J. Parasitol., 104(1), 2018, pp. 89–95

Development and Characterization of 12 Novel Polymorphic Microsatellite Loci for the Mammal Chewing Louse *Geomydoecus aurei* (Insecta: Phthiraptera) and a Comparison of Next-Generation Sequencing Approaches for Use in Parasitology

J. E. Light, S. E. Harper*, K. P. Johnson†, J. W. Demastes*, and T. A. Spradling*

Department of Wildlife and Fisheries Sciences, Texas A&M University, 534 John Kimbrough Blvd., College Station, Texas 77843. Correspondence should be sent to J. E. Light at: *jlight2@tamu.edu*

ABSTRACT: Next-generation sequencing methodologies open the door for evolutionary studies of wildlife parasites. We used 2 next-generation sequencing approaches to discover microsatellite loci in the pocket gopher chewing louse *Geomydoecus aurei* for use in population genetic studies. In one approach, we sequenced a library enriched for microsatellite loci; in the other approach, we mined microsatellites from genomic sequences. Following microsatellite discovery, promising loci were tested for amplification and polymorphism in 390 louse individuals from 13 pocket gopher hosts. In total, 12 loci were selected for analysis (6 from each methodology), none of which exhibited evidence of null alleles or heterozygote deficiencies. These 12 loci showed adequate genetic diversity for population-level analyses, with 3–9 alleles per locus with an average H_E per locus ranging from 0.32 to 0.70. Analysis of Molecular Variance (AMOVA) indicated that genetic variation among infrapopulations accounts for a low, but significant, percentage of the overall genetic variation, and individual louse infrapopulations showed F_{ST} values that were significantly different from zero in the majority of pairwise infrapopulation comparisons, despite all 13 infrapopulations being taken from the same locality. Therefore, these 12 polymorphic markers will be useful at the infrapopulation and population levels for future studies involving *G. aurei*. This study shows that next-generation sequencing methodologies can successfully be used to efficiently obtain data for a variety of evolutionary questions.

Molecular data from non-model organisms such as parasites can be efficiently and economically obtained via next-generation sequencing. Herein we use and compare 2 next-generation sequencing approaches (one that uses an enriched microsatellite library and one that uses the whole genome) to discover microsatellite loci in the chewing louse (Insecta: Phthiraptera: Ischnocera), *Geomyodecus aurei*. In the end, both approaches yielded similar numbers of useable microsatellite loci, even from a small-bodied parasite. Thus, next-generation sequences approaches are promising and can open new avenues of research that were previously considered impractical.

Parasites are often characterized by complex ecological interactions with their hosts that persist over long evolutionary timescales, such that understanding how and why these interactions originate and persist can yield valuable information about evolutionary processes. Pocket gophers (Rodentia: Geomyidae) and their chewing lice are model organisms in host-parasite studies. As such, multiple attempts have been made to understand the macroevolutionary patterns that drive host-parasite associations in pocket gophers and their chewing lice (Hafner and Nadler, 1988; Hafner et al., 1994). However, despite the role the pocket gopher-chewing louse assemblage plays as a model system in evolutionary biology, only a few studies to date have focused on interactions within or among populations of lice (e.g., Nadler et al., 1990; Demastes et al., 2012; Nessner et al., 2014; Harper et al., 2015), representing relatively little of the diversity of 122 species of chewing lice that inhabit pocket gophers. Although recent years have seen an increase in population genetic studies of parasites of wildlife (e.g., Detwiler and Criscione, 2011, 2017; Detwiler et al., 2017; Duffy et al., 2017; Jirsová et al., 2017; Opiro et al., 2017), even in lice (e.g., Johnson et al., 2002; Leo et al., 2005; Ascunce et al., 2013; Koop et al., 2014; Harper et al., 2015), generally few attempts have been made to explore host-parasite associations at the population level, despite these studies being of vital significance to the evolutionary and ecological dynamics of symbiotic systems (Thompson, 1994; Criscione et al., 2005; Huyse et al., 2005; McCoy, 2009; Mazé-Guilmo et al., 2016). The paucity of studies at the population level may be because of the perceived difficulty of obtaining informative markers in non-model organisms such as parasites. However, with the advent of next-generation sequencing technologies, multiple methods are available for obtaining molecular markers, even at the population scale, for non-model organisms lacking reference genomes (e.g., Peterson et al., 2012; Puritz et al., 2012; Riesgo et al., 2012). Herein we use and compare 2 different next-generation sequencing methodologies to obtain microsatellite loci for the chewing louse *Geomydoecus aurei*, a parasite of the Bottae's pocket gopher (*Thomomys bottae*), and examine these loci for their utility in a population genetic context.

