Journal of the American Association for Laboratory Animal Science
Copyright 2023
by the American Association for Laboratory Animal Science

Vol 99, No 99
Month 2023
Pages 55–63

Assembly of a Draft Genome for the Mouse Ectoparasite *Myocoptes musculinus*

Thomas A Randall[1] and David M Kurtz[2,*]

Myocoptes musculinus **is a common ectoparasite of wild mice and is occasionally found on research mice. Infestations of research mice are often subclinical but can cause severe dermatitis. Perhaps more importantly, infestations can cause immunologic reactions that may alter research outcomes, and most animal research facilities strive to prevent or eliminate mites from their mouse colonies. *M. musculinus* infestations are currently detected by using microscopic evaluation of the fur and skin and PCR assays of pelt swabs targeting the rRNA genes of this mite. In our facility, we encountered multiple, false-positive 18S rRNA PCR results from a closed mouse colony. We could not identify the source of the false positives even after performing PCR analysis of other *Myocoptes* gene targets using assays developed from the few other target genomic sequences available for *M. musculinus* or *Myocoptes japonensis* in public databases. This situation highlighted the limited genetic resources available for development of diagnostic tests specific for this ectoparasite. To expand the available genetic resources, we generated a metagenome of *M. musculinus* derived by sequencing from fur plucks of an infected mouse. We also determined the completeness of this metagenome and compared it with those of related mites.**

Abbreviation: ITS, internal transcribed spacer

DOI: 10.30802/AALAS-JAALAS-22-000066

Introduction

Myocoptes musculinus is a nonburrowing mouse ectoparasite in the superorder Acariformes and family Mycoptidae. Although its prevalence in research mouse populations is considered relatively low (0.12% of samples submitted to a large commercial diagnostic testing laboratory[32]), *M. musculinus* is the most common mite found on research mice.[4] Clinical signs of *M. musculinus* infestation can range from subclinical to severe dermatitis that warrants euthanasia.[4,5] In some mouse strains, *M. musculinus* can elicit Th2 cytokine-mediated responses and elevated IgE concentrations.[23,38,42,47,51] As such, most research institutions work to prevent or eliminate these mites from their mouse populations. The treatment of research rodents to eliminate ectoparasites, including *M. musculinus*, has primarily used the avermectin and milbemycin classes of macrocyclic lactones such as ivermectin, selamectin (avermectins), and moxidectin (milbemycin).[35]

Historically, murine ectoparasite infections was diagnosed by microscopic examination of the fur and skin and identification of the adult or larval stages.[5,22,24] In 2011, 2 commercial testing laboratories began offering PCR assays for the detection *M. musculinus*; both assays targeted the 18S rRNA sequences and likely were based on sequences reported in GenBank (accession no., KT384411.1 or JF834893.1). In 2013, we incorporated this commercial PCR testing of mouse pelt swabs into our regular animal health surveillance program. Our animal facility had been negative for murine ectoparasites for at least the preceding 35 y. In Fall 2016, we received results from one commercial testing laboratory indicating that a sentinel mouse was PCR-positive for *M. musculinus*. We sent backup pelt swabs from the same PCR-positive sentinel mouse to another commercial testing laboratory. This second testing lab obtained a PCR-positive result from their Myocoptes genus-level PCR assay, but the sample tested PCR-negative on their *M. musculinus* species-specific assay.

Our own *M. musculinus* 18S rRNA PCR assay was based on primer sequences provided by Dr. Susan Compton (Yale University, New Haven, CT). We obtained fur plucks from a confirmed *M. musculinus*-infected mouse colony (kindly gifted by Dr. Julie Watson, Johns Hopkins University, Baltimore, MD) and used the isolated DNA as our positive control. The same initial, positive sample from our sentinel mouse was likewise PCR-positive in our assay. We sequenced the amplicon and found it to be a 100% match to nucleotides 406 through 810 of the *M. musculinus* 18S rRNA GenBank entry KT384411.1. Over a period of 2 mo, we tested approximately 615 swabs from either mouse pelts or cages housing mice from PCR-positive colonies. Of these, 16 samples were PCR-positive for *M. musculinus* according to both outside testing labs and our inhouse PCR and sequencing tests. We also performed visual pelt or fur pluck examinations of 215 mice from these PCR-positive populations but found no physical evidence of *M. musculinus*. After multiple tests performed over approximately a 2 mo period, we could no longer detect PCR-positive samples from these experimental populations. Based on the lack of physical evidence of *M. musculinus*, the ambiguous PCR results from one commercial testing lab, and the abrupt halt in finding PCR-positive samples, we concluded that these previous PCR results were false positives.

Approximately 21 *Myocoptes* species are known, with a worldwide distribution and a range of susceptible rodent species. Accurate morphologic classification of mites is difficult.[56]

Submitted: 30 Jun 2022. Revision requested: 08 Aug 2022. Accepted: 22 Oct 2022.
[1]Integrative Bioinformatics and [2]Comparative Medicine Branch, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina
**Corresponding author. Email: david.kurtz@nih.gov*
This article contains supplemental materials online.

http://prime-pdf-watermark.prime-prod.pubfactory.com/ | 2025-02-25

Current PCR assays for classification are based on 18S rRNA, which presents known difficulties as a marker because the assays may not always have resolution down to the genus and species levels. Visual inspection and sequence confirmation can produce contradictory results.[15] In addition, only 2 of these species—*M. myocoptes* and *M. japonensis*—have genomic sequences in GeneBank. Of the 18 total GenBank entries for *Myocoptes* spp., 6 are for the 18S and 28S rRNA genes and one for the rRNA internal transcribed spacer (ITS) region. This situation means that the current development of useful PCR assays is severely limited to just a few well-conserved regions of these 2 mite species. The lack of diversity in terms of orthologous sequences available in multiple related species makes it difficult to design primers that are species-specific. For example, we developed PCR assays based on heat-shock protein cognate 5 (Hsc70-5; accession no. JQ001561.1) and elongation factor 1α (Ef1α) of *M. japonensis* (JQ000939.1). These assays yielded positive results with our known *M. musculinus* DNA sample that were confirmed by sequencing, thus illustrating the potential lack of specificity at the species level.

An additional complication to consider regarding mite identification is the considerable controversy concerning the phylogenetic placement of Astigmata mites in the Acariformes superorder. In part, this difficulty appears to be due to a faster rate of gene evolution in Astigmata mites compared with other genera, thus causing confusion, particularly in the relationship between the Astigmata and Oribatida suborders, because of long branch attraction.[11] Whether Acari are even monophyletic is in considerable doubt.[53,56]

Given the difficulties in accurately identifying *M. musculinus* by using currently available PCR assays, we decided to test the hypothesis that sequence generated from the positive-control fur sample mentioned earlier would be sufficient to generate a full draft metagenome of the *M. musculinus* mite. Our results below indicate we have done so. We anticipate that this information will be a valuable resource for the development of more diverse and accurate diagnostic tests for both *M. musculinus* and other murine ectoparasites.

Materials and Methods

Isolation of total genomic DNA and sequence alignment. Fur plucks collected from a B6 × 129 background mouse colony that had been confirmed as monoinfected with *M. musculinus* mites were obtained from Dr. Julie Watson at Johns Hopkins University.[45] Genomic DNA was isolated from the fur plucks (DNeasy Blood and Tissue Kit, Qiagen, Hilden, Germany) based on the kit instructions except that we extended the initial proteinase K digestion to overnight at 56 °C. Total genomic DNA concentration was quantified fluorometrically by using a spectrophotometer–fluorometer (model DS-11 FX, DeNovix, Wilmington, DE) and assay kit (dsDNA Broad Range Assay Kit, DeNovix) according to the manufacturer's instructions. All multiple-sequence alignments were generated by using Multiple Sequence Comparison by Log- Expectation (MUSCLE)[14] within CLC Genomics Workbench (version 11, Qiagen).

