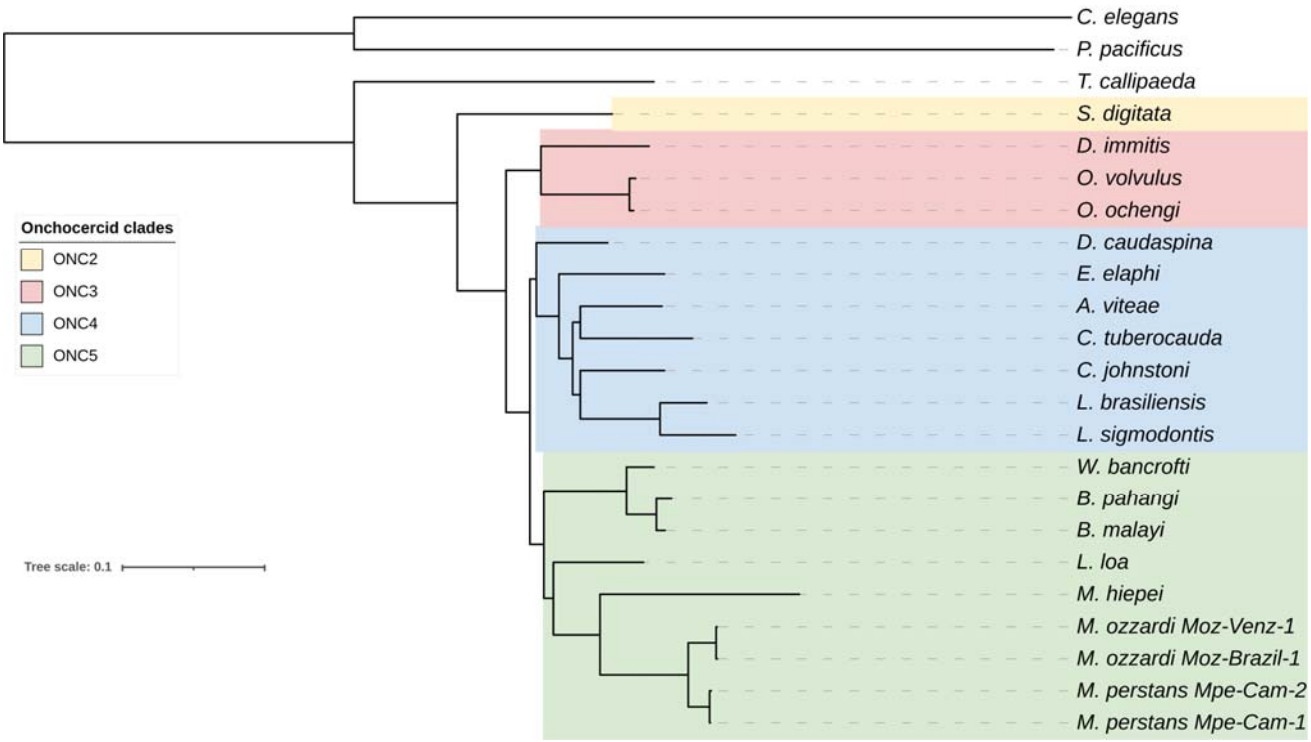
Homologs conserved across various filarial species, as well as the Onchocercid nematodes *S. digitata* and *T. callipaeda*, and free-living model nematodes *P. pacificus* and *C. elegans* from Clade V were identified using OrthoFinder. The set of all 263,019 proteins from these nematode genomes were grouped into 16,564 orthogroups. The number of orthogroups identified in each species are represented by horizontal bars in the Upset plot, and the number of orthogroups common across different species are represented by vertical bars. The presence of an orthogroup in a species is indicated by a filled, colored dot while the absence is indicated by empty gray dots. Genus-specific orthogroups are shown as orange dots/bars for *Mansonella*, and blue dots/bars for *Onchocerca* and *Brugia*.


Number of orthogroups common across species
4000
3000
2000
1000
0
3829
1499
988
922
619
529
320
245
235
207
158
157
149
142
134
130
116
110
93
93
83
81
79
79
69
65
63
58
57
55
50
50
50
48
45
42
41
37
37
36
C. elegans
P. pacificus
T. callipaeda
S. digitata
D. immitis
O. ochengi
O. volvulus
W. bancrofti
B. malayi
B. pahangi
D. caudaspina
E. elaphi
C. tuberocauda
A. viteae
C. johnstoni
L. sigmodontis
L. brasiliensis
L. loa
M. hispai
M. ozzardi
M. perstans
10000
7500
5000
2500
0
Number of orthogroups in a species

550

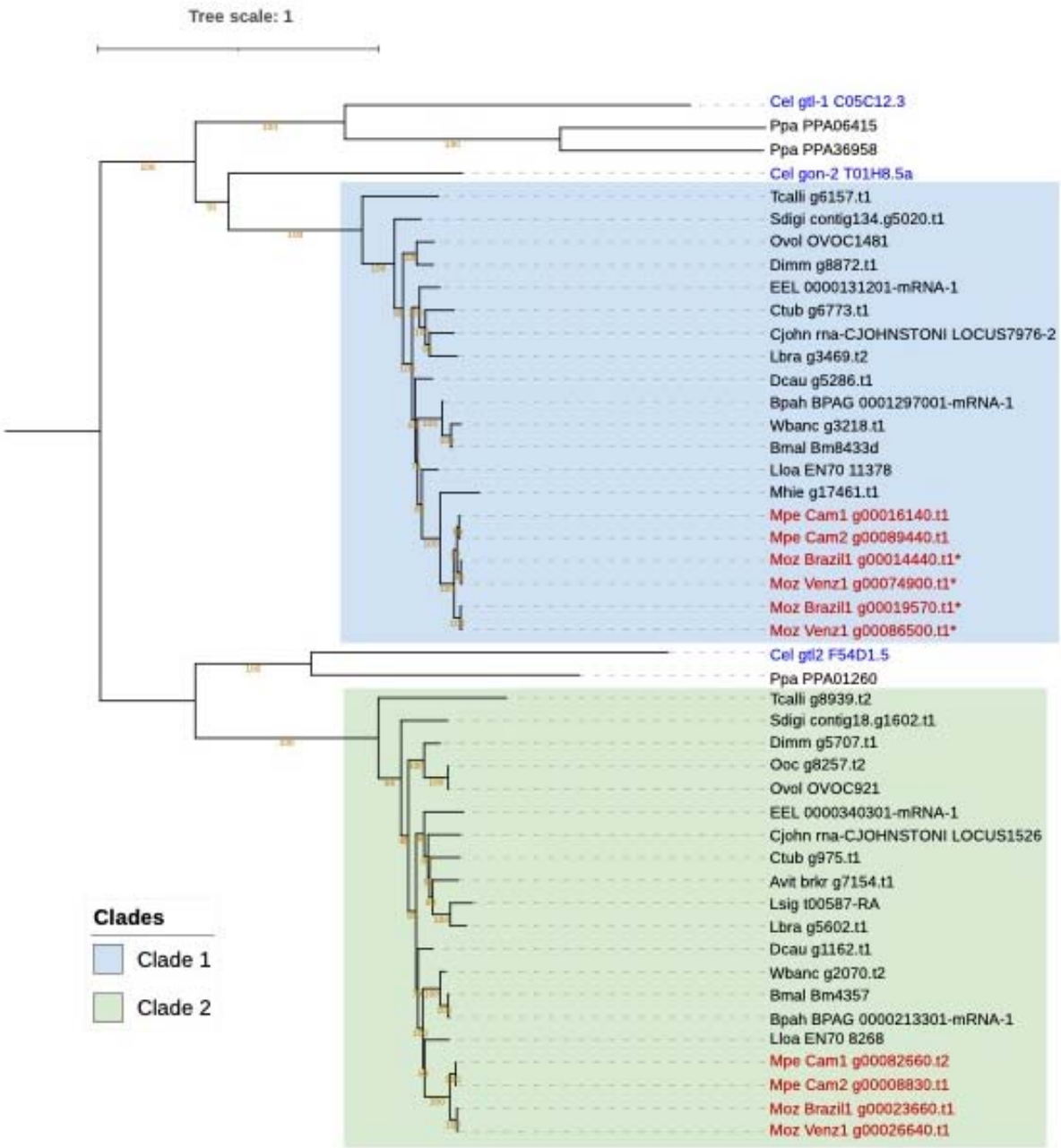
10.3 Figure 3. Phylogenomic analysis of filarial parasites.

Phylogenomic tree was generated based on protein sequences of 465 single copy orthologs identified across all nematode genomes. Onchocercid specie *S. digitata* and *T. callipaeda* were used as immediate outgroups, while free-living nematodes *C. elegans* and *P. pacificus* were used as more distant outgroups. The previously described Onchocercid clades ONC2 to ONC5 are highlighted.