Pocket gopher specimens were collected with approval of the New Mexico Department of Game and Fish using collection and care procedures that followed all guidelines set by the American Society of Mammalogists (Sikes et al., 2016) and that were approved by the University of Northern Iowa (UNI) Institutional Animal Care and Use Committee. Lice from each individual host (i.e., an infrapopulation) were placed in labeled Nunc CryoTube vials (Nalge Nunc International, Rochester, New York) and stored on dry ice or in liquid nitrogen until return to the laboratory, where they were stored at –80 C. Microsatellite regions of the *G. aurei* genome were detected and targeted using 2 methodologies: (1) utilizing services at Cornell University to enrich for microsatellite regions followed by 454-sequencing and

* Department of Biology, 144 McCollum Science Hall, University of Northern Iowa, Cedar Falls, Iowa 50614.

† Illinois Natural History Survey, 1614 South Oak Street, University of Illinois, Champaign, Illinois 61820.

DOI: 10.1645/17-130

(2) screening Illumina HiSeq–generated shot-gun genome (whole genome) sequences for repeats. In both cases, once microsatellite loci were identified and primers were designed for amplification, loci were tested for polymorphism either on lice from the same infrapopulation or on lice from multiple host individuals at the same geographic locality, with the rationale that polymorphic loci at these scales would likely be polymorphic at larger geographic scales.

Method 1 (sequencing an enriched library via 454-sequencing): At UNI, a DNeasy Blood & Tissue Kit (Qiagen, Valencia, California) was used to extract DNA from a pool of 209 *G. aurei* lice taken from 3 pocket gophers (n = 94, 72, and 43 individual lice from each host) from the same locality: New Mexico, Socorro County, 1.4 mile S, 0.8 mile W Las Nutrias (pocket gopher specimens are housed at the Biodiversity Research and Teaching Collections at Texas A&M University [BRTC], catalog numbers TCWC 64962, 64965, and 64968; TCWC is an older acronym for the BRTC collections, and all specimens retain this acronym). The manufacturer's recommendations were followed with the following exceptions: (1) lice were used directly from the ultra-cold freezer and use of liquid nitrogen was eliminated, (2) after 4 hr of incubation in ATL buffer, an additional 20 µl of Proteinase K was added with additional crushing performed before continued incubation overnight at 56 C, (3) final elution volume of AE Buffer was decreased from 200 µl to 50 µl, and (4) final elution incubation time was increased to 5 min.

At the Evolutionary Genetics Core Facility at Cornell University, pooled louse DNA was treated with the restriction enzyme Hinc II and then ligated to a double-stranded SNX linker. The ligation procedure was modified to generate *Pme I* sites if linkers ligated to themselves. Digested, ligated fragments were enriched for microsatellites by hybridization to 3′ biotinylated di-, tri-, and tetra-meric repeat probes. PCR amplified products were ligated to 1.0 µl of a Titanium Rapid Library MID adapter (10 µm adapter stock; Roche, Pleasanton, California), and Ampure beads were used to remove small fragments, resulting in the selection of fragments that were 400–900 bp in length. Libraries were submitted to the Sequencing and Genotyping Facility at Cornell Life Sciences Core Laboratory Center for FAM-quantification and Titanium 454-sequencing.