Genome sequencing and assembly. The initial sequencing dataset consisted of 113,349,552 paired-end, 151-bp reads generated by using the NextSeq 4500 (Illumina, San Diego, CA) at the NIEHS Epigenetics and DNA Sequencing Core Facility (Research Triangle Park, NC). The reads were trimmed to remove adapter sequence and aligned to mm10 with the STAR assembler[12] to filter out mouse-specific reads. The unaligned output was further filtered through a database consisting of de novo genome assemblies of C57BL/6NJ, 129S1/SvImJ, and C3H/HeJ[31] and available from The Mouse Genomes Project website,[39] which also used STAR and resulted in a paired-end dataset of 13,110,845 reads (NCBI Project PRJNA840911). This dataset was assembled by using SPAdes 3.11.0[6] and testing a range of kmers; an optimal assembly based on N50 was found by using a kmer of 55. This assembly had a size of 66.4 Mb consisting of 14,487 scaffolds greater than 300 bp, with an N50 of 10.5 kb and a GC content of 29.83%. These statistics were determined by using Quast 4.0.[20] We also tested 2 other de novo assemblers—soapdenovo[30] and CLC Genomics (version 12.0, Qiagen, Redwood City, CA); at any kmer tested, these programs produced inferior assemblies to that generated with SPAdes. No improvement of the assembly was found after analysis with EukRep, which filters eukaryotic from prokaryotic reads,[55] and we therefore considered the original assembly to be our final assembly. Repeat content was determined by using MISA for microsatellites (default settings)[52] and RepeatMasker.[49] tRNAs were predicted by using tRNAscan-SE 1.3.[33]

Protein prediction and comparison. We used 2 algorithms, Augustus 2.2.5[50] and SNAP (snap-2013-11-29),[26] to independently predict proteins sets from our genome assembly. Of the available training sets for SNAP, we had previously empirically determined that the jewel wasp *Nasonia vitripennis* training set was optimal for protein prediction from *Dermatophagoides pteronyssinus*[44] and therefore we used this set. For Augustus, we developed a training set based on the *Sarcoptes scabiei* protein dataset.[46] BUSCO 3.0.1[48] was used to assess the genome completeness. We used Orthovenn2 for genome-level ortholog prediction of the 6 mite species with available public proteomes, with an E value of 1e-05 and an inflation of 1.5.[54]

Phylogenetic tree analysis. Multiple sequences for each of the genes were aligned by using MAFFT v5[25]. This result was concatenated for analysis with MRBAYES[21] and by maximum likelihood using MEGA 7.[28] In MRBAYES, the protein substitution model was set to Jones and run until the convergence was less than 0.01. *D. melanogaster* was set as the outgroup. In MEGA, the protein substitution Jones–Taylor–Thornton matrix was used, and 100 bootstraps were done.

Sequence searches. To determine whether our predicted protein set contained any candidate allergen homologs, we downloaded the amino acid sequences of the official list of allergens from the dust mite *D. pteronyssinus* (www.allergen.org)[43] and used BLAST[2] to query these sequences against our predicted protein set in *M. musculinus* and against that of the most closely related mite, *S. scabiei*. A protein was considered to be a candidate allergen homolog when it contained at least 50% sequence identity to a known *D. pteronyssinus* allergen, an accepted cutoff used in the allergen community.[16] We then conducted a BLAST search against the Uniref90 database (https://www.uniprot.org/downloads to annotate the *M. musculinus*-specific set of 215 proteins described in the results.[10] The rDNA repeat was annotated by using RNAmmer.[29] From this rDNA repeat, the 18S (OP361390), 28S (OP361389), and the ITS1-5.8S-ITS2 (OP361388) regions were submitted to Genbank with the identifiers noted.

Results

Assembly statistics and gene content. Our source for the *M. musculinus* metagenome was DNA from a mouse fur pluck. Assuming that this sample was primarily mouse DNA, we first filtered out the reads from this sample with the standard mm10 mouse genome assembly from C57BL/6. We thought an additional filter would be useful, so we chose 3 from the Welcome Sanger Institute that represented commonly used mouse

http://prime-pdf-watermark.prime-prod.pubfactory.com/ | 2025-02-25

genomes. This strategy removed a few percent more of the reads, leaving us with a collection of 13,110,845 reads from the original set of 113,349,552 paired-end reads, so approximately 10% of the reads could not be assigned to any mouse genome and thus represented our initial dataset. From this dataset, we assembled a metagenome of 66 Mb containing 14,487 contigs greater than 300 bp with an N50 of 10.4 kb, thus representing 59.5× genome coverage. The GC content of this dataset was 29.83%, comparable to other related Astigmata mites, *S. scabiei* (33.3%), *D. pteronyssinus* (29.03%), and *Dermatophagoides farina* (30.8%). Two major questions that we considered were the completeness of the metagenome and its accurate genus and species identification. We addressed these issues in several ways. We performed a BUSCO analysis by using the conserved ortholog set for arthropods, bacteria, fungi, and vertebrates and for the 3 related mites, all derived from monoinfested colonies. We found that 92% of the arthropod BUSCO conserved ortholog dataset was identified in our metagenome of *M. musculinus*, and the conserved ortholog profile of the other datasets was similar to those of the other mites analyzed (Table 1). Thus, the completeness of our metagenome is comparable to those of 3 other mites that were derived from pure cultures. We used 2 tools to compile a set of predicted proteins for our metagenome. Augustus predicted 11,818 proteins whereas SNAP predicted 17,466 proteins. Combined, these 2 datasets had an overlap of 9,839 total predicted proteins in common. The merged, final set contained 13,481 unique proteins of at least 100 amino acids (Supplementary Table S1).

Phylogenetic placement. For a more definitive placement of our metagenome, we did a phylogenetic analysis. A set of 10 ribosomal proteins, each in a one-to-one orthologous relationship with no paralogs in the species used, was identified for each of the following organisms: the dust mites *D. pteronyssinus*[44] and *D. farinae*;[8] scabies mite, *Sarcoptes scabiei*;[3,37,46] spider mite, *Tetranychus urticae*;[18] and deer tick, *Ixodes scapularis*.[19,41] We included these organisms because the available and complete mite genomes were publicly available and had predicted protein datasets. We used the water flea *Daphnia pulex*[9] and fruit fly *D. melanogaster*[1] used as nonmite outgroups. As can be seen in Figure 1, our set of genes is clearly within the Astigmata clade containing *S. scabiei* and both *Dermatophagoides* species.

Mitochondrial genome, rRNA, tRNA, and repetitive elements. The completeness of the *M. musculinus* mitochondrial genome and rRNA repeat suggested the robustness of our metagenome. We found an intact mitochondrial genome of 13,692 bp as a single contig. This contig had 99% to 100% sequence identity to all *M. musculinus* mitochondrial sequences in Genbank. The mitochondrial genome of *S. scabiei* was 13,667 bp (NC031334.1). A comparison of the gene content is shown in Figure 2. Genes missing from the *M. musculinus* mitochondrial genome are *trnA*, *trnR*, and *OL*; genes missing from the *S. scabiei* mitochondrial genome are *trnK*, *OL*, *trnA*, *trnF*, *trnY*, and *atp8*; and genes missing from the *D. pteronyssinus* mitochondrial genome (NC012218.1) are *trnA*, *trnR*, and *OL*. All other mitochondrial genes annotated are present in all 3 genomes. A list of Acariformes mitochondrial genome annotations is provided in Supplementary Table S2.

