Published in final edited form as:

Vet Parasitol. 2020 October ; 286: 109225. doi:10.1016/j.vetpar.2020.109225.

Comparative sequences of the *Wolbachia* genomes of drug-sensitive and resistant isolates of *Dirofilaria immitis*

Pei-Tsz Shin^{a,b,1}, Rodrigo de Paula Baptista^{b,c}, Connor M. O'Neill^{a,b}, Connor Wallis^{a,b,2}, Barbara J. Reaves^{a,b,3}, Adrian J. Wolstenholme^{a,b,3,*}

^aDepartment of Infectious Diseases, University of Georgia, Athens, GA 30602 USA

^bCenter for Tropical and Emerging Global Diseases, University of Georgia, Athens, GA 30602 USA

^cInstitute of Bioinformatics, University of Georgia, Athens, GA 30602 USA

Abstract

The recent identification of isolates of *D. immitis* with confirmed resistance to the macrocyclic lactone preventatives presents an opportunity for comparative genomic studies using these isolates, and examining the genetic diversity within and between them. We studied the genomes of *Wolbachia* endosymbionts of five isolates of *D. immitis* maintained at the University of Georgia. Missouri and Georgia-2 are maintained as drug susceptible isolates, and JYD-27, Yazoo-2013 and Metairie-2014 are resistant to the macrocyclic lactone preventatives. We used whole genome amplification followed by Illumina-based sequencing from 8–12 individual microfilariae from each of the five isolates, obtaining a depth of coverage of approximately 40–75 fold for each. The Illumina sequences were used to create new genome assemblies for all the *Wolbachia* isolates studied. Comparisons of the *Wolbachia* sequences revealed more than 3,000 sequence variations in each isolate. We identified 67 loci specific in resistant isolates but not in susceptible isolates, including 18 genes affected. Phylogenetic analysis suggested that the endosymbionts of the drug-susceptible isolates are more closely related to each other than to those from any of the resistant parasites. This level of variation in the *Wolbachia* endosymbionts of *D. immitis* isolates suggests a potential for selection for resistance against drugs targeting them.

*Correspondence: Dr Adrian Wolstenholme, adrianw@uga.edu. Telephone +44 786 775 1908.

¹Current addresses: Department of Small Animal Medicine & Surgery, University of Georgia;

²University of Manchester, Manchester, UK;

³INRAE, Centre de recherche Val de Loire, UMR ISP – 213, 37380 Nouzilly, France

Author Contributions

P-TS, Investigation; Methodology; Writing - original draft. RPB, Data curation; Methodology, Writing – review & editing CO'N, Investigation; Methodology; Writing - review & editing. CW, Methodology. BJR; Supervision; Funding Acquisition; Writing - review & editing. AJW; Conceptualization; Supervision; Funding Acquisition; Writing - original draft; Writing – review & editing.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Declaration of Interest

The animal work was conducted under AUP A2017 05–002 from the University of Georgia Institutional Animal Care and Use Committee. The authors declare no financial conflicts of interest.

Keywords

Dirofilaria immitis; drug resistance; genome sequence; *Wolbachia*

Introduction

The existence of *Dirofilaria immitis* that are resistant to the macrocyclic lactone anthelmintics used to prevent heartworm disease has been established and thoroughly documented (Bourguinat et al., 2015; Cima, 2013; Pulaski, 2014; Wolstenholme et al., 2015), although their capacity to spread and the extent of the threat they present to canine health remains uncertain and is the subject of debate (Atkins et al., 2014). The discovery of genome sequence variations that strongly correlate with, or cause, drug resistance would greatly help resolve unanswered questions about the origin, spread and threat of resistance in this species. Some success in the development of such markers has been reported (Ballesteros et al., 2018), but the discovery of additional variations would still be of value, especially if they proved to be related to the mechanism of resistance.