A total of 21,713 single end sequence reads (ranging from 56 to 1,201 nucleotides in length, with a preferred length of 700–800 bp) from analysis of 454-sequencing data were trimmed and then assembled into 8,699 contigs using SeqMan Pro software (Lasergene v 8.1.1, DNASTAR, Inc.). The Pro Assembler option was selected, and match size was set to 120, and minimum match percentage was set to 94 (values more stringent than the default settings). Contigs were examined using MSATCOMMANDER software (Rozen and Skaletsky, 2000; Faircloth, 2008) to identify contigs with repeats and determine primer sites and sequences for those contigs. In total, 4,375 potential loci (2,536 di-, 1,502 tri-, 322 tetra-, and 16 pentameric repeats) ranging in size from 5 to 39 repeats were identified. Of these, 2,223 potential loci had sufficient sequence on the 5′ and 3′ end of the repeat structure to facilitate accurate amplification and primer design (1,054 di-, 927 tri-, 236 tetra-, and 7 pentameric repeats). Only trimeric and tetrameric repeat motifs with 7 or more repeats (but not more than 12) were examined at UNI. During this examination, surrounding sequence was carefully inspected for additional repeat structures that could complicate amplification and genotyping of the repeat of interest. Remaining contigs were aligned in Geneious PRO version 5.4.6 (http://geneious.com, BioMatters Ltd, Auckland, New Zealand; Kearse et al., 2012) to verify each contig was unique. In the end, 23 tetrameric and 11 trimeric repeat loci were chosen for tests of amplification and polymorphism at UNI, representing only 1.5% of the 2,223 potentially useful contigs initially identified.

For initial tests of the amplification potential of primers designed for the 34 loci above, genomic DNA was extracted from individual lice using the DNeasy Blood & Tissue Kit (Qiagen) following the manufacturer's recommendations with the following exceptions: (1) prior to DNA extraction, individual louse bodies were punctured once through the abdomen with a no. 2 insect pin or placed on a freezer block under the dissecting microscope and punctured 6 times through the abdomen using a no. 2 insect pin, (2) carrier RNA was added to AL Buffer before addition of ethanol, and (3) incubation before elution in 30 µl of AE was increased to 5 min. Following DNA extraction, cleared louse bodies were prepared as voucher specimens on microscope slides. A 3-primer amplification protocol was used for PCR of the 34 promising loci, where a long M13 tag (5′-CGAGTTTTCC-CAGTCACGAC-3′) was added to the 5′ end of all locus-specific forward primers to allow concurrent amplification with a fluorescent primer (Schuelke, 2000). A short tag (5′-GTTTCTT-3′) was added to all locus-specific reverse primers to promote adenylation and reduce stutter (Brownstein et al., 1996). A fluorescent tag (6-FAM) was added to the 5′ end of universal M13 primers (5′-CGAGTTTTCCCAGTCACGAC-3′) to allow 3-primer amplification of PCR products and subsequent multiplex genotyping (Schuelke, 2000). Microsatellite loci were amplified in 10 µl reactions containing 1.0 µl DNA (approximately 0.4–2.0 ng) and 1 X GoTaq Clear Hot Start Master Mix (Promega, Madison, Wisconsin) and 1 µl primer mix in a 10 µl reaction. Each primer mix contained 0.5 µM each forward primer, 2 µM each reverse primer, and 4 µM fluorophore-labeled M13 primer per primer pair included. Thermal cycles were as follows for all samples: 1 cycle of denaturation for 2 min at 95 C, 30 cycles of denaturation at 94 C for 40 sec, annealing at 53 C for 40 sec, and elongation at 72 C for 40 sec, followed by a final extension at 72 C for 45 min. PCR products were screened for amplification and contamination on 1.2% agarose gels. Cleanly amplified products were sent to the Iowa State University DNA Facility for analysis on an Applied Biosystems 3730 DNA Analyzer (Applied Biosystems, Foster City, California). Output fsa files were scored and verified manually at least twice per specimen using GeneMarker software.