Table 1. BUSCO analysis of Astigmata mite genomes

Genome analyzed	% of genes in each genome analyzed			
	Arthropods	Bacteria	Fungi	Vertebrates
M. musculinus	92.0	18.9	78.9	39.2
S. scabiei	89.7	8.7	68.9	34.8
D. pteronyssinus	88.7	43.9	72.4	40.6
D. farina	86.7	68.2	73.4	34.0

Percentage of genes in each category that are found in each genome analyzed. Total genes in ach BUSCO set of single-copy conserved orthologs: Arthropods; 1066 genes, Bacteria; 148 genes, Fungi; 290 genes, Vertebrates; 2586 genes.

We identified an intact rRNA sequence of 7440 bp that had 100% sequence identity to the 28S (KT384412.1) and 18S (KT38411.1) sequences in Genbank, and to the 5.8S region (Figure 3). tRNAscan found 63 tRNA transcripts within the *M. musculinus* genome, whereas the *S. scabiei* genome contained 66. RepeatMasker determined a repeat content of 11.0%, primarily simple and low complexity repeats, in our mite metagenome, whereas *S. scabiei* had 6.34% repeat content. MISA predicted 139,090 simple sequence repeats, a relatively high level of 2094 per megabase, a similar level to that in *S. scabiei* and *D. farina* (more than 2500 per megabase) and much higher than in other mites that have been examined (fewer than 500 per megabase).[46]

Proteomic comparison to other mite genomes. We compared the predicted proteome of *M. musculinus* with those of several closely related mites to identify candidate *Myocoptes*-specific genes. This is of interest in terms of developing new tools for the identification of *M. musculinus* in research and clinical laboratory settings. We used the OrthoVenn2 tool and input predicted proteomes from our metagenome; the closely related Astigmata mite genomes from *S. scabiei*, *D. farina*, *D. pteronyssinus*; and 2 outgroups, *I. scapulus* and *T. urticae* to identify candidate genes conserved within subgroups of these mites and unique to each organism (Figure 4). This process resulted in the identification of 215 clusters of gene families that are potentially unique to the *Myocoptes* lineage. A total of 130 of the *M. musculinus* proteins within these clusters matched at least one protein in the UniRef90 database with an E value of 1e-05 or lower. Supplementary Table S3 contains a representative *M. musculinus* gene and UniRef annotation from each cluster.

Identification of candidate allergen homologs in M. musculinus. Two of the mites most closely related to *M. musculinus*—*D. pteronyssinus* and *D. farinae*—are also 2 of the most common dust mites known to cause human allergic reactions.[36] We therefore examined whether our metagenome contained any candidate allergen genes. We determined candidate allergen genes from *S. scabiei* by using as a query the set of official *D. pteronyssinus* allergen genes (allergen.org; Table 2) We also found that our *M. musculinus* metagenome has a complement of allergens that is very similar to that of *S. scabiei*: of the 20 potential allergens identified in *M. musculinus* (Supplementary Table S4), 16 are also present in *S. scabiei*. A recent *S. scabiei* genome report identified 22 allergen genes in *S. scabiei* that were not present in other closely related mites.[27] Using BLAST at a threshold of 50% sequence identity, we found only one of these genes (KPM11750.1) in our *M. musculinus* predicted proteome. Using the same threshold, we found all but 3 (AY333085.1, AY333080.1, AY333073.1) of these *S. scabiei*-specific allergens in the *S. scabiei* protein set of our proteomic comparison.

Discussion

Several lines of evidence suggest that our metagenome for the mite *M. musculinus* has a high level of completeness. Our phylogenetic analysis places the *M. musculinus* metagenome clearly in the Astigmata suborder. Our analysis was not meant to resolve the difficult questions inherent to mite classification that we mentioned in the Introduction, because it was limited to a small number of genes in the few mite genomes that were

http://prime-pdf-watermark.prime-prod.pubfactory.com/ | 2025-02-25

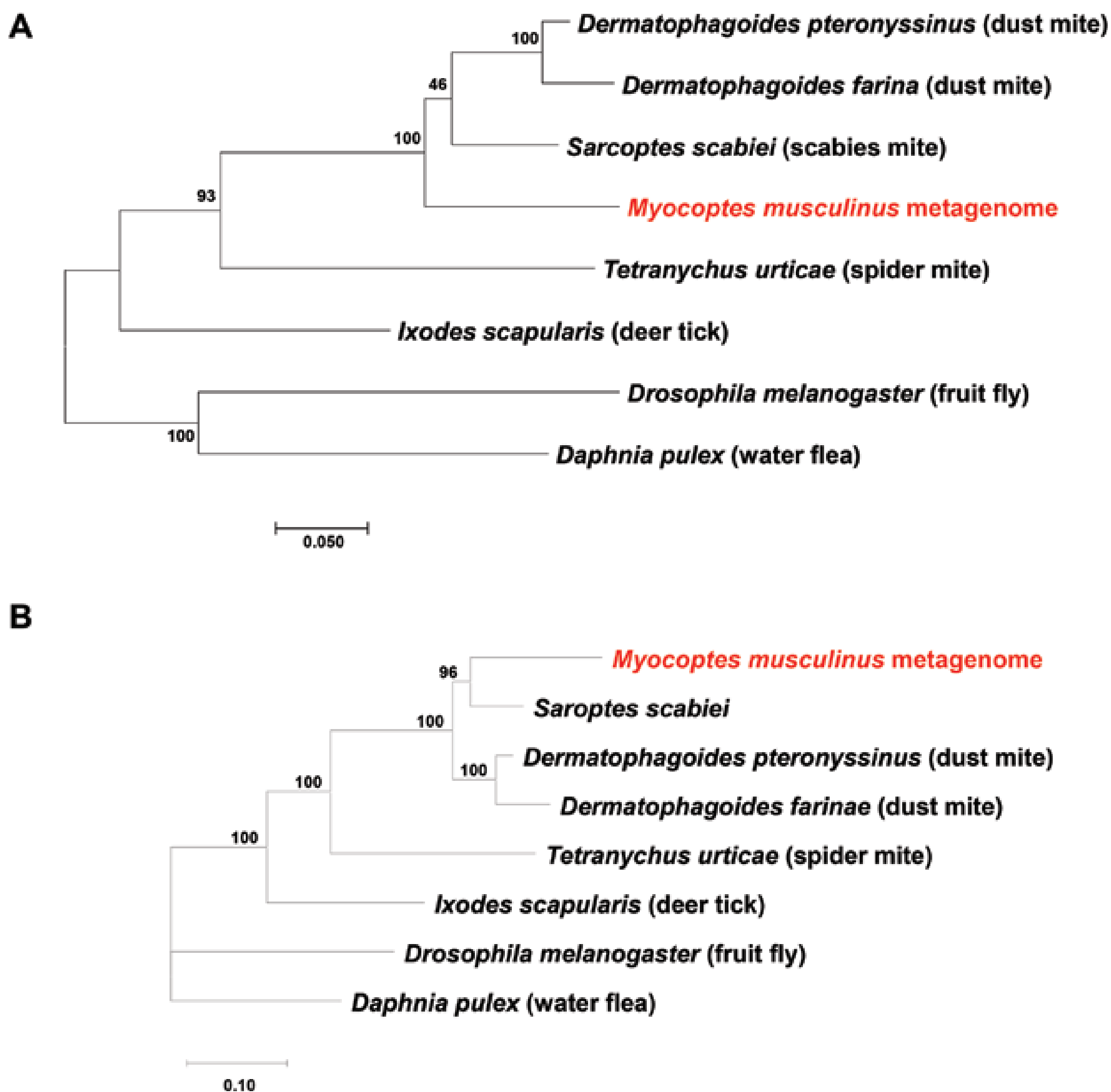

Figure 1. (A) The multiple-sequence alignment of the concatenated ribosomal proteins was input into MEGA7, and a maximum likelihood tree was constructed. The values at the nodes are bootstrapping values. Branch lengths are measured as the number of substitutions per site. (B) Using MRBAYES yielded a nearly identical tree. The posterior probabilities are indicated at the nodes. Branch lengths are measured as the number of substitutions per site.

available to us. Our analysis was designed only to help confirm the identity of our metagenome as *M. musculinus*.