Many filarial nematodes, including *D. immitis*, harbor a *Wolbachia* endosymbiont. Many aspects of this symbiotic relationship remain poorly understood, including the potential genetic and genomic variation between *Wolbachia* in different heartworm isolates, and their biological consequences. Variations in the endosymbiont copy number have been associated with biological variation in *Onchocerca volvulus* (Armoo et al., 2017), but no sequence comparisons of the complete *Wolbachia* genomes have been made between heartworm isolates. The use of doxycycline, which targets *Wolbachia*, is a standard component of treatment regimens for heartworm disease (Nelson et al., 2018; Savadelis et al., 2017), so sequence variations in the endosymbiont could theoretically impact the effectiveness of this treatment. Doxycycline targets the bacterial ribosome, and, in other species, resistance can be caused by mutations within the ribosomal binding site, mobile genetic elements carrying tetracycline-specific resistance genes, or mutations leading to increased expression of intrinsic resistance mechanisms (Grossman, 2016). Targeting *Wolbachia* is also a strategy being used in the development of novel anthelmintics (Clare et al., 2019), especially for the discovery of novel adulticide drugs, and sequence variations in the endosymbiont may lead to differences in the efficacy of such compounds on parasite species and isolates. Here, we report the level of variation between the *Wolbachia* sequences present in five *D. immitis* isolates, some of which are resistant to macrocyclic lactones. This variation raises the possibility that environmental pressures, including drug treatment, could act on the endosymbiont and result in phenotypic changes, including drug resistance.

2 Materials and Methods

2.1 Parasites

The *D. immitis* isolates, Missouri, Georgia-2, JYD-27, Yazoo-2013 and Metairie-2014, used in these studies were described previously (Berrafato et al., 2019; Maclean et al., 2017). Missouri and Georgia-2 are susceptible to macrocyclic lactone preventatives, JYD-27,

Yazoo-2013 and Metairie-2014 are resistant to them. Microfilariae (Mf) were isolated from the blood of infected animals as described previously (Maclean et al., 2017).

2.2 DNA extraction and sequencing

To extract the genomic DNA from single Mf of each isolate, the Mf were diluted in RPMI1640 to 0.05 Mf/mL so that a single Mf would be present in every 20µl. Individual Mf (12/isolate) were placed in a 0.2ml microcentrifuge tube and DNA extracted using the DNeasy Blood and Tissue Kit (Qiagen, Germantown, MD, USA). Whole Genome Amplification was performed using the TruPrime Whole Genome Amplification Kit (Lucigen, Middleton WI, USA). The isolated DNA from all single-Mf samples were sent to the Georgia Genomics and Bioinformatics Core (GGBC) at the University of Georgia to be sequenced via Illumina HiSeq using PE150. All sequence data are available from NCBI under Bioproject accession number PRJNA588995.

2.3 Genome assembly and analysis

The genome assembly for all samples was performed using a reference-based assembly method using as reference the *D. immitis* *Wolbachia* (wDi) version 2.2 described in Godel et al. (2012), available at http://genoweb.toulouse.inra.fr/~formation/M2_Phylogenomique/data/Wolbachia/Additional/. To perform this analysis all Illumina reads from each sample were aligned against the reference genome by using BWA mem (Li & Durbin, 2009). To assess the mapping percentage to the reference genome, SAMtools (v1.6) (Li and Durbin, 2009) flagstat tool was used. SAMtools mpileup (v1.6) (Li and Durbin, 2009) and BCFtools (v1.8) (<https://github.com/samtools/bcftools>) call and consensus tools were used to assemble the wDi genome, which was then checked with the quality assessment online tool QUAST (Gurevich et al., 2013). Gene annotation was performed using Prokka (v1.13) (Seemann, 2014) and a text-based format FASTA file for protein sequences was produced for each wDi isolate.

2.4 Wolbachia SNPs analysis

To identify single nucleotide polymorphism (SNP), all previously mapped, aligned and merged aligned files (BAM) were pre-processed with Picard (v2.16.0) (<http://broadinstitute.github.io/picard/>) and SAMtools (v1.6) (Li and Durbin, 2009). The Picard ValidateSamFile module was used before running GATK (v3.8–0) HaplotypeCaller recommended by GATK Best Practices (Van der Auwera et al., 2013). A hard filter was performed to remove false positive variants. Filter parameters were set at QD (variant confidence/quality by depth) < 2.0, FS (Phred-scaled p-value using Fisher's Exact Test to detect strand bias) > 25.0, MQ (Root Mean Square of mapping quality) < 40.0, SOR (StrandOddsRatio is used to detect strand bias) > 3.0, MQRankSum (Z-score from Wilcoxon rank sum test for mapping quality) < -7.5, ReadPosRankSum (Z-score from Wilcoxon rank sum test of alternative vs. reference read position bias) < -6.5, and DP (read depth) < 10 for single nucleotide variant (SNV), and QD (variant confidence/quality by depth) < 2.0, FS (Phred-scaled p-value using Fisher's Exact Test to detect strand bias) > 100.0, ReadPosRankSum (Z-score from Wilcoxon rank sum test of alternative vs. reference read position bias) < -15.0, SOR (StrandOddsRatio is used to detect strand bias) > 10.0, and DP (read depth) < 10 for indel separately. The GATK CombineVariants module was used to