Method 2 (whole-genome sequencing [WGS] via Illumina HiSeq): DNA was extracted from a pool of 15 individuals of *G. aurei* taken from a single host individual of *Thomomys bottae* from the same locality listed above (TCWC 64963). DNA was sonicated with a Covaris M220 to an average size of about 400 nucleotides, and a DNAseq library was prepared with the Kapa Library Preparation kit (Kapa Biosystems, Boston, Massachusetts). The library was pooled in equimolar concentration, quantitated by qPCR, and the pool was sequenced on 1 lane of an Illumina HiSeq2500 platform for 161 cycles from each end of the fragments using a TruSeq SBS sequencing Rapid kit version 1 at the University of Illinois Keck Center for Comparative and Functional Genomics. Fastq files were generated with Casava 1.8.2. The process yielded 58,340,400 paired reads of 160 nucleotides in length (NCBI SRA accession no. SRR5308121).

Paired-end sequence data were error-trimmed using SolexaQA (Cox et al., 2010), and 123,371 contigs were assembled from an 80% subset of the data using the Velvet (Zerbino and Birney, 2008) plugin of Geneious R8.1.9 (Kearse et al., 2012). At Texas A&M (TAMU), the program MSATCOMMANDER (Rozen and Skaletsky, 2000; Faircloth, 2008) was used to screen contigs for all di-, tri-, tetra-, and pentameric repeat motifs (with a minimum of 6 repeats) with suitable priming sites and lacking additional repeat structures (as described above). In total, 6,005 potential loci were identified (4,105 di-, 1,831 tri-, 68 tetra-, and 1 pentameric repeat).

Of the 6,005 loci identified from WGS, 155 loci (2.6% of all potential loci identified) containing penta-, tetra-, and trimeric repeat motifs were tested for amplification and polymorphism at TAMU except for 21 loci containing trimeric repeats that were tested at UNI (see Method 1 above). Tests of amplification and polymorphism were conducted in 24 *G. aurei* individuals collected from pocket gopher specimen TCWC 64963 from the vicinity of La Nutrias, Socorro County, New Mexico. Genomic DNA was extracted from individual lice as described above in Method 1 except louse exoskeletons were saved in ethanol in a –20 C freezer post-extraction. PCR amplifications of microsatellites used unlabeled 5′ (forward) and 3′ (reverse) primers and a fluorescently labeled 5′-tail-sequence primer following Karlsson et al. (2008) and Boutin-Ganache et al. (2001). All PCRs were performed in 10 μl reactions containing 3.75 μl Emerald Master Mix (Takara Bio Inc., Mountain View, California), 5.75 μl water, 0.5 μl fluorescently labeled tail primer (6-FAM, NED, PET, or VIC; 5′-GCCTCGTTTATCAGATGTGGA-3′; 10 μM), 0.5 μl forward primer (1 μM) with same tail as the fluorescently labeled tail primer, 0.5 μl reverse primer (10 μM), and 1 μl DNA. Thermal cycling parameters were as follows: 1 cycle of denaturation for 2 min at 95 C, 10 cycles of denaturation at 95 C for 30 sec, annealing at 58 C for 45 sec, and elongation at 72 C for 60 sec, 10 cycles of denaturation at 95 C for 30 sec, annealing at 56 C for 45 sec, and elongation at 72 C for 60 sec, 15 cycles of denaturation at 95 C for 30 sec, annealing at 52 C for 45 sec, and elongation at 72 C for 60 sec, followed by a final extension at 72 C for 45 min. Amplified PCR products were sent to the DNA Analysis Facility on Science Hill at Yale University, combined with Gel Company Liz-500 size standard, and run on a 3730xl 96-Capillary Genetic Analyzer (Applied Biosystems). Sizes of microsatellite fragments were visualized using GeneMarker software (version 1.90, SoftGenetics, State College, Pennsylvania). In total, 15 loci (13 tested at TAMU and 2 tested at UNI) amplified successfully and were polymorphic, leading to further screening.