The BUSCO analysis showed that the level of arthropod-specific genes in our metagenome is comparable to the genomes of mites derived from pure cultures. We recovered a nearly intact mitochondrial genome as a single contig, with a gene complement very similar to the mitochondrial genomes of related mites (Figure 2) and near-perfect identity to *M. musculinus* mitochondrial sequences in Genbank. In addition, the rDNA repeat unit is represented as a single contig of 7740 bp containing the 5S, 18S, and 28S genes with the ITS regions (Figure 3). Given these results, we conclude that we have assembled a metagenome for *M. musculinus*.

As noted in the Introduction, PCR assays based on sequences within the rDNA repeat may not always have specificity to the species level. Genomic resources for any species of *Myocoptes* are scarce in publicly available collections. The immediate value of a complete genomic sequence likely is that it enables the development of new diagnostic PCR tests for *M. musculinus* contamination that target other regions of the genome that may truly be species-specific. One recent example is the development of PCR tests for the SARS-CoV-2 virus: within 1 mo of the release of the genomic sequence to Genbank, the CDC developed tests to distinguish SARS-CoV-2 from other closely related coronaviruses[7,34]. Ultimately, the best molecular diagnostic tool for the detection of infectious organisms may be their own genomes. Using next-generation sequencing to identify infectious agents is already possible for the bacterial realm.[17,40] Published mite genomes are only 10 to 20 times larger than bacterial genomes; our metagenome is 66 Mb compared with 5+ Mb of most pathogenic *E. coli* strains, so as sequencing costs drop, the type of sequencing we describe here may become cost effective for use as a general tool. The availability of genomes of other murine ectoparasites would help to fulfill

http://prime-pdf-watermark.prime-prod.pubfactory.com/ | 2025-02-25

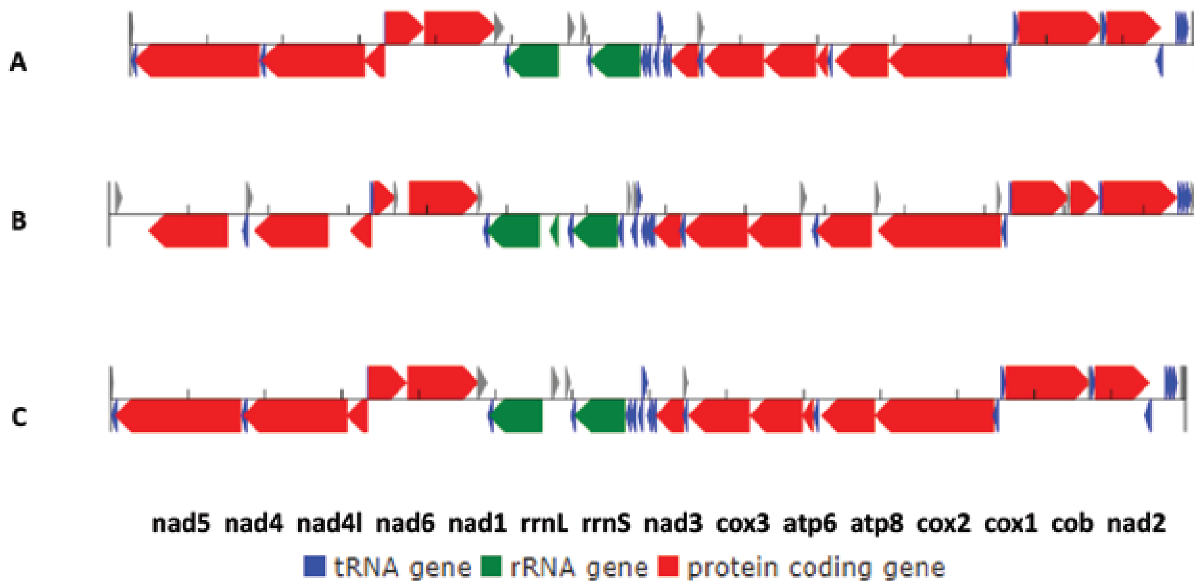

Figure 2. The mitochondrial (mt) genomes of *M. musculinus*, *S. scabiei*, and *D. pteronyssinus* were analyzed by using MITOS.[13] (A) Graphical representation of the gene content of the M. musculinus mt genome (13,962 bp). (B) Gene content of the *S. scabiei* mt genome (13,667 bp). (C) Gene content of the *D. pteronyssinus* mt genome (13,962 bp). In each case, the identities of the protein-coding genes and rRNA genes and color coding of the features are shown. Supplementary Table S2 details the complete gene content in tabular form.

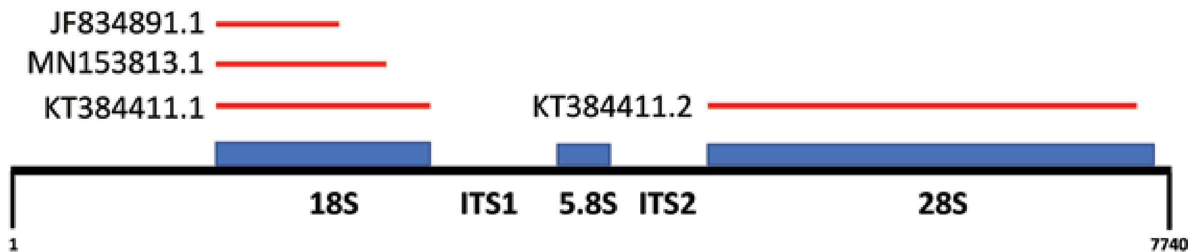

Figure 3. The positions of the major rRNA products of the rDNA repeat are noted in blue. Above the figure (in red) are the positions of known rRNA sequences from *M. musculinus* sequences currently available in Genbank.

the potential of these approaches. This availability would be valuable for molecular diagnostics that are based on PCR technology, because it would support the in-silico design of PCR primers with greater specificity and the ability to distinguish between ectoparasite genomes.

A complete genome is valuable for improving the design of PCR assays based on elements of the rDNA repeat,. Given that rDNA is a repeated feature in most eukaryotic genomes, sequence reads of rDNA are over-represented in any sequencing project. This situation means that the assembly of an intact rDNA repeat requires less sequencing effort than would a full organismal genome. Therefore even genomic survey sequencing (i.e., lower depth of reads than are necessary for a draft genome) of related *Myocoptes* species or of related murine ectoparasites likely would be sufficient for the assembly of an intact or nearly intact rDNA repeat, such as we have done with a more complete genome. The utility of this resource would be in the design of more species-specific primers for diagnostic PCR assays.[57] For example, the comparison of our 7740-bp rRNA element with the 4 rDNA sequences from *M. musculinus* available in Genbank (Figure 3) reveals that we have considerably extended the search space for the development of new PCR primers to include ITS1, ITS2, and the 5.8S rRNA.

Our proteomic comparison identifying *M. musculinus*-specific genes could be a good starting point for the development of improved molecular diagnostics. The collection of 215 gene families defined as unique to *M. musculinus* is dependent on the proteomes with which ours was compared. The collection does not distinguish between the various *Myocoptes* species, for example. Any of these genes could be found in multiple *Myocoptes* spp., or in all. Better genomic characterization of genomes of other *Myocoptes* spp., and other related murine ectoparasites would allow us to refine this set of genes into to a smaller set of truly species-specific *M. musculinus* genes that could be most valuable for a diagnostic PCR assay. The 85 gene families with no annotation might represent the best candidates for PCR assay development as some subset of these could actually represent mite-specific genes.