10.4 Figure 4. Phylogenetic analysis of genes encoding the targets of the anti-filarial drug Diethylcarbamazine (DEC).

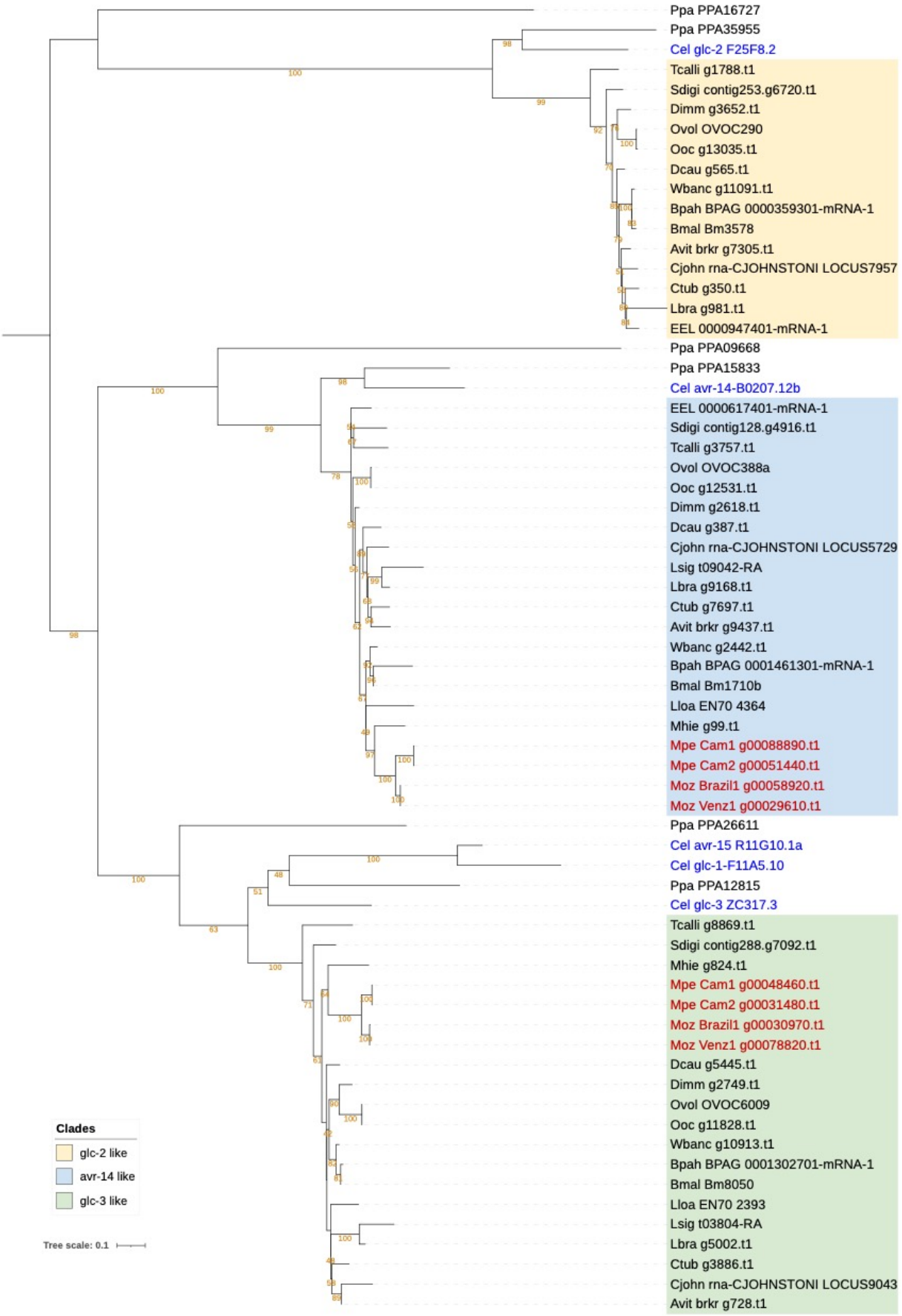
Phylogenetic analysis of the *gon-2*, *gtl-1* and *gtl-2* gene family was performed based on their protein sequences and the tree was rooted at the midpoint. The orthologs of the *Cel-gon-2* gene are placed in Clade 1, highlighted in blue, and orthologs of *Cel-gtl-2* orthologs are placed in Clade 2, highlighted green. The *gtl-1* gene has no orthologs in filarial nematodes. Genes from *M. perstans* and *M. ozzardi* are in red, and *C. elegans* genes are in blue. Asterisks (*) next to gene names indicate fragmented gene structures.



10.5 Figure 5. Phylogenetic analysis of genes encoding Glutamate Chloride channel subunits, targets of ivermectin.

Phylogenetic analysis of the GluCl gene family was based on their protein sequences and the tree was rooted at the midpoint. The clade with *Cel-glc-2* orthologs is highlighted in yellow, clade with *Cel-*

avr-14 orthologs is highlighted in blue, and the clade with *Cel-glc-3* orthologs is highlighted in green. *M. perstans* and *M. ozzardi* genes are in red, and *C. elegans* genes are in blue.


Ppa PPA16727
Ppa PPA35955
Cel glc-2 F25F8.2
Tcalli g1788.t1
Sdigi contig253.g6720.t1
Dimm g3652.t1
Ovol OVOC290
Ooc g13035.t1
Dcau g565.t1
Wbanc g11091.t1
Bpah BPAG 0000359301-mRNA-1
Bmal Bm3578
Avit brkr g7305.t1
Cjohn rna-CJOHNSTONI LOCUS7957
Ctub g350.t1
Lbra g981.t1
EEL 0000947401-mRNA-1
Ppa PPA09668
Ppa PPA15833
Cel avr-14-B0207.12b
EEL 0000617401-mRNA-1
Sdigi contig128.g4916.t1
Tcalli g3757.t1
Ovol OVOC388a
Ooc g12531.t1
Dimm g2618.t1
Dcau g387.t1
Cjohn rna-CJOHNSTONI LOCUS5729
Lsig t09042-RA
Lbra g9168.t1
Ctub g7697.t1
Avit brkr g9437.t1
Wbanc g2442.t1
Bpah BPAG 0001461301-mRNA-1
Bmal Bm1710b
Lloa EN70 4364
Mhie g99.t1
Mpe Cam1 g00088890.t1
Mpe Cam2 g00051440.t1
Moz Brazil1 g00058920.t1
Moz Venz1 g00029610.t1
Ppa PPA26611
Cel avr-15 R11G10.1a
Cel glc-1-F11A5.10
Ppa PPA12815
Cel glc-3 ZC317.3
Tcalli g8869.t1
Sdigi contig288.g7092.t1
Mhie g824.t1
Mpe Cam1 g00048460.t1
Mpe Cam2 g00031480.t1
Moz Brazil1 g00030970.t1
Moz Venz1 g00078820.t1
Dcau g5445.t1
Dimm g2749.t1
Ovol OVOC6009
Ooc g11828.t1
Wbanc g10913.t1
Bpah BPAG 0001302701-mRNA-1
Bmal Bm8050
Lloa EN70 2393
Lsig t03804-RA
Lbra g5002.t1
Ctub g3886.t1
Cjohn rna-CJOHNSTONI LOCUS9043
Avit brkr g728.t1
Clades
glc-2 like
avr-14 like
glc-3 like
Tree scale: 0.1

572

573

11 Tables

11.1 Table 1. Summary statistics of *Mansonella* genome assemblies

	M. perstans	*M. perstans*	*M. ozzardi*	*M. ozzardi*
Isolate Name	**Mpe1-Cam-1**	**Mpe-Cam-2**	**Moz-Brazil-1**	**Moz-Venz-1**
Sampling site	Cameroon	Cameroon	Brazil	Venezuela
Sequencing Platform	PacBio	Illumina	Illumina	Illumina
Assembled genome size	76 Mb	74.35 Mb	75.78 Mb	74.86 Mb
Number of scaffolds	581	6,206	4,789	4,561
Largest scaffold	1.25 Mb	217 kb	317 kb	258 kb
Scaffold N50 size	285.8 kb	29kb	37.75 kb	38.90 kb
%GC	30.31	30.06	29.65	29.67
% Repeats	9.70 %	7.77 %	8.57%	7.91 %
Predicted genes	9,477	9,984	9,834	9,860

11.2 Table 2. Structural and local differences between *Mansonella* genomes

	Mpe-Cam-2 versus Mpe-Cam-1	**Moz-Venz-1 versus Moz-Brazil-1**	**Moz-Brazil-1 versus Mpe-Cam-1**
gANI score	**99.51 %**	**99.50 %**	**89.69**

All local variants categorized by type (% of genome size)	**146,854 (0.19 %)**	**220, 466 (0.29 %)**	**4,726,827 (6.22 %)**
Substitutions, Single Nucleotide Variants	92,837 (0.12 %)	169, 489 (0.22 %)	3,575,532 (4.70 %)
Insertions	34,473 (0.05 %)	19,983 (0.03 %)	401,196 (0.53 %)
Deletions	14,419 (0.02 %)	20,018 (0.03 %)	429,726 (0.57 %)
Substitutions, Multiple Nucleotide variants	5,125 (0.01 %)	10,976 (0.01 %)	320,373 (0.42 %)
All local variants categorized by location in the genome	**146,854**	**220,466**	**4,726, 827**
Variants outside genes (% of all variants)	92,028 (62.7 %)	132, 043 (59.9 %)	2,495, 765 (52.8 %)
Variants within introns (% of all variants	47,318 (32.2 %)	63,934 (29.0 %)	1,786, 741 (37.8 %)
Variants within exons (% of all variants)	7,508 (5.1 %)	24,471 (8.4 %)	444,321 (9.4 %)
Total number of synonymous and non-synonymous changes caused by variants	**8,757**	**18,502**	**577,029**
Synonymous changes	4,797 (54.8 %)	11,112 (60%)	330,755 (57.32 %)
Non-Synonymous, Missense changes	3,893 (44.5 %)	7,300 (39.5 %)	245,069 (42.47 %)
Non-Synonymous, Nonsense changes	67 (0.77%)	90 (0.48 %)	1,205 (0.21 %)

Total number of Structural differences	**1,273**	**4,114**	**3,828**
Translocations	1,139	4,060	2,975
Relocations	63	1	292
Reshufflings	29	15	176
Inversions	42	38	385

11.3 Table 3. Nuclear mitochondria Transfers (nuMTs) annotated across all genomes

Species	**Isolate**	**Total number of nuMTs**	**Total size of nuMTs**	**Median size of nuMTs**	**Largest nuMT**	**Median percent identity to mitochondria**
M. perstans	Mpe-Cam1	48	19.240 kb	238 bp	2,591 bp	87.81 %
M. perstans	Mpe-Cam2	39	14.330 kb	249 bp	2,351 bp	86.48 %
M. ozzardi	Moz-Brazil-1	82	35.070 kb	271 bp	3,170 bp	89.57 %
M. ozzardi	Moz-Venz-1	94	35.851 kb	223 bp	3,697 bp	89.50 %

11.4 Table 4. Nuclear Wolbachia Transfers (nuWTs) annotated across all genomes

Species	**Isolate**	**Total number of nuWTs**	**Total size of nuWTs**	**Median size of nuWTS**	**Largest nuWT**	**Median percent identity to *Wolbachia***
M. perstans	Mpe-Cam1	328	87.83 kb	173 bp	2,733 bp	91.51 %

M. perstans	Mpe-Cam2	340	76.26 kb	171 bp	1,236 bp	91.84 %
M. ozzardi	Moz-Brazil-1	308	77.33 kb	198 bp	1,515 bp	91.37 %
M. ozzardi	Moz-Venz-1	294	73.54 kb	195 bp	2,151 bp	90.97 %

12 References

Basano, S. de A., Camargo, J. de S. A. A., Fontes, G., Pereira, A. R., Medeiros, J. F., Laudisse, M. C. de O., et al. (2018). Phase III Clinical Trial to Evaluate Ivermectin in the Reduction of *Mansonella ozzardi* infection in the Brazilian Amazon. *The American Journal of Tropical Medicine and Hygiene* 98, 786–790. doi: 10.4269/ajtmh.17-0698.