After the initial screening of loci at both universities, microsatellite loci that were determined polymorphic in either of our 2 approaches were assessed further in a single laboratory using the same approach to test for population-level genetic variability and reliable scoring. Data included here resulted from 390 individual chewing lice collected from 13 pocket gopher individuals from within approximately 1 mile of La Joya, Socorro County, New Mexico (all pocket gopher hosts are housed at the Louisiana State University Museum of Natural Science [LSUMZ] or the BRTC; catalog numbers are LSUMZ 30743, 30744, 30785, and 33915; TCWC 64297–64299, 64301, 64302, 64971, and 64973–64975). For all these louse individuals, genomic DNA was extracted and amplified as described above in Method 1, with louse bodies mounted on microscope slides for preservation as voucher specimens following the extraction process. After sizes of amplification products were initially determined, most loci were amplified in multiplex (Table I) with 1 μl multiplex-primer mix added to the 10 μl reaction. Each multiplex-primer mix contained 0.5 μM each forward primer, 2 μM each reverse primer, and 4 μM fluorophore-labeled M13 primer per primer pair included. Thermal cycling parameters, post-PCR electrophoresis, and genotyping were the same as described above in Method 1. The software Convert (version 1.31; Glaubitz, 2004) was used to reformat all data files for use in additional genetic analysis programs. During data analysis, loci were eliminated from consideration if they amplified sporadically, produced alleles that could not easily be binned into expected sizes based on simple repeat structures, or showed evidence of null alleles or heterozygote deficit in multiple infrapopulations of lice. In total, 12 microsatellite loci were chosen for analysis, six from each methodology (i.e., from the microsatellite enriched library or from WGS; Table I).

Locus quality was evaluated with tests for genotypic disequilibrium between each pair of loci in each infrapopulation, which were performed using Genepop 4.6 (Rousset, 2008) and corrected for multiple tests using the Benjamini and Hochberg (1995) False Discovery Rate correction (http://www.sdmproject.com/utilities/?show=FDR). These tests yielded 11 pairs of loci with statistically significant genotypic disequilibrium after False Discovery Rate correction for 885 tests (Benjamini and Hochberg, 1995). None of the pairs of loci with significant disequilibrium were in disequilibrium in more than 1 infrapopulation of the 13 tested. Among the 20 genotypic comparisons with the lowest *p*-values (significant or not), there was no apparent pattern of pairs of loci that trended toward disequilibrium in multiple populations, with the single exception of a pair of loci that was in genotypic disequilibrium in 2 of the 13 infrapopulations. Therefore, chromosomal linkage of loci was deemed unlikely, and all 12 loci tested were retained for further analysis.

The Brookfield (1996) method was used to test for null alleles by using the null.all option of PopGenReport (Adamack and Gruber, 2014) in the R environment (version 3.3.3; R Core Team, 2013) implemented using RStudio (version 1.0.136; RStudio Team, 2016). These tests indicated a low frequency of null alleles for the 12 loci of interest (range –0.02–0.03; Table I). Null allele estimates for all 12 loci fell within the 2.5 and 97.5 percentiles. Heterozygote deficit and F_{IS}, both possible indicators of null alleles, were calculated using Genepop 4.6 (Rousset, 2008). Tests for heterozygote deficit run for all loci and all populations were not corrected for multiple tests, yet none of the 12 loci tested showed statistically significant heterozygote deficit in more than 1 or 2 of the 13 populations tested. Furthermore, F_{IS} averaged over all infrapopulations was low for each locus (Table I), also suggesting that these markers do not exhibit null alleles for these lice.