A complete genome has additional biologic value in and of itself. For example, we studied the allergen-related gene component of our metagenome, which is very similar to the complement of allergen-related genes of *S. scabiei*, which is the most closely related mite to *M. musculinus* that is not a house-dust mite responsible for common allergies. In house-dust mites, the group 1 and 2 allergens have the highest rates of sensitization[36] and the broadest level of IgE reactivity in hu-

http://prime-pdf-watermark.prime-prod.pubfactory.com/ | 2025-02-25

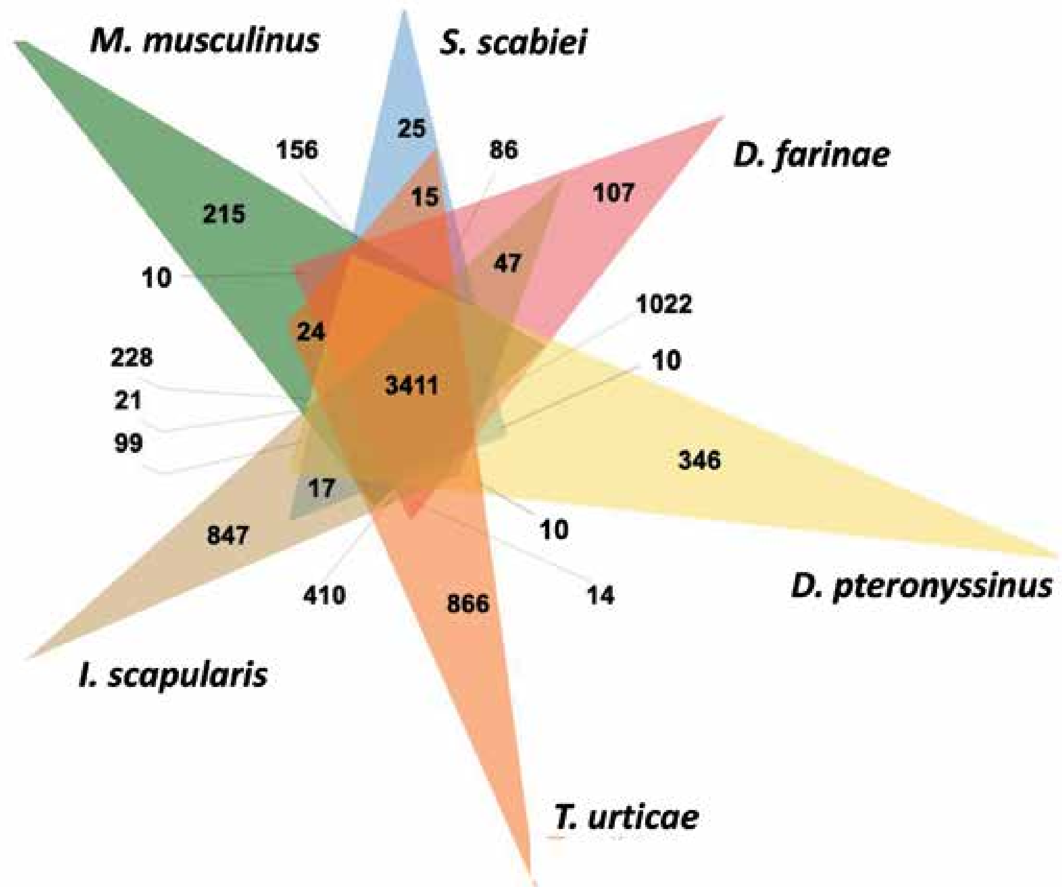

Figure 4. The predicted proteomes of our *M. musculinus* metagenome, along with *D. farinae*, *D. pteronnysinus*, *S. scabeie*, *T. urticus*, and *I. scapularis*, were input into Orthovenn2. The numbers indicate the number of conserved gene families within each region of the diagram.

Table 2. Candidate allergens in *M. musculinus*

Allergen ID	Function	*M. musculinus* ID	%ID	*S. scabiei* ID	%ID
Der p 3	Trypsin	MUSC_06977	50.57		
Der p 4	Alpha-amylase	MUSC_03488	72.43		
Der p 8	Glutathione S-transferase	MUSC_07635	66.36	KPM11587.1	65.44
Der p 10	Tropomyosin	MUSC_05430	96.3	KPM09025.1	97.55
Der p 11	Paramyosin	MUSC_01749	94.88	KPM04483.1	94.43
Der p 13	Cytosolic fatty acid binding protein	MUSC_08973	90.08	KPM07763.1	89.84
Der p 14	Apolipophorin			KPM11048.1	58.97
Der p 15	Chitinase-like protein	MUSC_03198	80.65		
Der p 18	Chitin-binding protein	MUSC_04343	62.19		
Der p 20	Arginine kinase	MUSC_02266	91.57	KPM07362.1	90.17
Der p 24	Ubiquinol-cytochrome c reductase binding protein	MUSC_09186	71.19	KPM06595.1	93.22
Der p 25	Triosphosphate isomerase	MUSC_06740	73.45	KPM10468.1	80.57
Der p 26	Myosin light chain	MUSC_08521	93.12	KPM03769.1	98.75
Der p 28	Heat shock protein 70	MUSC_02578	95.92	KPM03927.1	80.1
Der p 29	Cyclophilin	MUSC_07139	77.56	KPM10308.1	85
Der p 30	Ferritin	MUSC_11869	86.63	KPM04725.1	83.24
Der p 31	Cofilin	MUSC_08732	97.96	KPM08623.1	97.3
Der p 32	inorganic pyrophosphatase	MUSC_04908	67.01	KPM05552.1	68.37
Der p 33	α-tubulin	MUSC_04425	84.9	KPM02536.1	82.61
Der p 39	Troponin C	MUSC_08614	98.04	KPM08126.1	98.04
Der p 40	Thioredoxin like protein	MUSC_01637	69.52	KPM09467.1	70.87

The list of allergens from *D. pteronyssinus* and their functions, if known, and the gene ID and percentage identity of their closest homologs in *M. musculinus* and *S. scabiei*. A threshold of 50% sequence identity to the known *D. pteronyssinus* allergen protein was used as a cutoff.

http://prime-pdf-watermark.prime-prod.pubfactory.com/ | 2025-02-25

man patients. Both our metagenome and the *S. scabiei* genome appear to lack these 2 major allergen proteins, which could account in part for their lack of allergenicity in humans. In addition, our *M. musculinus* metagenome lacks most of the *S. scabiei*-specific allergen genes, suggesting that the allergenic potential for *M. musculinus* in humans would be even lower than that of *S. scabiei*. However, *M. musculinus* infestations can cause severe dermatitis in mice, indicating allergens to which rodents are likely sensitized.

Supplementary Materials

Table S1. Predicted *M. musculinus* proteins and UniRef annotation.

Table S2. Acariformes mitochondrial genome annotation.

Table S3. *M. musculinus* specific proteins.

Table S4. Amino acid sequence of candidate allergen homologs in *M. musculinus*.

Acknowledgments

This research was supported in part by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences. We thank Gregory Solomon and the NIEHS Sequencing Core, as well as the staff of the NIEHS Quality Assurance Laboratory. Finally, we thank Dr James Ward and Dr C Tyler Long for the review of this manuscript.