identify SNPs specific in the resistant (JYD-27, Metairie-2014, and Yazoo-2013) and susceptible isolates (Georgia-2 and Missouri). SNP impact was predicted using SnpEff (v4.3t) (Cingolani et al., 2012). SNPs density of *w*Di genome was calculated with the window set at 500 bp using BEDTools (v2.26.0) (Quinlan and Hall, 2010). The SNPs density for each *w*Di isolate and the specific loci in resistant isolates were plotted using circos (v0.69–6) (Krzywinski et al., 2009). The specific genes in the resistant and susceptible isolates were blasted using NCBI BLAST (<http://blast.ncbi.nlm.nih.gov/>), and GO term for genes was found using Blast2GO InterPro module (Conesa and Gotz, 2008).

2.5 Wolbachia phylogeny

To construct the phylogenetic tree for *w*Di from five *D. immitis* isolates, *Wolbachia* from *O. ochengi* (*w*Oo) (accession number: HE660029.1 in the NCBI nucleotide database), *Wolbachia* in *O. volvulus* (*w*Ov) (accession number: HG810405.1 in the NCBI nucleotide database) and *Wolbachia* from *B. malayi* (*w*Bm) (accession number AE017321) in the supergroup C were selected as the outgroups. A dataset of 564 single-copy orthologous genes, shared among the eight *Wolbachia* strains was obtained using OrthoFinder (v2.1.2) (Emms and Kelly, 2015). According to the protein FASTA files generated from Prokka (v1.13) (Seemann, 2014), corresponding amino acid sequences for those orthologs were aligned. An amino acid supermatrix of the 8 *Wolbachia* was concatenated and aligned, and then generating a single phylip format alignment file with MAFFT (v7) (Katoh and Standley, 2013). Maximum-likelihood analysis of the amino acid supermatrix was performed with PhyML (v3.0) and the substitution model was selected by Modeltest-NG (Darriba et al., 2020) using the AIC criteria, and the robustness of nodes was assessed with bootstrap 1,000 replicates (Guindon et al., 2009). PhyML (v.3) was used to construct the final Maximum Likelihood (ML) tree (Guindon et al., 2010).

3 Results

After merging the HiSeq genomic data and mapping to the *w*Di genome reference, the mean read depth of coverage of the *w*Di genome was approximately 40–75 X (Table 1). We used the *w*Di genome assembly data from each isolate to generate a text-based format FASTA file for protein sequences to construct phylogenetic tree in the following step. To build data for the following SNP effect prediction, gene annotation of the *w*Di reference (version 2.2) (Godel et al., 2012) was performed. The results are shown in Table 2. To compare the *w*Di gene annotation results, *w*Oo and *w*Ov reference genomes were annotated with the same pipeline.

3.1 Wolbachia SNPs analysis

We used GATK (v3.8–0) (Van der Auwera et al., 2013), which primarily focuses on variant discovery and genotyping purposes, to find variations between the *w*Di reference and each *w*Di isolate. The generated text-based variant call format (VCF) file from GATK contains different parameters indicating input data quality, which allowed us to assess the results and remove the false positive variants. By combining these variants from different *w*Di isolates, we found positions in the *w*Di genome specific to resistant or susceptible *w*Di isolates. The generated VCF file from GATK is compatible with SnpEff, a software which can annotate