Bélard, S., and Gehringer, C. (2021). High Prevalence of *Mansonella* Species and Parasitic Coinfections in Gabon Calls for an End to the Neglect of *Mansonella* Research. *The Journal of Infectious Diseases* 223, 187–188. doi: 10.1093/infdis/jiaa671.

Bensasson, D., Feldman, M. W., and Petrov, D. A. (2003). Rates of DNA Duplication and Mitochondrial DNA Insertion in the Human Genome. *J Mol Evol* 57, 343–354. doi: 10.1007/s00239-003-2485-7.

Bregani, E. R., Rovellini, A., Mbaïdoum, N., and Magnini, M. G. (2006). Comparison of different anthelminthic drug regimens against Mansonella perstans filariasis. *Transactions of The Royal Society of Tropical Medicine and Hygiene* 100, 458–463. doi: 10.1016/j.trstmh.2005.07.009.

Brůna, T., Hoff, K. J., Lomsadze, A., Stanke, M., and Borodovsky, M. (2021). BRAKER2: automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database. *NAR Genomics and Bioinformatics* 3. doi: 10.1093/nargab/lqaa108.

Calvopina, M., Chiluisa-Guacho, C., Toapanta, A., Fonseca, D., and Villacres, I. (2019). High Prevalence of *Mansonella ozzardi* Infection in the Amazon Region, Ecuador. *Emerging Infectious Diseases* 25, 2081–2083. doi: 10.3201/eid2511.181964.

Chadee, D. D., Tilluckdharry, C. C., Rawlins, S. C., Doon, R., and Nathan, M. B. (1995). Mass Chemotherapy with Diethylcarbamazine for the Control of Bancroftian Filariasis: A Twelve-Year Follow-Up in Northern Trinidad, Including Observations on Mansonella ozzardi. *The American Journal of Tropical Medicine and Hygiene* 52, 174–176. doi: 10.4269/ajtmh.1995.52.174.

Chung, M., Aluvathingal, J., Bromley, R. E., Nadendla, S., Fombad, F. F., Kien, C. A., et al. (2020). Complete Mitochondrial Genome Sequence of Mansonella perstans. *Microbiol Resour Announc* 9. doi: 10.1128/MRA.00490-20.

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., et al. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly* 6, 80–92. doi: 10.4161/fly.19695.

Cotton, J. A., Bennuru, S., Grote, A., Harsha, B., Tracey, A., Beech, R., et al. (2016). The genome of *Onchocerca volvulus*, agent of river blindness. *Nature Microbiology* 2, 16216. doi: 10.1038/nmicrobiol.2016.216.

Coulibaly, Y. I., Dembele, B., Diallo, A. A., Lipner, E. M., Doumbia, S. S., Coulibaly, S. Y., et al. (2009). A Randomized Trial of Doxycycline for *Mansonella perstans* Infection. *New England Journal of Medicine* 361, 1448–1458. doi: 10.1056/NEJMoa0900863.

Crainey, J. L., Costa, C. H. A., de Oliveira Leles, L. F., Ribeiro da Silva, T. R., de Aquino Narzetti, L. H., Serra dos Santos, Y. V., et al. (2020). Deep Sequencing Reveals Occult Mansonellosis Coinfections in Residents From the Brazilian Amazon Village of São Gabriel da Cachoeira. *Clinical Infectious Diseases* 71, 1990–1993. doi: 10.1093/cid/ciaa082.

Crainey, J. L., Marín, M. A., Silva, T. R. R. da, Medeiros, J. F. de, Pessoa, F. A. C., Santos, Y. V., et al. (2018). *Mansonella ozzardi* mitogenome and pseudogene characterisation provides new perspectives on filarial parasite systematics and CO-1 barcoding. *Scientific Reports* 8, 6158. doi: 10.1038/s41598-018-24382-3.

Debrah, L. B., Phillips, R. O., Pfarr, K., Klarmann-Schulz, U., Opoku, V. S., Nausch, N., et al. (2019). The Efficacy of Doxycycline Treatment on *Mansonella perstans* Infection: An Open-Label, Randomized Trial in Ghana. *The American Journal of Tropical Medicine and Hygiene* 101, 84–92. doi: 10.4269/ajtmh.18-0491.

Dent, J. A., Smith, M. M., Vassilatis, D. K., and Avery, L. (2000). The genetics of ivermectin resistance in Caenorhabditis elegans. *Proc Natl Acad Sci U S A* 97, 2674–2679.

Desjardins, C. A., Cerqueira, G. C., Goldberg, J. M., Hotopp, J. C. D., Haas, B. J., Zucker, J., et al. (2013). Genomics of Loa loa, a Wolbachia-free filarial parasite of humans. *Nat. Genet.* doi: 10.1038/ng.2585.

Downes, B., and Jacobsen, K. (2010). A systematic review of the epidemiology of mansonelliasis. *African Journal of Infectious Diseases* 4. doi: 10.4314/ajid.v4i1.55085.

Driscoll, M., Dean, E., Reilly, E., Bergholz, E., and Chalfie, M. (1989). Genetic and molecular analysis of a Caenorhabditis elegans beta-tubulin that conveys benzimidazole sensitivity. *Journal of Cell Biology* 109, 2993–3003. doi: 10.1083/jcb.109.6.2993.

Emms, D. M., and Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biology* 20, 238. doi: 10.1186/s13059-019-1832-y.

Ferreira, M. U., Crainey, J. L., and Luz, S. L. B. (2021). *Mansonella ozzardi*. *Trends in Parasitology* 37, 90–91. doi: 10.1016/j.pt.2020.03.005.

Flynn, J. M., Hubley, R., Goubert, C., Rosen, J., Clark, A. G., Feschotte, C., et al. (2020). RepeatModeler2 for automated genomic discovery of transposable element families. *PNAS* 117, 9451–9457. doi: 10.1073/pnas.1921046117.

Foster, J. M., Grote, A., Mattick, J., Tracey, A., Tsai, Y.-C., Chung, M., et al. (2020). Sex chromosome evolution in parasitic nematodes of humans. *Nature Communications* 11, 1964. doi: 10.1038/s41467-020-15654-6.

Ghedin, E., Wang, S., Spiro, D., Caler, E., Zhao, Q., Crabtree, J., et al. (2007). Draft Genome of the Filarial Nematode Parasite Brugia malayi. *Science* 317, 1756–1760. doi: 10.1126/science.1145406.

Godel, C., Kumar, S., Koutsovoulos, G., Ludin, P., Nilsson, D., Comandatore, F., et al. (2012). The genome of the heartworm, Dirofilaria immitis, reveals drug and vaccine targets. *FASEB J* 26, 4650–4661. doi: 10.1096/fj.12-205096.

Grote, A., Lustigman, S., and Ghedin, E. (2017). Lessons from the genomes and transcriptomes of filarial nematodes. *Molecular and Biochemical Parasitology* 215, 23–29. doi: 10.1016/j.molbiopara.2017.01.004.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Systematic Biology* 59, 307–321. doi: 10.1093/sysbio/syq010.

Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29, 1072–1075. doi: 10.1093/bioinformatics/btt086.

Hazkani-Covo, E., Zeller, R. M., and Martin, W. (2010). Molecular Poltergeists: Mitochondrial DNA Copies (numts) in Sequenced Nuclear Genomes. *PLOS Genetics* 6, e1000834. doi: 10.1371/journal.pgen.1000834.

Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., and Vinh, L. S. (2018). UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution* 35, 518–522. doi: 10.1093/molbev/msx281.

Hotopp, J. C. D., Slatko, B. E., and Foster, J. M. (2017). Targeted Enrichment and Sequencing of Recent Endosymbiont-Host Lateral Gene Transfers. *Sci Rep* 7, 1–10. doi: 10.1038/s41598-017-00814-4.

Howe, K. L., Bolt, B. J., Shafie, M., Kersey, P., and Berriman, M. (2017). WormBase ParaSite − a comprehensive resource for helminth genomics. *Molecular and Biochemical Parasitology* 215, 2–10. doi: 10.1016/j.molbiopara.2016.11.005.

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., and Jermiin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* 14, 587–589. doi: 10.1038/nmeth.4285.

Khelik, K., Lagesen, K., Sandve, G. K., Rognes, T., and Nederbragt, A. J. (2017). NucDiff: in-depth characterization and annotation of differences between two sets of DNA sequences. *BMC Bioinformatics* 18, 338. doi: 10.1186/s12859-017-1748-z.

Klasson, L., Kumar, N., Bromley, R., Sieber, K., Flowers, M., Ott, S. H., et al. (2014). Extensive duplication of the Wolbachia DNA in chromosome four of Drosophila ananassae. *BMC Genomics* 15, 1097. doi: 10.1186/1471-2164-15-1097.

Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., and Phillippy, A. M. (2017). Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res.* 27, 722–736. doi: 10.1101/gr.215087.116.

Kozek, W. J., Palma, G., Henao, A., García, H., and Hoyos, M. (1983). Filariasis in Colombia: Prevalence and Distribution of *Mansonella Ozzardi* and *Mansonella (=Dipetalonema) Perstans* Infections in the Comisaría Del Guainía. *The American Journal of Tropical Medicine and Hygiene* 32, 379–384. doi: 10.4269/ajtmh.1983.32.379.