References

1. **Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, Scherer SE, Li PW, Hoskins RA, Galle RF, George RA, Lewis SE, Richards S, Ashburner M, Henderson SN, Sutton GG, Wortman JR, Yandell MD, Zhang Q, Chen LX, Brandon RC, Rogers YH, Blazej RG, Champe M, Pfeiffer BD, Wan KH, Doyle C, Baxter EG, Helt G, Nelson CR, Gabor GL, Abril JF, Agbayani A, An HJ, Andrews-Pfannkoch C, Baldwin D, Ballew RM, Basu A, Baxendale J, Bayraktaroglu L, Beasley EM, Beeson KY, Benos PV, Berman BP, Bhandari D, Bolshakov S, Borkova D, Botchan MR, Bouck J, Brokstein P, Brottier P, Burtis KC, Busam DA, Butler H, Cadieu E, Center A, Chandra I, Cherry JM, Cawley S, Dahlke C, Davenport LB, Davies P, de Pablos B, Delcher A, Deng Z, Mays AD, Dew I, Dietz SM, Dodson K, Doup LE, Downes M, Dugan-Rocha S, Dunkov BC, Dunn P, Durbin KJ, Evangelista CC, Ferraz C, Ferriera S, Fleischmann W, Fosler C, Gabrielian AE, Garg NS, Gelbart WM, Glasser K, Glodek A, Gong F, Gorrell JH, Gu Z, Guan P, Harris M, Harris NL, Harvey D, Heiman TJ, Hernandez JR, Houck J, Hostin D, Houston KA, Howland TJ, Wei MH, Ibegwam C, Jalali M, Kalush F, Karpen GH, Ke Z, Kennison JA, Ketchum KA, Kimmel BE, Kodira CD, Kraft C, Kravitz S, Kulp D, Lai Z, Lasko P, Lei Y, Levitsky AA, Li J, Li Z, Liang Y, Lin X, Liu X, Mattei B, McIntosh TC, McLeod MP, McPherson D, Merkulov G, Milshina NV, Mobarry C, Morris J, Moshrefi A, Mount SM, Moy M, Murphy B, Murphy L, Muzny DM, Nelson DL, Nelson DR, Nelson KA, Nixon K, Nusskern DR, Pacleb JM, Palazzolo M, Pittman GS, Pan S, Pollard J, Puri V, Reese MG, Reinert K, Remington K, Saunders RD, Scheeler F, Shen H, Shue BC, Sidén-Kiamos I, Simpson M, Skupski MP, Smith T, Spier E, Spradling AC, Stapleton M, Strong R, Sun E, Svirskas R, Tector C, Turner R, Venter E, Wang AH, Wang X, Wang ZY, Wassarman DA, Weinstock GM, Weissenbach J, Williams SM, Woodage T, Worley KC, Wu D, Yang S, Yao QA, Ye J, Yeh RF, Zaveri JS, Zhan M, Zhang G, Zhao Q, Zheng L, Zheng XH, Zhong FN, Zhong W, Zhou X, Zhu S, Zhu X, Smith HO, Gibbs RA, Myers EW, Rubin GM, Venter JC**. 2000. The genome sequence of *Drosophila melanogaster*. Science **287**:2185–2195. https://doi.org/10.1126/science.287.5461.2185.
2. **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ**. 1990. Basic local alignment search tool. J Mol Biol **215**:403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.
3. **Arlian LG, Morgan MS, Rider SD**. 2016. *Sarcoptes scabiei*: Genomics to proteomics to biology. Parasit Vectors **9**:380. https://doi.org/10.1186/s13071-016-1663-6.
4. **Baker DG**. 2006. Chapter 23 - Arthropods, p 565–579. In: **Fox JG, Davisson MT, Quimby FW, Barthold SW, Newcomer CE, Smith AL**, editors. The mouse in biomedical research (2nd edition). Burlington (VT): Academic Press.
5. **Baker DG**. 2007. Parasites of rats and mice, p 303–397. Flynn's parasites of laboratory animals. Ames (IA): Blackwell Publishing.
6. **Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA**. 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol **19**:455–477. https://doi.org/10.1089/cmb.2012.0021.
7. Centers for Disease Control and Prevention (CDC). [Internet]. 2020. CDC's Influenza SARS-CoV-2 Multiplex Assay. [Cited 18 November 2022]. Available at: https://www.cdc.gov/coronavirus/2019-ncov/lab/multiplex.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fcoronavirus%2F2019-ncov%2Flab%2Ftesting.html.
8. **Chan TF, Ji KM, Yim AK, Liu XY, Zhou JW, Li RQ, Yang KY, Li J, Li M, Law PT, Wu YL, Cai ZL, Qin H, Bao Y, Leung RK, Ng PK, Zou J, Zhong XJ, Ran PX, Zhong NS, Liu ZG, Tsui SK**. 2015. The draft genome, transcriptome, and microbiome of *Dermatophagoides farinae* reveal a broad spectrum of dust mite allergens. J Allergy Clin Immunol **135**:539–548. https://doi.org/10.1016/j.jaci.2014.09.031.
9. **Colbourne JK, Pfrender ME, Gilbert D, Thomas WK, Tucker A, Oakley TH, Tokishita S, Aerts A, Arnold GJ, Basu MK, Bauer DJ, Cáceres CE, Carmel L, Casola C, Choi JH, Detter JC, Dong Q, Dusheyko S, Eads BD, Fröhlich T, Geiler-Samerotte KA, Gerlach D, Hatcher P, Jogdeo S, Krijgsveld J, Kriventseva EV, Kültz D, Laforsch C, Lindquist E, Lopez J, Manak JR, Muller J, Pangilinan J, Patwardhan RP, Pitluck S, Pritham EJ, Rechtsteiner A, Rho M, Rogozin IB, Sakarya O, Salamov A, Schaack S, Shapiro H, Shiga Y, Skalitzky C, Smith Z, Souvorov A, Sung W, Tang Z, Tsuchiya D, Tu H, Vos H, Wang M, Wolf YI, Yamagata H, Yamada T, Ye Y, Shaw JR, Andrews J, Crease TJ, Tang H, Lucas SM, Robertson HM, Bork P, Koonin EV, Zdobnov EM, Grigoriev IV, Lynch M, Boore JL**. 2011. The ecoresponsive genome of *Daphnia pulex*. Science **331**:555–561. https://doi.org/10.1126/science.1197761.
10. **Consortium TU**. 2021. UniProt: The universal protein knowledgebase in 2021. Nucleic Acids Res **49** D1:D480–D489. https://doi.org/10.1093/nar/gkaa1100.
11. **Dabert M, Witalinski W, Kazmierski A, Olszanowski Z, Dabert J**. 2010. Molecular phylogeny of acariform mites (Acari, Arachnida): Strong conflict between phylogenetic signal and long-branch attraction artifacts. Mol Phylogenet Evol **56**:222–241. https://doi.org/10.1016/j.ympev.2009.12.020.
12. **Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR**. 2013. STAR: Ultrafast universal RNA-seq aligner. Bioinformatics **29**:15–21. https://doi.org/10.1093/bioinformatics/bts635.
13. **Donath A, Jühling F, Al-Arab M, Bernhart SH, Reinhardt F, Stadler PF, Middendorf M, Bernt M**. 2019. Improved annotation of protein-coding genes boundaries in metazoan mitochondrial genomes. Nucleic Acids Res **47**:10543–10552. https://doi.org/10.1093/nar/gkz833.
14. **Edgar RC**. 2004. MUSCLE: A multiple-sequence alignment method with reduced time and space complexity. BMC Bioinformatics **5**:113. https://doi.org/10.1186/1471-2105-5-113.
15. **Glowska E, Dragun-Damian A, Dabert J**. 2013. DNA-barcoding contradicts morphology in quill mite species *Torotrogla merulae* and *T. rubeculi* (Prostigmata: Syringophilidae). Folia Parasitol (Praha) **60**:51–60. https://doi.org/10.14411/fp.2013.007.
16. **Goodman RE, Ebisawa M, Ferreira F, Sampson HA, van Ree R, Vieths S, Baumert JL, Bohle B, Lalithambika S, Wise J, Taylor SL**. 2016. AllergenOnline: A peer-reviewed, curated allergen database to assess novel food proteins for potential cross-reactivity. Mol Nutr Food Res **60**:1183–1198. https://doi.org/10.1002/mnfr.201500769.