the variants and predict their effects. Following this pipeline, we found that more than 3,000 raw variants were generated for each isolate using the GATK HaplotypeCaller module. Hard filter results showed an average of 86% of the variants are passed, which means these variants are confident and can be included for the following analysis (Table 3). To show the SNP density in the *wDi* genome, we set the window size at 500 bp and constructed the line plots for each *wDi* isolate. In Figure 1, the five rings stand for the SNPs density for each *wDi* isolate and the red dots on the outmost scale represent the positions showed variation only in the resistant isolates, JYD-27, Metairie-2014, and Yazoo-2013, but not in the susceptible isolates. The distribution of these positions is not even. Around positions 255 kb, 265 kb, 360 kb, 450 kb, and 710 kb, there are more than 6 variations at each defined region. GATK CombineVariants and SnpEff results revealed, excluding upstream and downstream gene variants, 67 SNPs specific to the resistant isolates with 18 genes affected (Table 3). Among these 67 SNPs, there are 26 SNPs where the coverage was less than 10 from all three resistant isolates (marked with * and ** in Table 3). We kept them in our data because either coverage of two of the isolates were higher than 10 or the same gene was affected by many SNPs. Of these variations, 69% are non-synonymous variants resulting in protein sequence changes or premature stop codon. Blast2GO was used to further investigate the functions of the affected genes and the gene ontology relationships. The molecular function results from Blast2GO showed that, in the resistant isolates, there are 5 genes related to ATP binding (GO:0005524), which can further affect drug binding (GO:0008144). The five affected genes in the resistant isolates are the type I secretion system ATP-binding protein PrsD (*prsD*), transcription termination factor Rho (*rho*), phosphoribosylformylglycinamide synthase subunit PurL (*purL_2*), arginyl-tRNA synthetase (*argS*), and isoleucyl-tRNA synthetase (*ileS*). There are four genes, *rpoBC*, *rho*, *purL_2*, and *rplQ*, with variations in both resistant and susceptible isolates. The effects of SNPs resulting in *rpoBC* mutation in resistant isolates is one non-synonymous variant, and in susceptible isolates are five non-synonymous and one synonymous variants. For *rho* mutation, three non-synonymous and four synonymous variants in the resistant isolates whereas only one synonymous variant in the susceptible isolates. The SNPs affecting *purL_2* gene in the resistant and susceptible isolates are both non-synonymous variants. For the SNPs in the resistant and susceptible isolates affecting *rplQ* are synonymous and non-synonymous variants respectively. The allelic frequency of these variants is shown in Supplementary Table 1.

3.2 Wolbachia phylogeny

To construct the rooted phylogenetic tree for *wDi* from the five *D. immitis* isolates, *wBm*, *wOo* and *wOv* were used as outgroups. A total of 564 orthologous single-copy genes among these eight *Wolbachia* were identified using Orthofinder (v2.1.2) (Emms and Kelly, 2015). For each *Wolbachia* isolate, the corresponding amino acid sequences for those orthologs were concatenated and aligned. We then generated a single phylip alignment format file containing an amino acid supermatrix of the 8 *Wolbachia*, and then constructed a phylogenetic tree with the maximum likelihood method. The best substitution model predicted by Modeltest-NG for this dataset was Jones-Taylor-Thornton with 5 categories of discrete gamma distribution and frequencies estimated from the data (JTT+G+F). The resulted phylogenetic tree confirmed the separation of different *wDi* isolates and showed that *wDi* from two susceptible *D. immitis* isolates are indeed genetically different from the *wDi*

from three resistant isolates (Figure 2). In addition, compared to the resistant isolates, the relatedness of *w*Di from the two susceptible isolates is closer to each other. The results of the maximum likelihood method showed bootstrap values presented in percent at the nodes are high, 93–100%, to support the phylogenetic tree (Figure 2).

4 Discussion

We determined the sequence of the *Wolbachia* genome from five isolates of *D. immitis* all of which were isolated from the South-eastern United States during the current century. The sequences were compared with other and with that deposited by Godel et al (2012). These comparisons revealed over 3,000 variations between the isolates, about 0.3% of the genome. Some of these variations were specific to the drug-resistant isolates studied, though the number of such isolates was too small to draw any inferences from this. However, the *Wolbachia* sequences from the two drug-susceptible isolates were more closely related to each other than to those from any of the resistant ones, despite a considerable geographical and temporal separation in their original isolation. Several of the variations we discovered could have a significant impact on the functions of the proteins that would be produced by the endosymbiont, in particular the premature stop codons in *nusB* and *acnA*, and the frameshift mutations observed in *gph*, *ubiE*, *rpe* and *virD4*. Interestingly, *O. volvulus* parasites that responded sub-optimally to ivermectin did carry reduced *Wolbachia* loads (Armoo et al., 2017), though the authors did not report if any sequence variations correlated with either the copy-number of the endosymbiont or the drug response of the parasite, so it is possible that variations in *Wolbachia* are associated with drug resistance. However, this level of sequence variation does indicate that the *Wolbachia* genome has the potential to evolve in response to selection pressures; one such pressure could be the application of antibiotics as therapy for heartworm or other filarial diseases. Doxycycline is an established component of treatment for heartworm-infected dogs; in this instance selection for resistance may be limited as the drug is one component of the three recommended in the American Heartworm Society guidelines (Nelson et al., 2018). However, if only a combination of a macrocyclic lactone, such as moxidectin, and the antibiotic are used to treat infections with macrocyclic lactone-resistant parasites (Savadelis et al., 2017) then the potential for the evolution of double-resistant heartworms may exist. Resistance to tetracyclines, including doxycycline, is established in mycoplasmas, which are intracellular bacterial pathogens related to *Wolbachia* (Yatoo et al., 2019). In human filarial disease, *Wolbachia* have been proposed as a good target for the development of adulticidal drugs (Slatko et al., 2010; Taylor et al., 2019) and in this case a sole reliance on such drugs could also theoretically promote the development of resistance.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by grants from the American Heartworm Society and the National Institutes of Health (R01AI103140 & R21AI125899). None of the funders had any influence on the study design or the decision to publish. We should like to thank Andrew Moorhead, Pablo Jimenez Castro and the staff of the Filarial Research