Kriventseva, E. V., Tegenfeldt, F., Petty, T. J., Waterhouse, R. M., Simão, F. A., Pozdnyakov, I. A., et al. (2015). OrthoDB v8: update of the hierarchical catalog of orthologs and the underlying free software. *Nucl. Acids Res.* 43, D250–D256. doi: 10.1093/nar/gku1220.

Kwarteng, A., Sylverken, A., Asiedu, E., and Ahuno, S. T. (2021). Genome editing as control tool for filarial infections. *Biomedicine & Pharmacotherapy* 137, 111292. doi: 10.1016/j.biopha.2021.111292.

Lacey, E. (1990). Mode of action of benzimidazoles. *Parasitology Today* 6, 112–115. doi: 10.1016/0169-4758(90)90227-U.

Laetsch, D. R., and Blaxter, M. L. (2017). BlobTools: Interrogation of genome assemblies. *F1000Research* 6, 1287. doi: 10.12688/f1000research.12232.1.

Lam, K.-K., LaButti, K., Khalak, A., and Tse, D. (2015). FinisherSC: a repeat-aware tool for upgrading de novo assembly using long reads. *Bioinformatics* 31, 3207–3209. doi: 10.1093/bioinformatics/btv280.

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods* 9, 357–359. doi: 10.1038/nmeth.1923.

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

Lefoulon, E., Bain, O., Makepeace, B. L., d'Haese, C., Uni, S., Martin, C., et al. (2016). Breakdown of coevolution between symbiotic bacteria *Wolbachia* and their filarial hosts. *PeerJ* 4, e1840. doi: 10.7717/peerj.1840.

Lefoulon, E., Clark, T., Guerrero, R., Cañizales, I., Cardenas-Callirgos, J. M., Junker, K., et al. (2020). Diminutive, degraded but dissimilar: *Wolbachia* genomes from filarial nematodes do not conform to a single paradigm. *Microbial Genomics* 6, e000487. doi: 10.1099/mgen.0.000487.

Letunic, I., and Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Research* 49, W293–W296. doi: 10.1093/nar/gkab301.

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34, 3094–3100. doi: 10.1093/bioinformatics/bty191.

Lima, N. F., Aybar, C. A. V., Juri, M. J. D., and Ferreira, M. U. (2016). *Mansonella ozzardi*: a neglected New World filarial nematode. *Pathogens and Global Health* 110, 97–107. doi: 10.1080/20477724.2016.1190544.

Lima, N. F., Gonçalves-Lopes, R. M., Kruize, Y. C. M., Yazdanbakhsh, M., and Ferreira, M. U. (2018). CD39 and immune regulation in a chronic helminth infection: The puzzling case of *Mansonella ozzardi*. *PLOS Neglected Tropical Diseases* 12, e0006327. doi: 10.1371/journal.pntd.0006327.

Liu, C., Grote, A., Ghedin, E., and Unnasch, T. R. (2020). CRISPR-mediated Transfection of Brugia malayi. *PLOS Neglected Tropical Diseases* 14, e0008627. doi: 10.1371/journal.pntd.0008627.

Lok, J. B., Shao, H., Massey, H. C., and Li, X. (2016). Transgenesis in Strongyloides and related parasitic nematodes: historical perspectives, current functional genomic applications and progress towards gene disruption and editing. *Parasitology*, 1–16. doi: 10.1017/S0031182016000391.

Lu, S., Wang, J., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R., et al. (2020). CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Research* 48, D265–D268. doi: 10.1093/nar/gkz991.

Marçais, G., Delcher, A. L., Phillippy, A. M., Coston, R., Salzberg, S. L., and Zimin, A. (2018). MUMmer4: A fast and versatile genome alignment system. *PLOS Computational Biology* 14, e1005944. doi: 10.1371/journal.pcbi.1005944.

McCann, K., Grant, W., and Doyle, S. R. (2021). The genome sequence of the Australian filarial nematode, Cercopithifilaria johnstoni. *Wellcome Open Res* 6, 259. doi: 10.12688/wellcomeopenres.17258.2.

McNulty, S. N., Foster, J. M., Mitreva, M., Hotopp, J. C. D., Martin, J., Fischer, K., et al. (2010). Endosymbiont DNA in Endobacteria-Free Filarial Nematodes Indicates Ancient Horizontal Genetic Transfer. *PLOS ONE* 5, e11029. doi: 10.1371/journal.pone.0011029.

McNulty, S. N., Mitreva, M., Weil, G. J., and Fischer, P. U. (2013). Inter and intra-specific diversity of parasites that cause lymphatic filariasis. *Infection, Genetics and Evolution* 14, 137–146. doi: 10.1016/j.meegid.2012.11.002.

Mediannikov, O., and Ranque, S. (2018). Mansonellosis, the most neglected human filariasis. *New Microbes New Infect* 26, S19–S22. doi: 10.1016/j.nmni.2018.08.016.

Nakamura, T., Yamada, K. D., Tomii, K., and Katoh, K. (2018). Parallelization of MAFFT for large-scale multiple sequence alignments. *Bioinformatics* 34, 2490–2492. doi: 10.1093/bioinformatics/bty121.

Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Molecular Biology and Evolution* 32, 268–274. doi: 10.1093/molbev/msu300.

Njouendou, A. J., Ritter, M., Ndongmo, W. P. C., Kien, C. A., Narcisse, G. T. V., Fombad, F. F., et al. (2017). Successful long-term maintenance of Mansonella perstans in an in vitro culture system. *Parasites & Vectors* 10, 563. doi: 10.1186/s13071-017-2515-8.

Nurk, S., Meleshko, D., Korobeynikov, A., and Pevzner, P. A. (2017). metaSPAdes: a new versatile metagenomic assembler. *Genome Res.* 27, 824–834. doi: 10.1101/gr.213959.116.

Pallotto, L. M., Dilks, C. M., Park, Y.-J., Smit, R. B., Lu, B. T., Gopalakrishnan, C., et al. (2022). Interactions of Caenorhabditis elegans β-tubulins with the microtubule inhibitor and anthelmintic drug albendazole. *Genetics* 221, iyac093. doi: 10.1093/genetics/iyac093.

Poole, C. B., Sinha, A., Ettwiller, L., Apone, L., McKay, K., Panchapakesa, V., et al. (2019). In Silico Identification of Novel Biomarkers and Development of New Rapid Diagnostic Tests for the Filarial Parasites *Mansonella perstans* and *Mansonella ozzardi*. *Scientific Reports* 9, 10275. doi: 10.1038/s41598-019-46550-9.

Pryszcz, L. P., and Gabaldón, T. (2016). Redundans: an assembly pipeline for highly heterozygous genomes. *Nucleic Acids Res* 44, e113–e113. doi: 10.1093/nar/gkw294.

Qing, X., Kulkeaw, K., Wongkamchai, S., and Tsui, S. K.-W. (2021). Mitochondrial Genome of Brugia malayi Microfilariae Isolated From a Clinical Sample. *Frontiers in Ecology and Evolution* 9. Available at: https://www.frontiersin.org/article/10.3389/fevo.2021.637805 [Accessed May 17, 2022].

Quinlan, A. R., and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. doi: 10.1093/bioinformatics/btq033.

Quinzo, M. J., Perteguer, M. J., Brindley, P. J., Loukas, A., and Sotillo, J. (2022). Transgenesis in parasitic helminths: a brief history and prospects for the future. *Parasites & Vectors* 15, 110. doi: 10.1186/s13071-022-05211-z.

Raccurt, C. P. (2018). *Mansonella ozzardi* and its vectors in the New World: an update with emphasis on the current situation in Haiti. *Journal of Helminthology* 92, 655–661. doi: 10.1017/S0022149X17000955.

Ritter, M., Ndongmo, W. P. C., Njouendou, A. J., Nghochuzie, N. N., Nchang, L. C., Tayong, D. B., et al. (2018). *Mansonella perstans* microfilaremic individuals are characterized by enhanced type 2 helper T and regulatory T and B cell subsets and dampened systemic innate and adaptive immune responses. *PLOS Neglected Tropical Diseases* 12, e0006184. doi: 10.1371/journal.pntd.0006184.

Sandri, T. L., Kreidenweiss, A., Cavallo, S., Weber, D., Juhas, S., Rodi, M., et al. (2021). Molecular Epidemiology of *Mansonella* Species in Gabon. *The Journal of Infectious Diseases* 223, 287–296. doi: 10.1093/infdis/jiaa670.

Sankaranarayanan, G., Berriman, M., and Rinaldi, G. (2021). An uneven race: genome editing for parasitic worms. *Nat Rev Microbiol* 19, 621–621. doi: 10.1038/s41579-021-00625-5.

Saunders, G. I., Wasmuth, J. D., Beech, R., Laing, R., Hunt, M., Naghra, H., et al. (2013). Characterization and comparative analysis of the complete Haemonchus contortus β-tubulin gene

family and implications for benzimidazole resistance in strongylid nematodes. *International Journal for Parasitology* 43, 465–475. doi: 10.1016/j.ijpara.2012.12.011.

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31, 3210–3212. doi: 10.1093/bioinformatics/btv351.

Simonsen, P. E., Onapa, A. W., and Asio, S. M. (2011). *Mansonella perstans* filariasis in Africa. *Acta Tropica* 120, S109–S120. doi: 10.1016/j.actatropica.2010.01.014.

Sinha, A., Li, Z., Poole, C. B., Ettwiller, L., Lima, N. F., Ferreira, M. U., et al. (2021). Genomes of Wolbachia endosymbionts from the human filarial parasites Mansonella perstans and Mansonella ozzardi. *bioRxiv*, 2021.03.23.436630. doi: 10.1101/2021.03.23.436630.

Sinha, A., Li, Z., Poole, C. B., Ettwiller, L., Lima, N. F., Ferreira, M. U., et al. (2022). Multiple origins of nematode-Wolbachia symbiosis in supergroup F and convergent loss of bacterioferritin in filarial Wolbachia. 2021.03.23.436630. doi: 10.1101/2021.03.23.436630.

Song, H., Buhay, J. E., Whiting, M. F., and Crandall, K. A. (2008). Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. *Proceedings of the National Academy of Sciences* 105, 13486–13491. doi: 10.1073/pnas.0803076105.