http://prime-pdf-watermark.prime-prod.pubfactory.com/ | 2025-02-25

17. **Grad YH, Lipsitch M, Feldgarden M, Arachchi HM, Cerqueira GC, Fitzgerald M, Godfrey P, Haas BJ, Murphy CI, Russ C, Sykes S, Walker BJ, Wortman JR, Young S, Zeng Q, Abouelleil A, Bochicchio J, Chauvin S, Desmet T, Gujja S, McCowan C, Montmayeur A, Steelman S, Frimodt-Møller J, Petersen AM, Struve C, Krogfelt KA, Bingen E, Weill FX, Lander ES, Nusbaum C, Birren BW, Hung DT, Hanage WP**. 2012. Genomic epidemiology of the *Escherichia coli* O104:H4 outbreaks in Europe, 2011. Proc Natl Acad Sci USA **109**:3065–3070. https://doi.org/10.1073/pnas.1121491109.
18. **Grbić M, Van Leeuwen T, Clark RM, Rombauts S, Rouze P, Grbic V, Osborne EJ, Dermauw W, Ngoc PC, Ortego F, Hernandez-Crespo P, Diaz I, Martinez M, Navajas M, Sucena E, Magalhaes S, Nagy L, Pace RM, Djuranovic S, Smagghe G, Iga M, Christiaens O, Veenstra JA, Ewer J, Villalobos RM, Hutter JL, Hudson SD, Velez M, Yi SV, Zeng J, Pires-daSilva A, Roch F, Cazaux M, Navarro M, Zhurov V, Acevedo G, Bjelica A, Fawcett JA, Bonnet E, Martens C, Baele G, Wissler L, Sanchez-Rodriguez A, Tirry L, Blais C, Demeestere K, Henz SR, Gregory TR, Mathieu J, Verdon L, Farinelli L, Schmutz J, Lindquist E, Feyereisen R, Van de Peer Y**. 2011. The genome of *Tetranychus urticae* reveals herbivorous pest adaptations. Nature **479**:487–492. https://doi.org/10.1038/nature10640.
19. **Gulia-Nuss M, Nuss AB, Meyer JM, Sonenshine DE, Roe RM, Waterhouse RM, Sattelle DB, de la Fuente J, Ribeiro JM, Megy K, Thimmapuram J, Miller JR, Walenz BP, Koren S, Hostetler JB, Thiagarajan M, Joardar VS, Hannick LI, Bidwell S, Hammond MP, Young S, Zeng QD, Abrudan JL, Almeida FC, Ayllon N, Bhide K, Bissinger BW, Bonzon-Kulichenko E, Buckingham SD, Caffrey DR, Caimano MJ, Croset V, Driscoll T, Gilbert D, Gillespie JJ, Giraldo-Calderon GI, Grabowski JM, Jiang D, Khalil SMS, Kim D, Kocan KM, Koci J, Kuhn RJ, Kurtti TJ, Lees K, Lang EG, Kennedy RC, Kwon H, Perera R, Qi YM, Radolf JD, Sakamoto JM, Sanchez-Gracia A, Severo MS, Silverman N, Simo L, Tojo M, Tornador C, Van Zee JP, Vazquez J, Vieira FG, Villar M, Wespiser AR, Yang YL, Zhu JW, Arensburger P, Pietrantonio PV, Barker SC, Shao RF, Zdobnov EM, Hauser F, Grimmelikhuijzen CJP, Park Y, Rozas J, Benton R, Pedra JHF, Nelson DR, Unger MF, Tubio JMC, Tu ZJ, Robertson HM, Shumway M, Sutton G, Wortman JR, Lawson D, Wikel SK, Nene VM, Fraser CM, Collins FH, Birren B, Nelson KE, Caler E, Hill CA**. 2016. Genomic insights into the *Ixodes scapularis* tick vector of Lyme disease. Nat Commun **7**:10507. https://doi.org/10.1038/ncomms10507.
20. **Gurevich A, Saveliev V, Vyahhi N, Tesler G**. 2013. QUAST: Quality assessment tool for genome assemblies. Bioinformatics **29**:1072–1075. https://doi.org/10.1093/bioinformatics/btt086.
21. **Huelsenbeck JP, Ronquist F**. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics **17**:754–755. https://doi.org/10.1093/bioinformatics/17.8.754.
22. **Jacoby RO, Fox JG**. 1984. Biology and diseases of mice, p 31–89. In: **Fox JG, Cohen BJ, Loew FM**, editors. Laboratory animal medicine. Orlando (FL): Academic Press. https://doi.org/10.1016/B978-0-12-263620-2.50009-4.
23. **Jungmann P, Freitas A, Bandeira A, Nobrega A, Coutinho A, Marcos MA, Minoprio P**. 1996. Murine acariasis. II. Immunological dysfunction and evidence for chronic activation of Th2 lymphocytes. Scand J Immunol **43**:604–612. https://doi.org/10.1046/j.1365-3083.1996.d01-259.x.
24. **Karlsson EM, Pearson LM, Kuzma KM, Burkholder TH**. 2014. Combined evaluation of commonly used techniques, including PCR, for diagnosis of mouse fur mites. J Am Assoc Lab Anim Sci **53**:69–73.
25. **Katoh K, Kuma K, Toh H, Miyata T**. 2005. MAFFT version 5: Improvement in accuracy of multiple sequence alignment. Nucleic Acids Res **33**:511–518. https://doi.org/10.1093/nar/gki198.
26. **Korf I**. 2004. Gene finding in novel genomes. BMC Bioinformatics **5**:59. https://doi.org/10.1186/1471-2105-5-59.
27. **Korhonen PK, Gasser RB, Ma G, Wang T, Stroehlein AJ, Young ND, Ang C-S, Fernando DD, Lu HC, Taylor S, Reynolds SL, Mofiz E, Najaraj SH, Gowda H, Madugundu A, Renuse S, Holt D, Pandey A, Papenfuss AT, Fischer K**. 2020. High-quality nuclear genome for *Sarcoptes scabiei*—a critical resource for a neglected parasite. PLoS Negl Trop Dis **14**:e0008720. https://doi.org/10.1371/journal.pntd.0008720.
28. **Kumar S, Stecher G, Tamura K**. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol Biol Evol **33**:1870–1874. https://doi.org/10.1093/molbev/msw054.
29. **Lagesen K, Hallin P, Rødland EA, Staerfeldt HH, Rognes T, Ussery DW**. 2007. RNAmmer: Consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res **35**:3100–3108. https://doi.org/10.1093/nar/gkm160.
30. **Li R, Zhu H, Ruan J, Qian W, Fang X, Shi Z, Li Y, Li S, Shan G, Kristiansen K, Li S, Yang H, Wang J, Wang J**. 2010. De novo assembly of human genomes with massively parallel short read sequencing. Genome Res **20**:265–272. https://doi.org/10.1101/gr.097261.109.
31. **Lilue J, Doran AG, Fiddes IT, Abrudan M, Armstrong J, Bennett R, Chow W, Collins J, Collins S, Czechanski A, Danecek P, Diekhans M, Dolle DD, Dunn M, Durbin R, Earl D, Ferguson-Smith A, Flicek P, Flint J, Frankish A, Fu B, Gerstein M, Gilbert J, Goodstadt L, Harrow J, Howe K, Ibarra-Soria X, Kolmogorov M, Lelliott CJ, Logan DW, Loveland J, Mathews CE, Mott R, Muir P, Nachtweide S, Navarro FCP, Odom DT, Park N, Pelan S, Pham SK, Quail M, Reinholdt L, Romoth L, Shirley L, Sisu C, Sjoberg-Herrera M, Stanke M, Steward C, Thomas M, Threadgold G, Thybert D, Torrance J, Wong K, Wood J, Yalcin B, Yang F, Adams DJ, Paten B, Keane TM**. 2018. Sixteen diverse laboratory mouse reference genomes define strain-specific haplotypes and novel functional loci. Nat Genet **50**:1574–1583. https://doi.org/10.1038/s41588-018-0223-8.
32. **Livingston RS, Riley LK**. 2003. Diagnostic testing of mouse and rat colonies for infectious agents. Lab Anim (NY) **32**:44–51. https://doi.org/10.1038/laban0503-44.
33. **Lowe TM, Eddy SR**. 1997. tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res **25**:955–964. https://doi.org/10.1093/nar/25.5.955.
34. **Lu X, Wang L, Sakthivel SK, Whitaker B, Murray J, Kamili S, Lynch B, Malapati L, Burke SA, Harcourt J, Tamin A, Thornburg NJ, Villanueva JM, Lindstrom S**. 2020. US CDC real-time reverse transcription PCR panel for detection of Severe Acute Respiratory Syndrome Coronavirus 2. Emerg Infect Dis **26**:1654–1665. https://doi.org/10.3201/eid2608.201246.
35. **Merola VM, Eubig PA**. 2012. Toxicology of avermectins and milbemycins (macrocylic lactones) and the role of P-glycoprotein in dogs and cats. Vet Clin North Am Small Anim Pract **42**:313–333. https://doi.org/10.1016/j.cvsm.2011.12.005.
36. **Miller JD**. 2019. The role of dust mites in allergy. Clin Rev Allergy Immunol **57**:312–329. https://doi.org/10.1007/s12016-018-8693-0.
37. **Mofiz E, Holt DC, Seemann T, Currie BJ, Fischer K, Papenfuss AT**. 2016. Genomic resources and draft assemblies of the human and porcine varieties of scabies mites, *Sarcoptes scabiei* var. *hominis* and var. *suis*. Gigascience **5**:23. https://doi.org/10.1186/s13742-016-0129-2.
38. **Morita E, Kaneko S, Hiragun T, Shindo H, Tanaka T, Furukawa T, Nobukiyo A, Yamamoto S**. 1999. Fur mites induce dermatitis associated with IgE hyperproduction in an inbred strain of mice, NC/Kuj. J Dermatol Sci **19**:37–43. https://doi.org/10.1016/S0923-1811(98)00047-4.
39. Mouse Genomes Project. [Internet]. 2022. Sequencing read data. [Cited 18 November 2022]. Available at: https://www.mousegenomes.org/mouse-strains-sequenced/.
40. **Mustapha MM, Srinivasa VR, Griffith MP, Cho ST, Evans DR, Waggle K, Ezeonwuka C, Snyder DJ, Marsh JW, Harrison LH, Cooper VS, Van Tyne D**. 2022. genomic diversity of hospital-acquired infections revealed through prospective whole-genome sequencing-based surveillance. mSystems **7**:e0138421. https://doi.org/10.1128/msystems.01384-21.
41. **Pagel Van Zee JP, Geraci NS, Guerrero FD, Wikel SK, Stuart JJ, Nene VM, Hill CA**. 2007. Tick genomics: The Ixodes genome project and beyond. Int J Parasitol **37**:1297–1305. https://doi.org/10.1016/j.ijpara.2007.05.011.