As a measure of diversity, expected heterozygosity (H_E) per locus was calculated in Arlequin (version 3.5.1.2; Excoffier and Lischer, 2010), as was F_{ST} and AMOVA for measures of differentiation among infrapopulations. The 12 loci examined showed adequate genetic diversity for population-level analyses, with 3–9 alleles per locus in the 390 individuals sampled from a single locality (Table I). Average H_E per locus ranged 0.32–0.70 for these individuals (Table I). Individual infrapopulations of lice showed F_{ST} values that were low (ranging 0.01–0.05), but

TABLE I. *Geomydoecus aurei* microsatellite loci, identification (ID) methodology (enriched library or whole genome sequencing [WGS]), multiplex group, repeat sequence, primer sequence, and measures of locus reliability and diversity. Number of alleles was determined via the examination of 390 louse individuals.

Locus	ID method, multiplex group, and repeat motif	Primer sequences (underlined portions of sequence are tags described in methods)	No. of alleles, allele range	Null allele frequency estimate	Mean F_{IS}	Mean H_E
Ga3702	Library, 1, AAGT	F: 5′-CGAGTTTTCCCAGTCACGACCCCTTTACTTGCTGCTGATTTAC R: 5′-GTTTCTTCTCATCAGCATGGTCTACACATG	5,172–5,188	0.005	−0.01	0.61
Ga4103	Library, 1, AAAG	F: 5′-CGAGTTTTCCCAGTCACGACAACTCCGGGCGATTCGAATAC R: 5′-GTTTCTTACTCTGACAATAATCGTATCGTTG	5,202–5,218	0.032	0.11	0.32
Ga6020	Library, 1, AAAG	F: 5′-CGAGTTTTCCCAGTCACGACACATTGAAGTGACAGAATCGTCC R: 5′-GTTTCTTTCTTCTATTGGTTTACGAAGGCAC	4,290–4,302	0.025	0.04	0.55
Ga4282	Library, 2, AAAG	F: 5′-CGAGTTTTCCCAGTCACGACCACATAATGGGAGCGAAAGGAG R: 5′-GTTTCTTGCTCCCGAATATGATGGATGAAC	8,194–8,213	0.013	0.01	0.72
Ga4911	Library, 2, AAAT	F: 5′-CGAGTTTTCCCAGTCACGACACCTGAGAAGTTTCACCATTCAC R: 5′-GTTTCTTCGGGTTGTTGTGCTTTAATTGTC	3,247–3,256	0.007	0.01	0.43
Ga4863	Library, 3 or 5, AAG	F: 5′-CGAGTTTTCCCAGTCACGACTAGTAGTCGTCTAGAAAGTCCGC R: 5′-GTTTCTTCAAACTTTCACCACCTCGCTATC	6,210–6,224	0.024	0.04	0.70
Ga33816	WGS, 4, ATC	F: 5′-CGAGTTTTCCCAGTCACGACGCCCATGTAAACTGCCACC R: 5′-GTTTCTTTTATGTGCTGGAGTTTGGAGG	9,274–9,309	−0.011	−0.04	0.84
Ga29676	WGS, 4, AAT	F: 5′-CGAGTTTTCCCAGTCACGACTGTCTTTGTTACCGTCTGCAG R: 5′-GTTTCTTGAAGGTTCAATCAGTGGGCTAC	6,406–6,421	0.006	0.01	0.39
Ga4737	WGS, 5, AAT	F: 5′-CGAGTTTTCCCAGTCACGACTTTCGACGTAAAGCTGCCAC R: 5′-GTTTCTTATTAGTCCTCTCGCGCCTTC	6,267–6,282	0.022	0.05	0.60
Ga43595	WGS, 5, AGGC	F: 5′-CGAGTTTTCCCAGTCACGACACTCTCCGGCCCGATATTTC R: 5′-GTTTCTTATGATCAGGGCGGTTAAAGC	5,328–5,344	−0.017	−0.07	0.58
Ga739	WGS, 5, AGG	F: 5′-CGAGTTTTCCCAGTCACGACTCACTGTGTTCTGACTCTTCG R: 5′-GTTTCTTGAGGAACGTTTGAGTGAAGAAG	5,411–5,422	−0.012	−0.06	0.54
Ga51656	WGS, 5, AAGG	F: 5′-CGAGTTTTCCCAGTCACGACTGCCTGCATCTGAAGTGATTG R: 5′-GTTTCTTACTTTGAGCTTTGCACTGAAG	6,452–6,475	−0.025	−0.10	0.45