Storer, J., Hubley, R., Rosen, J., Wheeler, T. J., and Smit, A. F. (2021). The Dfam community resource of transposable element families, sequence models, and genome annotations. *Mobile DNA* 12, 2. doi: 10.1186/s13100-020-00230-y.

Talavera, G., and Castresana, J. (2007). Improvement of Phylogenies after Removing Divergent and Ambiguously Aligned Blocks from Protein Sequence Alignments. *Systematic Biology* 56, 564–577. doi: 10.1080/10635150701472164.

Tallon, L. J., Liu, X., Bennuru, S., Chibucos, M. C., Godinez, A., Ott, S., et al. (2014). Single molecule sequencing and genome assembly of a clinical specimen of Loa loa, the causative agent of loiasis. *BMC Genomics* 15, 788. doi: 10.1186/1471-2164-15-788.

Ta-Tang, T.-H., Crainey, J. L., Post, R. J., Luz, S. L., and Rubio, J. M. (2018). Mansonellosis: current perspectives. *Res Rep Trop Med* 9, 9–24. doi: 10.2147/RRTM.S125750.

Ta-Tang, T.-H., Luz, S. L., Crainey, J. L., and Rubio, J. M. (2021). An Overview of the Management of Mansonellosis. *RRTM* 12, 93–105. doi: 10.2147/RRTM.S274684.

Tavares da Silva, L. B., Crainey, J. L., Ribeiro da Silva, T. R., Suwa, U. F., Vicente, A. C. P., Fernandes de Medeiros, J., et al. (2017). Molecular Verification of New World *Mansonella perstans* Parasitemias. *Emerging Infect. Dis.* 23, 545–547. doi: 10.3201/eid2303.161159.

Tvedte, E. S., Gasser, M., Zhao, X., Tallon, L. J., Sadzewicz, L., Bromley, R. E., et al. (2022). Accumulation of endosymbiont genomes in an insect autosome followed by endosymbiont replacement. *Current Biology* 32, 2786-2795.e5. doi: 10.1016/j.cub.2022.05.024.

Verma, S., Kashyap, S. S., Robertson, A. P., and Martin, R. J. (2020). Diethylcarbamazine activates TRP channels including TRP-2 in filaria, Brugia malayi. *Commun Biol* 3, 1–11. doi: 10.1038/s42003-020-01128-4.

Vianna, L. M. M., Martins, M., Cohen, M. J., Cohen, J. M., and Belfort, R. (2012). Mansonella ozzardi corneal lesions in the Amazon: a cross-sectional study. *BMJ Open* 2, e001266. doi: 10.1136/bmjopen-2012-001266.

Wolstenholme, A. J., and Rogers, A. T. (2005). Glutamate-gated chloride channels and the mode of action of the avermectin/milbemycin anthelmintics. *Parasitology* 131, S85–S95. doi: 10.1017/S0031182005008218.

Yoon, S.-H., Ha, S., Lim, J., Kwon, S., and Chun, J. (2017). A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie van Leeuwenhoek* 110, 1281–1286. doi: 10.1007/s10482-017-0844-4.

Zamanian, M., and Andersen, E. C. (2016). Prospects and challenges of CRISPR/Cas genome editing for the study and control of neglected vector-borne nematode diseases. *The FEBS Journal* 283, 3204–3221. doi: 10.1111/febs.13781.

Zimin, A. V., and Salzberg, S. L. (2020). The genome polishing tool POLCA makes fast and accurate corrections in genome assemblies. *PLOS Computational Biology* 16, e1007981. doi: 10.1371/journal.pcbi.1007981.