http://prime-pdf-watermark.prime-prod.pubfactory.com/ | 2025-02-25

42. **Pochanke V, Hatak S, Hengartner H, Zinkernagel RM, McCoy KD**. 2006. Induction of IgE and allergic-type responses in fur mite-infested mice. Eur J Immunol **36**:2434–2445. https://doi.org/10.1002/eji.200635949.
43. **Pomés A, Davies JM, Gadermaier G, Hilger C, Holzhauser T, Lidholm J, Lopata AL, Mueller GA, Nandy A, Radauer C, Chan SK, Jappe U, Kleine-Tebbe J, Thomas WR, Chapman MD, van Hage M, van Ree R, Vieths S, Raulf M, Goodman RE**. 2018. WHO/IUIS allergen nomenclature: Providing a common language. Mol Immunol **100**:3–13. https://doi.org/10.1016/j.molimm.2018.03.003.
44. **Randall TA, Mullikin JC, Mueller GA**. 2018. The draft genome assembly of *Dermatophagoides pteronyssinus* supports identification of novel allergen isoforms in *Dermatophagoides* species. Int Arch Allergy Immunol **175**:136–146. https://doi.org/10.1159/000481989.
45. **Rice KA, Albacarys LK, Metcalf Pate KA, Perkins C, Henderson KS, Watson J**. 2013. Evaluation of diagnostic methods for *Myocoptes musculinus* according to age and treatment status of mice (*Mus musculus*). J Am Assoc Lab Anim Sci **52**:773–781.
46. **Rider SD, Morgan MS, Arlian LG**. 2015. Draft genome of the scabies mite. Parasit Vectors **8**:585. https://doi.org/10.1186/s13071-015-1198-2.
47. **Roble GS, Boteler W, Riedel E, Lipman NS**. 2012. Total IgE as a serodiagnostic marker to aid murine fur mite detection. J Am Assoc Lab Anim Sci **51**:199–208.
48. **Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM**. 2015. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics **31**:3210–3212. https://doi.org/10.1093/bioinformatics/btv351.
49. **Smit AFA, Hubley R, Green P**. [Internet]. RepeatMasker. [Cited 24 October 2016]. Available at: http://www.repeatmasker.org.
50. **Stanke M, Waack S**. 2003. Gene prediction with a hidden Markov model and a new intron submodel. Bioinformatics **19** Suppl 2: ii215–ii225. https://doi.org/10.1093/bioinformatics/btg1080.
51. **Sugiura K, Sugiura M, Hayakawa R, Shamoto M, Takahashi H**. 2002. Study of the patch test reactions to *Myocoptes musculinus* (Kekuidani) of NC/F mice (atopic model mice). Exogenous Dermatology **1**:87–91. https://doi.org/10.1159/000058338.
52. **Thiel T, Michalek W, Varshney RK, Graner A**. 2003. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). Theor Appl Genet **106**:411–422. https://doi.org/10.1007/s00122-002-1031-0.
53. **Van Dam MH, Trautwein M, Spicer GS, Esposito L**. 2019. Advancing mite phylogenomics: Designing ultraconserved elements for Acari phylogeny. Mol Ecol Resour **19**:465–475. https://doi.org/10.1111/1755-0998.12962.
54. **Wang Y, Coleman-Derr D, Chen GP, Gu YQ**. 2015. OrthoVenn: A web server for genome wide comparison and annotation of orthologous clusters across multiple species. Nucleic Acids Res **43** W1:W78–W84. https://doi.org/10.1093/nar/gkv487.
55. **West PT, Probst AJ, Grigoriev IV, Thomas BC, Banfield JF**. 2018. Genome reconstruction for eukaryotes from complex natural microbial communities. Genome Res **28**:569–580. https://doi.org/10.1101/gr.228429.117.
56. **Zhang YX, Chen X, Wang JP, Zhang ZQ, Wei H, Yu HY, Zheng HK, Chen Y, Zhang LS, Lin JZ, Sun L, Liu DY, Tang J, Lei Y, Li XM, Liu M**. 2019. Genomic insights into mite phylogeny, fitness, development, and reproduction. BMC Genomics **20**:954. https://doi.org/10.1186/s12864-019-6281-1.
57. **Zhao Y, Zhang W-Y, Wang R-L, Niu D-L**. 2020. Divergent domains of 28S ribosomal RNA gene: DNA barcodes for molecular classification and identification of mites. Parasit Vectors **13**:251. https://doi.org/10.1186/s13071-020-04124-z.

http://prime-pdf-watermark.prime-prod.pubfactory.com/ | 2025-02-25