Reagent Resource Center for their help in maintaining the heartworm infections and in Mf isolation. We should also like to thank the anonymous reviewers for their helpful and constructive comments.

References

- Armoo S, Doyle SR, Osei-Atweneboana MY, Grant WN, 2017 Significant heterogeneity in *Wolbachia* copy number within and between populations of *Onchocerca volvulus*. *Parasites & Vectors* 10, 188. [PubMed: 28420428]
- Atkins CE, Murray MJ, Olavessen LJ, Burton KW, Marshall JW, Brooks CC, 2014 Heartworm 'lack of effectiveness' claims in the Mississippi delta: Computerized analysis of owner compliance-2004–2011. *Vet. Parasitol* 206, 106–113. [PubMed: 25440944]
- Ballesteros C, Pulaski CN, Bourguinat C, Keller K, Prichard RK, Geary TG, 2018 Clinical validation of molecular markers of macrocyclic lactone resistance in *Dirofilaria immitis*. *Int.J. Parasitol.* *Drugs Drug Res* 8, 596–606.
- Berrafato T, Coates R, Reaves BJ, Kulke D, Wolstenholme AJ, 2019 Macrocyclic lactone anthelmintic-induced leukocyte binding to *Dirofilaria immitis* microfilariae: influence of the drug resistance status of the parasite. *Int. J. Parasitol.* *Drugs Drug Res* 10, 45–50.
- Bourguinat C, Lee ACY, Lizundia R, Blagburn BL, Liotta JL, Kraus MS, Keller K, Epe C, Letourneau L, Kleinman CL, Paterson T, Gomez EC, Montoya-Alonso JA, Smith H, Bhan A, Peregrine AS, Carmichael J, Drake J, Schenker R, Kaminsky R, Bowman DD, Geary TG, Prichard RK, 2015 Macrocyclic lactone resistance in *Dirofilaria immitis*: Failure of heartworm preventives and investigation of genetic markers for resistance. *Vet. Parasitol* 210, 167–178. [PubMed: 25936435]
- Cima G, 2013 Heartworm preventive resistance is real. *J. Am. Vet. Med.Assoc* 243, 1230–1230. [PubMed: 24288828]
- Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM, 2012 A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly* 6, 80–92. [PubMed: 22728672]
- Clare RH, Clark R, Bardelle C, Harper P, Collier M, Johnston KL, Plant H, Plant D, McCall E, Slatko BE, Cantin L, Wu B, Ford L, Murray D, Rich K, Wigglesworth M, Taylor MJ, Ward SA, 2019 Development of a high-throughput cytometric screen to identify anti-*Wolbachia* compounds: The power of public–private partnership. *SLAS Discov.* 24, 537–547. [PubMed: 30958712]
- Conesa A, Gotz S, 2008 Blast2GO: A comprehensive suite for functional analysis in plant genomics. *Int. J. Plant. Genom* 2008, 619832.
- Darriba D, Posada D, Kozlov AM, Stamatakis A, Morel B, Fluori T, 2020 ModelTest-NG: A new and scalable tool for the selection of DNA and protein evolutionary models. *Mol. Biol. Evol* 37:291–294. [PubMed: 31432070]
- Emms DM, Kelly S, 2015 OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* 16, 157. [PubMed: 26243257]
- Godel C, Kumar S, Koutsovoulos G, Ludin P, Nilsson D, Comandatore F, Wrobel N, Thompson M, Schmid CD, Goto S, Bringaud F, Wolstenholme A, Bandi C, Epe C, Kaminsky R, Blaxter M, Maser P, 2012 The genome of the heartworm, *Dirofilaria immitis*, reveals drug and vaccine targets. *FASEB J* 26, 4650–4661. [PubMed: 22889830]
- Grossman TH, 2016 Tetracycline antibiotics and resistance. *Cold Spring Harb Perspect Med.* 6, a025387. [PubMed: 26989065]
- Guindon S, Delsuc F, Dufayard JF, Gascuel O, 2009 Estimating maximum likelihood phylogenies with PhyML. *Methods Mol. Biol* 537, 113–137. [PubMed: 19378142]
- Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O, 2010 *Syst. Biol* 59, 307–21 [PubMed: 20525638]
- Gurevich A, Saveliev V, Vyahhi N, Tesler G, 2013 QUASt: quality assessment tool for genome assemblies. *Bioinformatics* 29, 1072–1075. [PubMed: 23422339]
- Katoh K, Standley DM, 2013 MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol. Biol. Evol* 30, 772–780. [PubMed: 23329690]

- Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA, 2009 Circos: an information aesthetic for comparative genomics. *Genome Res* 19, 1639–1645. [PubMed: 19541911]
- Kumar S, Stecher G, Tamura K, 2016 MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol* 33, 1870–1874. [PubMed: 27004904]
- Li H, Durbin R, 2009 Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. [PubMed: 19451168]
- Maclean MJ, Savadelis MD, Coates R, Dzimirski MT, Jones C, Benbow C, Storey BE, Kaplan RM, Moorhead AR, Wolstenholme AJ, 2017 Does evaluation of in vitro microfilarial motility reflect the resistance status of *Dirofilaria immitis* isolates to macrocyclic lactones? *Parasites & Vectors* 10, 480.
- Nelson CT, McCall J, Jones S, Moorhead A 2018 Current canine guidelines for the prevention, diagnosis, and management of heartworm (*Dirofilaria immitis*) infection in dogs, American Heartworm Society, (Wilmington, DE, USA).
- Pulaski CN, Malone JB, Bourguinat C, Prichard R, Geary T, Ward D, Klei TR, Guidry T, Smith GB, Delcambre B, Bova J, Pepping J, Carmichael J, Schenker R, Pariaut R, 2014 Establishment of macrocyclic lactone resistant *Dirofilaria immitis* isolates in experimentally infected laboratory dogs. *Parasites & Vectors* 7, 494. [PubMed: 25376278]
- Quinlan AR, Hall IM, 2010 BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. [PubMed: 20110278]
- Savadelis MD, Ohmes CM, Hostetler JA, Settje TL, Zolynas R, Dzimirski MT, Moorhead AR, 2017 Assessment of parasitological findings in heartworm- infected beagles treated with Advantage Multi (R) for dogs (10% imidacloprid+2.5% moxidectin) and doxycycline. *Parasites & Vectors* 10, 245. [PubMed: 28526088]
- Seemann T, 2014 Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. [PubMed: 24642063]
- Slatko BE, Taylor MJ, Foster JM, 2010 The *Wolbachia* endosymbiont as an anti-filarial nematode target. *Symbiosis* 51, 55–65. [PubMed: 20730111]
- Taylor MJ, von Geldern TW, Ford L, Hubner MP, Marsh K, Johnston KL, Sjoberg HT, Specht S, Pionnier N, Tyrer HE, Clare RH, Cook DAN, Murphy E, Steven A, Archer J, Bloemker D, Lenz F, Koschel M, Ehrens A, Metuge HM, Chunda VC, Chounna PWN, Njouendou AJ, Fombad FF, Carr R, Morton HE, Aljayyousi G, Hoerauf A, Wanji S, Kempf DJ, Turner JD, Ward SA, 2019 Preclinical development of an oral anti- *Wolbachia* macrolide drug for the treatment of lymphatic filariasis and onchocerciasis. *Science Translational Medicine* 11, eaau2086. [PubMed: 30867321]
- Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, Jordan T, Shakir K, Roazen D, Thibault J, Banks E, Garimella KV, Altshuler D, Gabriel S, DePristo MA, 2013 From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics* 43, 11 10 11–33. [PubMed: 25431634]
- Wolstenholme AJ, Evans CC, Jimenez PD, Moorhead AR, 2015 The emergence of macrocyclic lactone resistance in the canine heartworm, *Dirofilaria immitis*. *Parasitology* 142, 1249–1259. [PubMed: 26040450]
- Yatoo MI, Parray OR, Bhat RA, Muheet, Gopalakrishnan A, Saxena A, Chakraborty S, Tiwari R, Khurana SK, Singh SV, Dhama K, 2019 Emerging antibiotic resistance in mycoplasma microorganisms, designing effective and novel drugs / therapeutic targets: current knowledge and futuristic prospects. *J. Pure Appl. Microbiol* 13, 27–44.