Figure 1

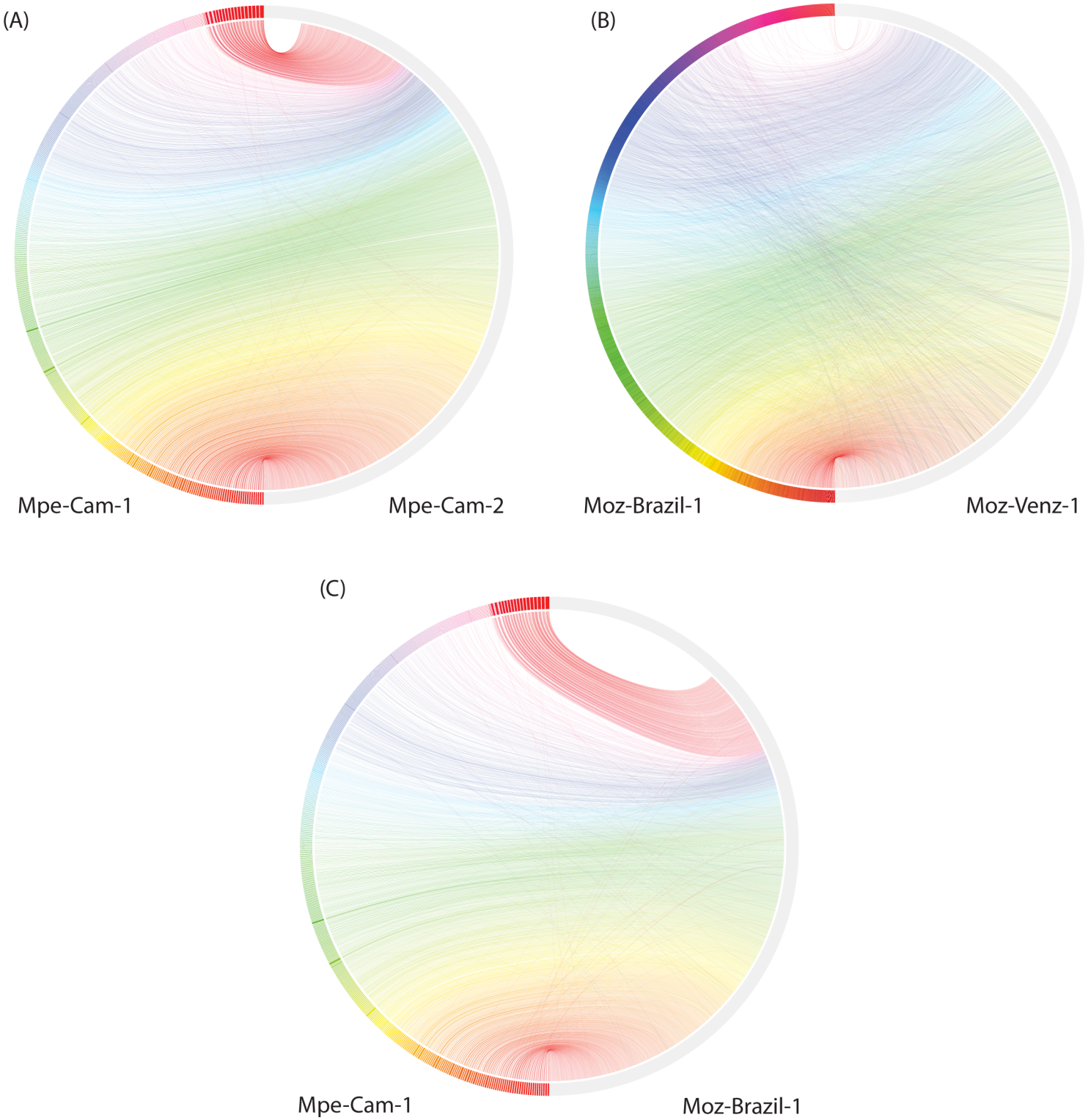

(A)
Mpe-Cam-1
Mpe-Cam-2
(B)
Moz-Brazil-1
Moz-Venz-1
(C)
Mpe-Cam-1
Moz-Brazil-1

Figure 2

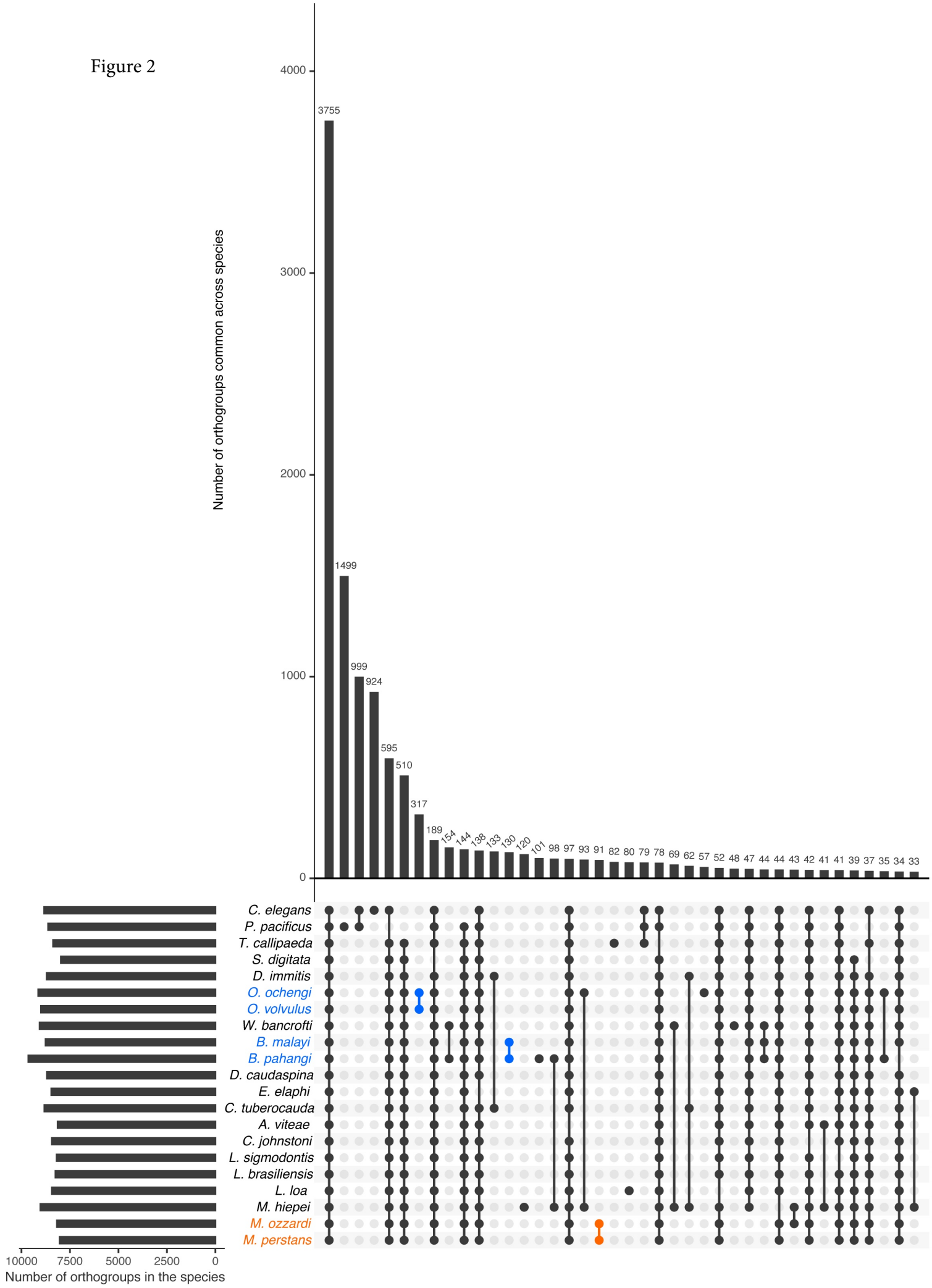

Number of orthogroups common across species
4000
3000
2000
1000
0
3755
1499
999
924
595
510
317
189
154
144
138
133
130
120
101
98
97
93
91
82
80
79
78
69
62
57
52
48
47
44
44
43
42
41
41
39
37
35
34
33
C. elegans
P. pacificus
T. callipaeda
S. digitata
D. immitis
O. ochengi
O. volvulus
W. bancrofti
B. malayi
B. pahangi
D. caudaspina
E. elaphi
C. tuberocauda
A. viteae
C. johnstoni
L. sigmodontis
L. brasiliensis
L. loa
M. hiepei
M. ozzardi
M. perstans
10000
7500
5000
2500
0
Number of orthogroups in the species

Figure 3

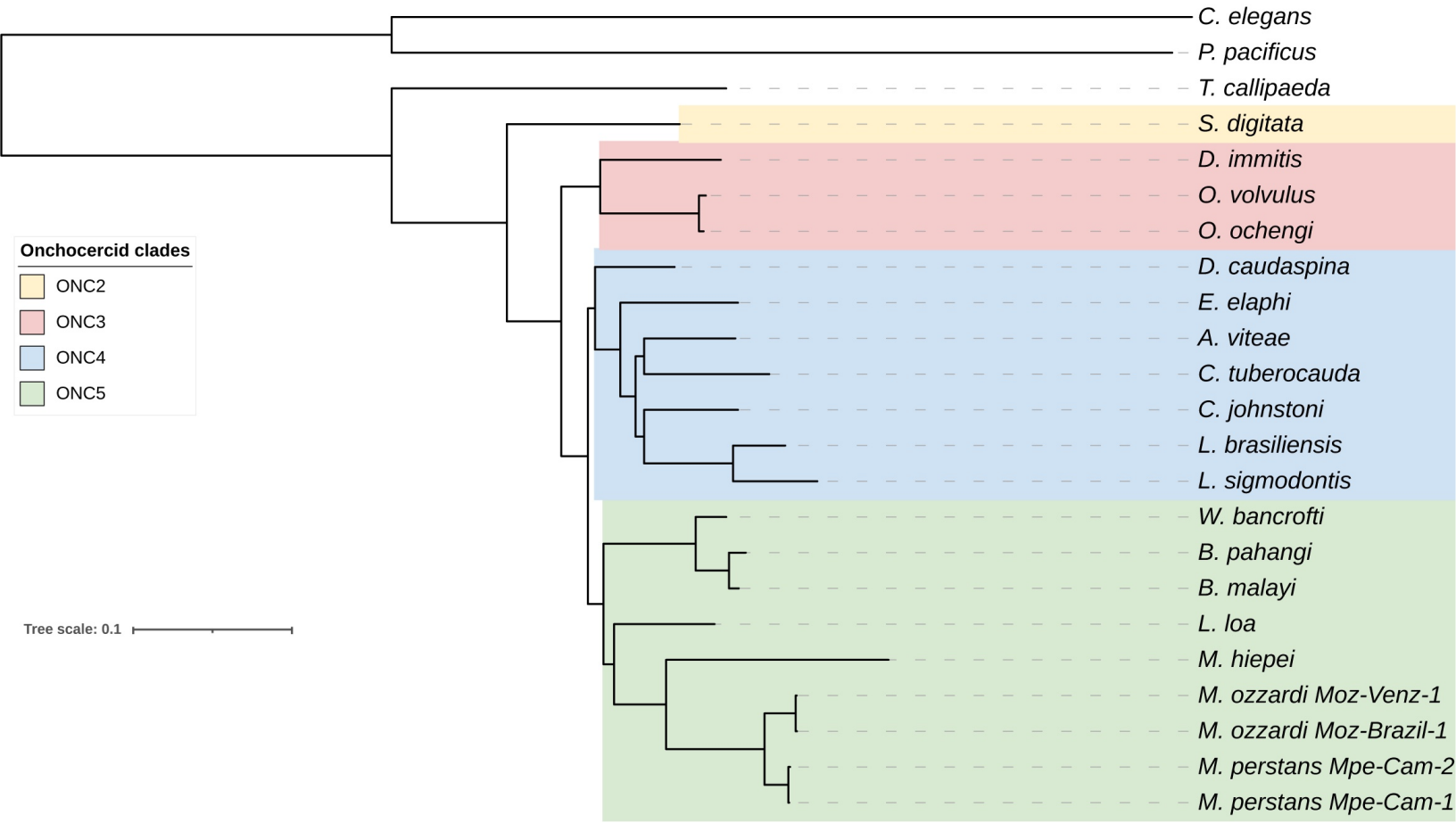

C. elegans
P. pacificus
T. callipaeda
S. digitata
D. immitis
O. volvulus
O. ochengi
D. caudaspina
E. elaphi
A. viteae
C. tuberocauda
C. johnstoni
L. brasiliensis
L. sigmodontis
W. bancrofti
B. pahangi
B. malayi
L. loa
M. hiepei
M. ozzardi Moz-Venz-1
M. ozzardi Moz-Brazil-1
M. perstans Mpe-Cam-2
M. perstans Mpe-Cam-1
Onchocercid clades
ONC2
ONC3
ONC4
ONC5
Tree scale: 0.1

Figure 4

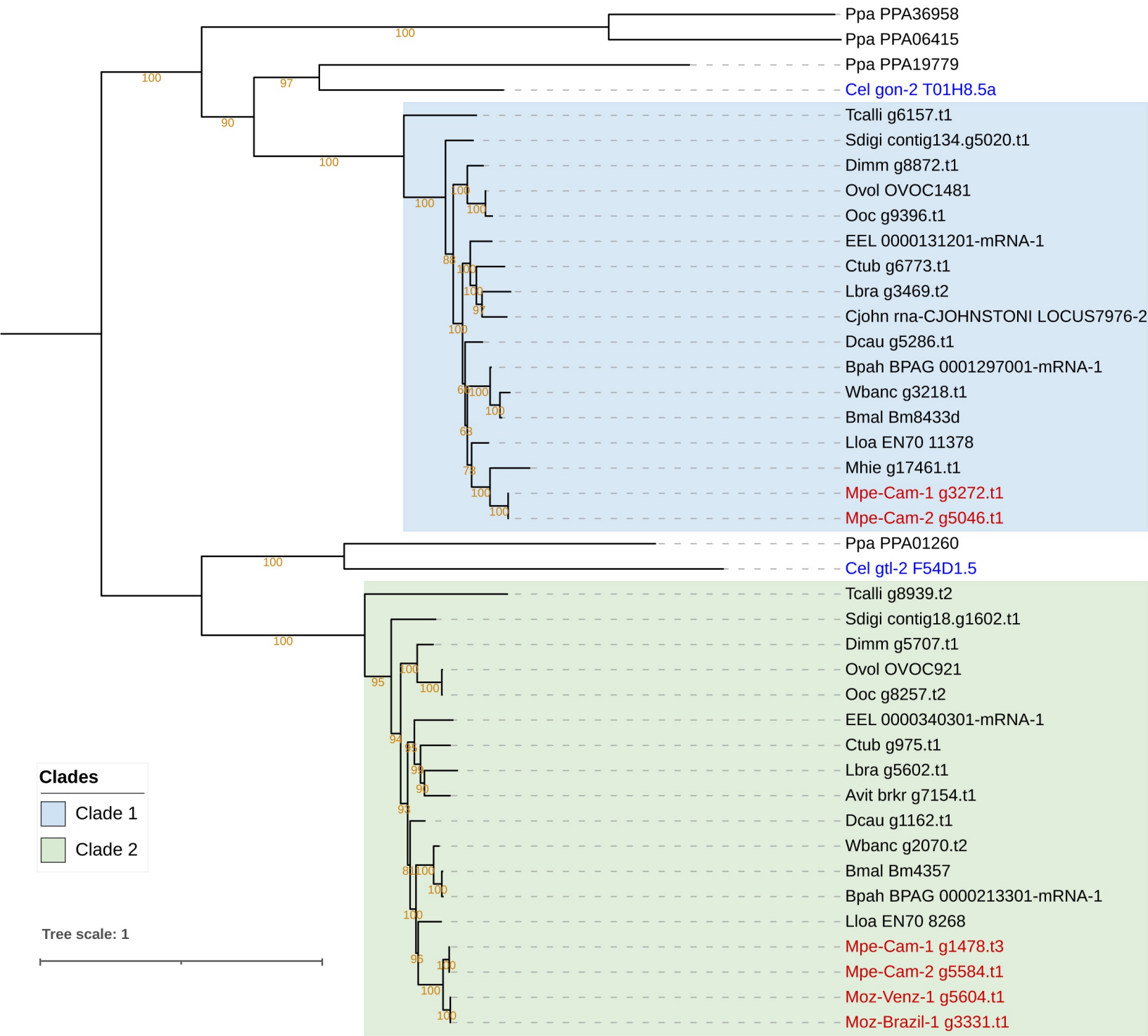
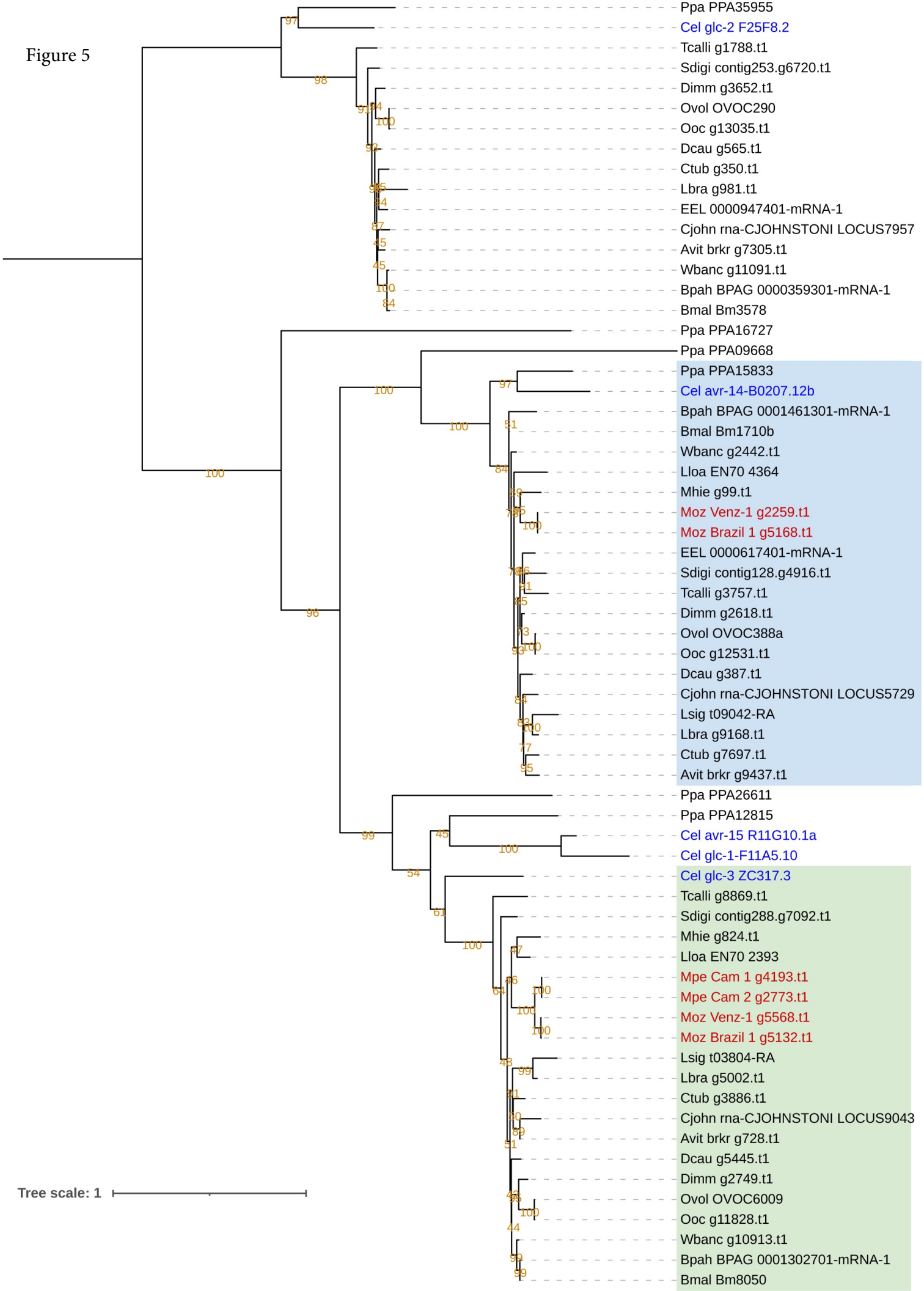
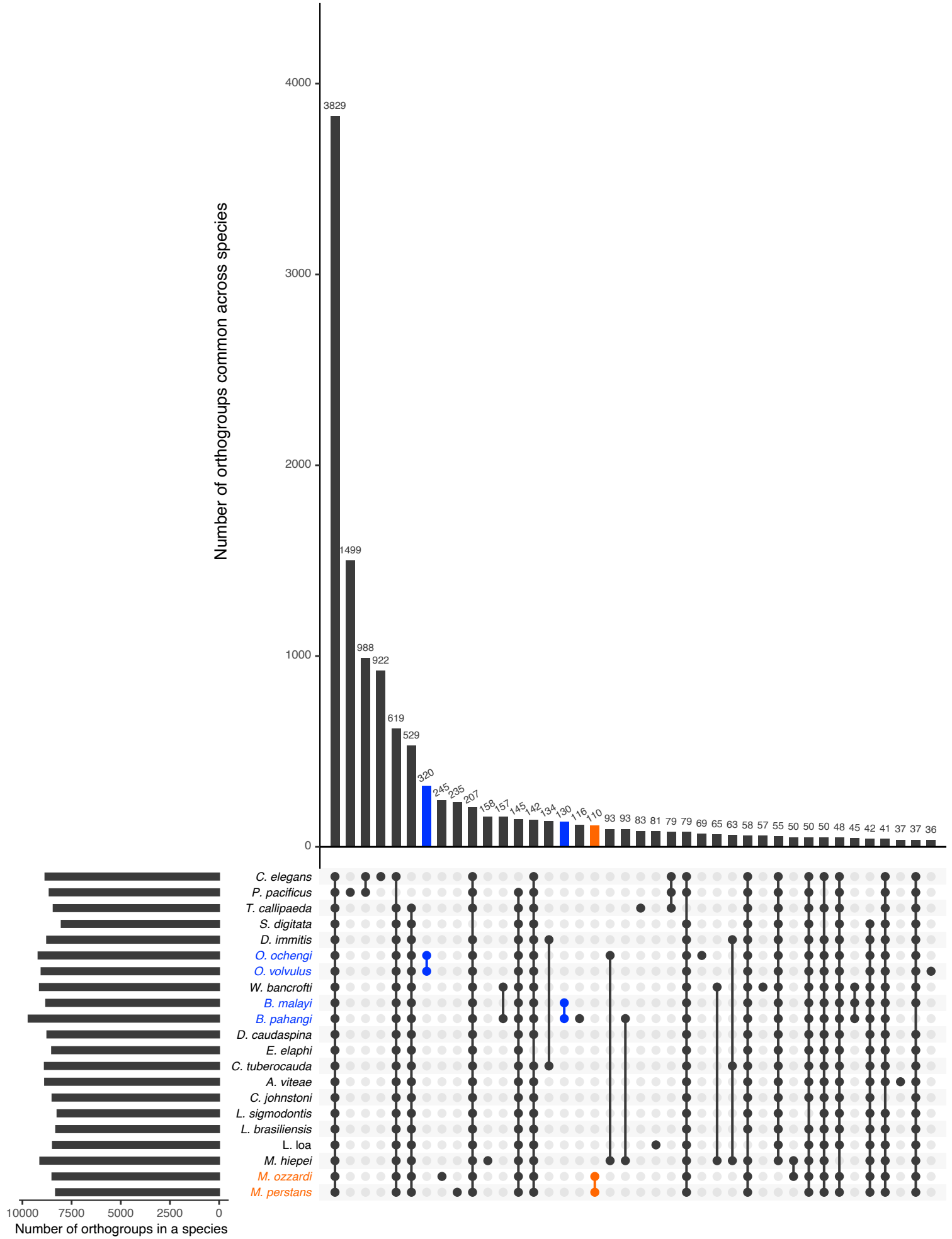
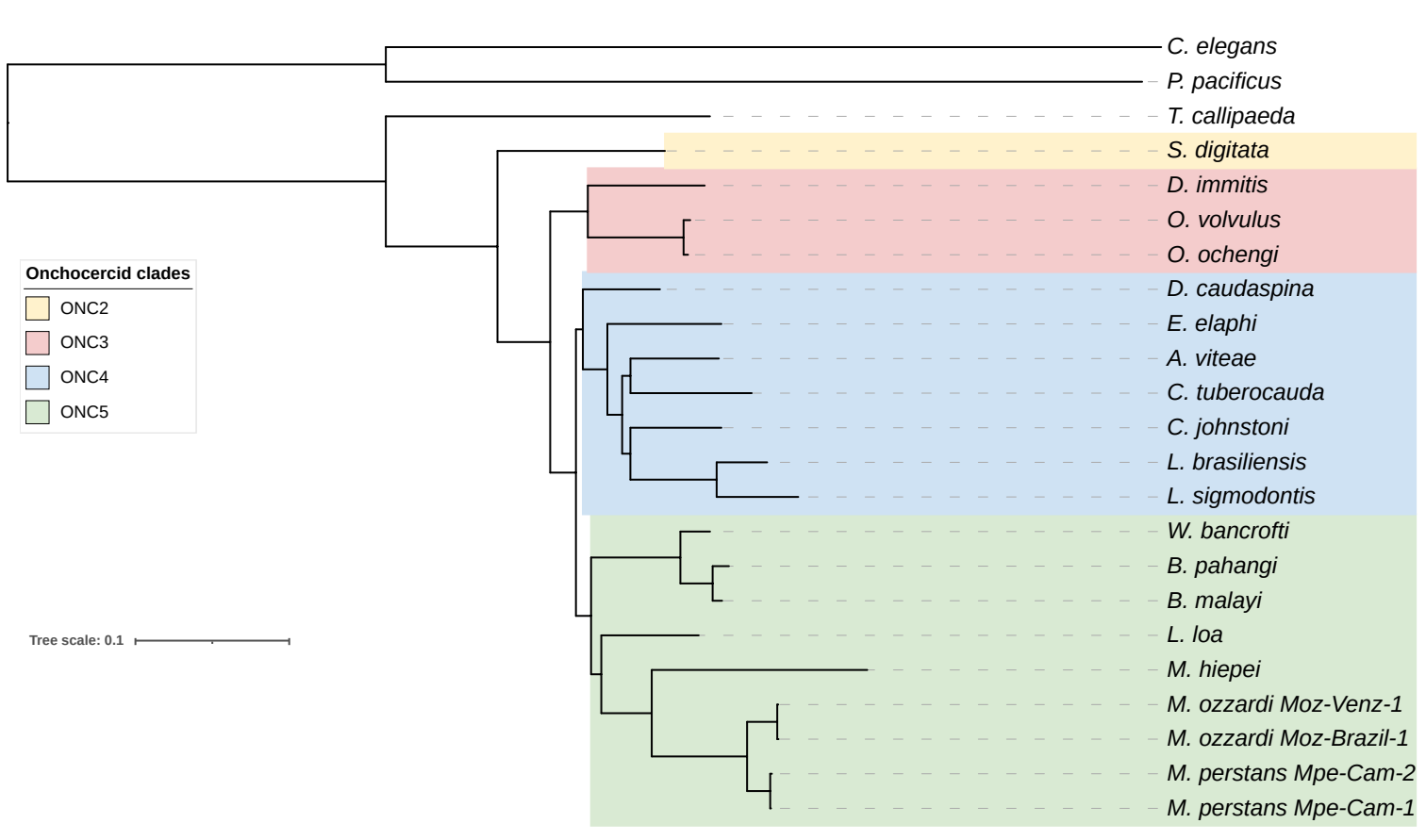