Figure 1. The density of sequence variations in the *wDi* genome from each *D. immitis* isolate. The outermost scale indicates coordinates in kilobase pairs (kbp) in the *wDi* genome. These five rings represent the SNPs density in *wDi* from five *D. immitis* isolates using a window size of 500 bp. The red dots on the scale show the loci found specific to the resistant isolates. (Green: Yazoo-2013, orange: Metairie-2014, blue: JYD-27, purple: Missouri, yellow: Georgia-2).

Figure 2. Phylogenetic tree showing the relationships between the *Wolbachia* sequences.

The topology and branch lengths are based on the construction of the maximum likelihood method. Values at nodes are bootstrap values calculated from the maximal likelihood methods and given in percent as numbers. *wBm*, *wOo* and *wOv* are used as outgroups to root the phylogenetic tree. (GA2: Georgia-2; MO: Missouri; MET: Metairie-2014; JYD: JYD-27; Yazoo: Yazoo-2013; *wBm*: *Wolbachia* from *B. malayi*; *wOo*: *Wolbachia* from *O. ochengi*; *wOv*: *Wolbachia* from *O. volvulus*)

Table 1.Depth of coverage of *wDi* genome sequences

<i>wDi</i> isolate	Genome Length (bp)	Mean Depth of Coverage
Georgia 2	920663	49 X
Missouri	920577	38 X
JYD-27	920683	77 X
Metairie-2014	920720	64 X
Yazoo-2013	920763	69 X

Table 2.

The number of sequence variations between the *wDi* sequences of the isolates studied here and the reference sequence (Godel et al., 2012b).

<i>wDi</i> isolate	SNPs	SNPs in coding regions	Indels	Total Variations
Georgia 2	3311	2937	661	3972
Missouri	2944	2579	509	3453
JYD-27	2222	1931	429	2651
Metairie-2014	2736	2501	611	3347
Yazoo-2013	2858	2518	607	3465

Table 3.

Summary of 67 sequence variations specific to the resistant isolates.

Position	Reference allele	Alternative allele	SNP effect	Gene	Protein	Predicted amino- acid change
798 *	T	TC	Upstream gene variant	tlyC	Hemolysin C	None
127027	T	G	Missense variant		Hypothetical protein	Lys476Gln
218531	C	T	Missense variant	rpoBC	Bifunctional DNA-directed RNA polymerase subunit beta-beta'	Val1338Ile
237140 *	C	G	Missense variant	prsD	Type I secretion system ATP-binding protein PrsD	Pro440Arg
237170 *	C	T	Missense variant	prsD	Type I secretion system ATP-binding protein PrsD	Thr450Ile
248697 **	G	T	Synonymous variant	acnA	Aconitate hydratase A	None
248717 **	G	A	Missense variant	acnA	Aconitate hydratase A	Gly226Asp
248737 **	G	T	Stop gained	acnA	Aconitate hydratase A	Glu233 *
248745 **	A	G	Synonymous variant	acnA	Aconitate hydratase A	None
255515	G	T	Missense variant	nusA	Transcription termination/antitermination protein NusA	Lys200Asn
255516	G	A	Missense variant	nusA	Transcription termination/antitermination protein NusA	Val201Ile
255537	G	A	Missense variant	nusA	Transcription termination/antitermination protein NusA	Val208Ile
255541	A	G	Missense variant	nusA	Transcription termination/antitermination protein NusA	Lys209Arg
255557	A	G	Synonymous variant	nusA	Transcription termination/antitermination protein NusA	None
255559	G	A	Missense variant	nusA	Transcription termination/antitermination protein NusA	Arg215His
264389 *	A	G	Synonymous variant	rho	Transcription termination factor Rho	None
264392 *	A	G	Synonymous variant	rho	Transcription termination factor Rho	None
264407 *	T	C	Synonymous variant	rho	Transcription termination factor Rho	None
264409 *	C	T	Missense variant	rho	Transcription termination factor Rho	Ser283Phe
264413 *	G	A	Synonymous variant	rho	Transcription termination factor Rho	None
264418 *	A	T	Missense variant	rho	Transcription termination factor Rho	Asp286Val
264424 *	C	T	Missense variant	rho	Transcription termination factor Rho	Pro288Leu
301681 *	TC	T	Frame-shift	rpe	Ribulose-phosphate 3-epimerase	Asp13fs
301684 *	G	A	Missense variant	rpe	Ribulose-phosphate 3-epimerase	Ala12Val
301685 *	C	T	Missense variant	rpe	Ribulose-phosphate 3-epimerase	Ala12Thr