Ppa PPA36958
Ppa PPA06415
Ppa PPA19779
Cel gon-2 T01H8.5a
Tcalli g6157.t1
Sdigi contig134.g5020.t1
Dimm g8872.t1
Ovol OVOC1481
Ooc g9396.t1
EEL 0000131201-mRNA-1
Ctub g6773.t1
Lbra g3469.t2
Cjohn rna-CJOHNSTONI LOCUS7976-2
Dcau g5286.t1
Bpah BPAG 0001297001-mRNA-1
Wbanc g3218.t1
Bmal Bm8433d
Lloa EN70 11378
Mhie g17461.t1
Mpe-Cam-1 g3272.t1
Mpe-Cam-2 g5046.t1
Ppa PPA01260
Cel gtl-2 F54D1.5
Tcalli g8939.t2
Sdigi contig18.g1602.t1
Dimm g5707.t1
Ovol OVOC921
Ooc g8257.t2
EEL 0000340301-mRNA-1
Ctub g975.t1
Lbra g5602.t1
Avit brkr g7154.t1
Dcau g1162.t1
Wbanc g2070.t2
Bmal Bm4357
Bpah BPAG 0000213301-mRNA-1
Lloa EN70 8268
Mpe-Cam-1 g1478.t3
Mpe-Cam-2 g5584.t1
Moz-Venz-1 g5604.t1
Moz-Brazil-1 g3331.t1
Clades
Clade 1
Clade 2
Tree scale: 1

Figure 5


Ppa PPA35955
Cel glc-2 F25F8.2
Tcalli g1788.t1
Sdigi contig253.g6720.t1
Dimm g3652.t1
Ovol OVOC290
Ooc g13035.t1
Dcau g565.t1
Ctub g350.t1
Lbra g981.t1
EEL 0000947401-mRNA-1
Cjohn rna-CJOHNSTONI LOCUS7957
Avit brkr g7305.t1
Wbanc g11091.t1
Bpah BPAG 0000359301-mRNA-1
Bmal Bm3578
Ppa PPA16727
Ppa PPA09668
Ppa PPA15833
Cel avr-14-B0207.12b
Bpah BPAG 0001461301-mRNA-1
Bmal Bm1710b
Wbanc g2442.t1
Lloa EN70 4364
Mhie g99.t1
Moz Venz-1 g2259.t1
Moz Brazil 1 g5168.t1
EEL 0000617401-mRNA-1
Sdigi contig128.g4916.t1
Tcalli g3757.t1
Dimm g2618.t1
Ovol OVOC388a
Ooc g12531.t1
Dcau g387.t1
Cjohn rna-CJOHNSTONI LOCUS5729
Lsig t09042-RA
Lbra g9168.t1
Ctub g7697.t1
Avit brkr g9437.t1
Ppa PPA26611
Ppa PPA12815
Cel avr-15 R11G10.1a
Cel glc-1-F11A5.10
Cel glc-3 ZC317.3
Tcalli g8869.t1
Sdigi contig288.g7092.t1
Mhie g824.t1
Lloa EN70 2393
Mpe Cam 1 g4193.t1
Mpe Cam 2 g2773.t1
Moz Venz-1 g5568.t1
Moz Brazil 1 g5132.t1
Lsig t03804-RA
Lbra g5002.t1
Ctub g3886.t1
Cjohn rna-CJOHNSTONI LOCUS9043
Avit brkr g728.t1
Dcau g5445.t1
Dimm g2749.t1
Ovol OVOC6009
Ooc g11828.t1
Wbanc g10913.t1
Bpah BPAG 0001302701-mRNA-1
Bmal Bm8050
Tree scale: 1


Number of orthogroups common across species
4000
3000
2000
1000
0
3829
1499
988
922
619
529
320
245
235
207
158
157
145
142
134
130
116
110
93
93
83
81
79
79
69
65
63
58
57
55
50
50
50
48
45
42
41
37
37
36
C. elegans
P. pacificus
T. callipaeda
S. digitata
D. immitis
O. ochengi
O. volvulus
W. bancrofti
B. malayi
B. pahangi
D. caudaspina
E. elaphi
C. tuberocauda
A. viteae
C. johnstoni
L. sigmodontis
L. brasiliensis
L. loa
M. hiepei
M. ozzardi
M. perstans
10000
7500
5000
2500
0
Number of orthogroups in a species


C. elegans
P. pacificus
T. callipaeda
S. digitata
D. immitis
O. volvulus
O. ochengi
D. caudaspina
E. elaphi
A. viteae
C. tuberocauda
C. johnstoni
L. brasiliensis
L. sigmodontis
W. bancrofti
B. pahangi
B. malayi
L. loa
M. hiepei
M. ozzardi Moz-Venz-1
M. ozzardi Moz-Brazil-1
M. perstans Mpe-Cam-2
M. perstans Mpe-Cam-1
Onchocercid clades
ONC2
ONC3
ONC4
ONC5
Tree scale: 0.1

Tree scale: 1
Cel gtl-1 C05C12.3
Ppa PPA06415
Ppa PPA36958
Cel gon-2 T01H8.5a
Tcalli g6157.t1
Sdigi contig134.g5020.t1
Ovol OVOC1481
Dimm g8872.t1
EEL 0000131201-mRNA-1
Ctub g6773.t1
Cjohn rna-CJOHNSTONI LOCUS7976-2
Lbra g3469.t2
Dcau g5286.t1
Bpah BPAG 0001297001-mRNA-1
Wbanc g3218.t1
Bmal Bm8433d
Lloa EN70 11378
Mhie g17461.t1
Mpe Cam1 g00016140.t1
Mpe Cam2 g00089440.t1
Moz Brazil1 g00014440.t1*
Moz Venz1 g00074900.t1*
Moz Brazil1 g00019570.t1*
Moz Venz1 g00086500.t1*
Cel gtl2 F54D1.5
Ppa PPA01260
Tcalli g8939.t2
Sdigi contig18.g1602.t1
Dimm g5707.t1
Ooc g8257.t2
Ovol OVOC921
EEL 0000340301-mRNA-1
Cjohn rna-CJOHNSTONI LOCUS1526
Ctub g975.t1
Avit brkr g7154.t1
Lsig t00587-RA
Lbra g5602.t1
Dcau g1162.t1
Wbanc g2070.t2
Bmal Bm4357
Bpah BPAG 0000213301-mRNA-1
Lloa EN70 8268
Mpe Cam1 g00082660.t2
Mpe Cam2 g00008830.t1
Moz Brazil1 g00023660.t1
Moz Venz1 g00026640.t1
Clades
Clade 1
Clade 2

Ppa PPA16727
Ppa PPA35955
Cel glc-2 F25F8.2
Tcalli g1788.t1
Sdigi contig253.g6720.t1
Dimm g3652.t1
Ovol OVOC290
Ooc g13035.t1
Dcau g565.t1
Wbanc g11091.t1
Bpah BPAG 0000359301-mRNA-1
Bmal Bm3578
Avit brkr g7305.t1
Cjohn rna-CJOHNSTONI LOCUS7957
Ctub g350.t1
Lbra g981.t1
EEL 0000947401-mRNA-1
Ppa PPA09668
Ppa PPA15833
Cel avr-14-B0207.12b
EEL 0000617401-mRNA-1
Sdigi contig128.g4916.t1
Tcalli g3757.t1
Ovol OVOC388a
Ooc g12531.t1
Dimm g2618.t1
Dcau g387.t1
Cjohn rna-CJOHNSTONI LOCUS5729
Lsig t09042-RA
Lbra g9168.t1
Ctub g7697.t1
Avit brkr g9437.t1
Wbanc g2442.t1
Bpah BPAG 0001461301-mRNA-1
Bmal Bm1710b
Lloa EN70 4364
Mhie g99.t1
Mpe Cam1 g00088890.t1
Mpe Cam2 g00051440.t1
Moz Brazil1 g00058920.t1
Moz Venz1 g00029610.t1
Ppa PPA26611
Cel avr-15 R11G10.1a
Cel glc-1-F11A5.10
Ppa PPA12815
Cel glc-3 ZC317.3
Tcalli g8869.t1
Sdigi contig288.g7092.t1
Mhie g824.t1
Mpe Cam1 g00048460.t1
Mpe Cam2 g00031480.t1
Moz Brazil1 g00030970.t1
Moz Venz1 g00078820.t1
Dcau g5445.t1
Dimm g2749.t1
Ovol OVOC6009
Ooc g11828.t1
Wbanc g10913.t1
Bpah BPAG 0001302701-mRNA-1
Bmal Bm8050
Lloa EN70 2393
Lsig t03804-RA
Lbra g5002.t1
Ctub g3886.t1
Cjohn rna-CJOHNSTONI LOCUS9043
Avit brkr g728.t1
Clades
glc-2 like
avr-14 like
glc-3 like
Tree scale: 0.1