Position	Reference allele	Alternative allele	SNP effect	Gene	Protein	Predicted amino- acid change
301698 *	A	G	Synonymous variant	rpe	Ribulose-phosphate 3-epimerase	None
333916	C	T	Missense variant	purL_2	Phosphoribosylformylglycine midline synthase subunit PurL	Pro71Ser
348374	A	AC	Frame-shift	virD4	Protein VirD4	His7fs
355534	A	T	Stop gained	nusB	Transcription antitermination protein NusB	Tyr126 *
363850 *	G	A	Missense variant	prmC	Release factor glutamine methyltransferase	Gly122Ser
363851 *	G	A	Missense variant	prmC	Release factor glutamine methyltransferase	Gly122Asp
363873 *	G	A	Synonymous variant	prmC	Release factor glutamine methyltransferase	None
363882 *	G	A	Synonymous variant	prmC	Release factor glutamine methyltransferase	None
363886 *	A	G	Missense variant	prmC	Release factor glutamine methyltransferase	Ser134Gly
363892 *	T	C	Missense variant	prmC	Release factor glutamine methyltransferase	Tyr136His
455084	C	T	Synonymous variant	argS	Arginine--tRNA ligase	None
455085	T	C	Missense variant	argS	Arginine--tRNA ligase	Lys486Arg
455092	A	G	Synonymous variant	argS	Arginine--tRNA ligase	None
455120	A	G	Synonymous variant	argS	Arginine--tRNA ligase	
455135	AAG	A	Frame-shift	argS	Arginine--tRNA ligase	Leu469fs
455138	G	GAA	Frame-shift	argS	Arginine--tRNA ligase	Leu469fs
455145	C	T	Missense variant	argS	Arginine--tRNA ligase	Ser466Asn
455152	G	A	Missense variant	argS	Arginine--tRNA ligase	Leu464Phe
455169	C	T	Missense variant	argS	Arginine--tRNA ligase	Arg458His
460335	C	T	Synonymous variant	ubiE	Ubiquinone/menaquinone biosynthesis C-methyltransferase UbiE	None
460336	A	G	Missense variant	ubiE	Ubiquinone/menaquinone biosynthesis C-methyltransferase UbiE	Leu142Pro
460343	AT	A	Frame-shift	ubiE	Ubiquinone/menaquinone biosynthesis C-methyltransferase UbiE	Leu140fs
460356	C	T	Synonymous variant	ubiE	Ubiquinone/menaquinone biosynthesis C-methyltransferase UbiE	None
636639	T	G	Upstream gene variant	lolE	Lipoprotein-releasing system transmembrane protein LolE	None
636699	AGCACC	A	Upstream gene variant	lolE	Lipoprotein-releasing system transmembrane protein LolE	None
636706	C	CTGA	Upstream gene variant	lolE	Lipoprotein-releasing system transmembrane protein LolE	None
636813	ACC	A	Upstream gene variant	lolE	Lipoprotein-releasing system transmembrane protein LolE	None

Position	Reference allele	Alternative allele	SNP effect	Gene	Protein	Predicted amino- acid change
643957	G	T	Synonymous variant	rplQ	50S ribosomal protein L17	None
690200	C	T	Upstream gene variant	pipB2	Secreted effector protein PipB2	None
713047	C	A	Upstream gene variant	ctaA	Heme A synthase	None
713054	G	A	Upstream gene variant	ctaA	Heme A synthase	None
713065	A	G	Upstream gene variant	ctaA	Heme A synthase	None
713077	G	A	Upstream gene variant	ctaA	Heme A synthase	None
713091	C	T	Upstream gene variant	ctaA	Heme A synthase	None
713113	TA	T	Upstream gene variant	ctaA	Heme A synthase	None
713143	T	C	Upstream gene variant	ctaA	Heme A synthase	None
713148	T	TG	Upstream gene variant	ctaA	Heme A synthase	None
775024	C	G	Missense variant	fumC	Fumarate hydratase class II	Lys211Asn
809307	ATT	A	Frame-shift	gph	Phosphoglycolate phosphatase	Ile179fs
823794 *	T	C	Upstream gene variant		Hypothetical protein	None
823798 *	A	G	Upstream gene variant		Hypothetical protein	None
848154	A	C	Missense variant	ileS	Isoleucine--tRNA ligase	Tyr651Asp

* indicates that coverage in one of the resistant isolates was <10 for this polymorphism.

** indicates that coverage for this polymorphism was <10 for all three resistant isolates.