significantly different from zero in 68 of 78 pairwise infrapopulation comparisons, indicating that these highly polymorphic markers will be useful at the infrapopulation and population levels for *G. aurei*. Similarly, AMOVA analysis indicated that genetic variation among infrapopulations accounts for a low, but significant, percentage (2%, $P < 0.001$) of the overall genetic variation.

As a final measure of locus reliability, locus-specific genotyping error rates were estimated for 64 lice from 8 different gophers (8 lice each), which were randomly chosen from *G. aurei* from La Joya and nearby localities to re-amplify and verify genotypes. These tests showed no genotyping errors in which allele size was incorrectly identified in individuals after reamplification and rescoring, and no cases of heterozygotes being misidentified as homozygotes or vice versa, indicating a locus-specific genotyping error rate of zero for each of the 12 loci.

Overall, the 2 next-generation sequencing approaches used in this study generated large numbers of potential microsatellite loci for population genetics purposes. The library enriched for microsatellites (via 454-sequencing) of course yielded fewer reads (21,713 single-sequence reads) and fewer potentially useful loci (4,375) than WGS via Illumina HiSeq (58,340,400 paired reads and 6,005 loci, respectively), but both approaches yielded more loci than could be feasibly tested. For example, we tested only 0.8% and 2.6% of all possible loci identified with enriched library sequencing and WGS, respectively). Only 12 of the 189 loci tested (6%; Table I) met all our criteria for inclusion (e.g., consistent amplification, alleles that could be binned into expected sizes based on simple repeats, no evidence of null alleles); therefore, a large proportion of loci were tested but ultimately were not useful, resulting in a loss of time and money. Such inefficiency in the search for informative microsatellite loci and overall low success rate could be considered concerning. However, this aspect of designing useful markers is likely common regardless of how loci are obtained, whether that be via traditional, library-building methods or with next-generation sequencing methods and is a vital part of ensuring quality data for population analysis.

At this particular point, WGS remains expensive enough that there is a cost savings that can be derived from a more limited sequencing run of an enriched library. However, given the continued decline in prices, WGS may eventually become a more economical approach for obtaining molecular markers. Perhaps more importantly, WGS offers the additional advantage of providing data that could potentially be leveraged for other applications in addition to development of microsatellite markers. For example, read mapping from WGS data against assemblies could identify single-nucleotide polymorphisms (SNPs), which also can be used in population genetic studies. Beyond the population level, targeted local assemblies of 1:1 ortholog genes can be built and used for phylogenomic studies. If the study organism has a small genome size, then sequencing can be performed at relatively high coverage (Allen et al., 2017; Boyd et al., 2017). In fact, a total of 1,107 protein-coding genes have already been assembled from the WGS data used here and incorporated in a higher-level phylogenetic framework along with orthologous sequences from other species (K. P. Johnson et al., unpub. data). Genomes of the bacterial symbionts of parasites (e.g., Boyd et al., 2014) can also be assembled from WGS data, and it should be feasible to assemble complete mitochondrial genomes from these libraries, as has been done for the human body louse (*Pediculus humanus*; Shao et al., 2009). Thus, a WGS approach opens a multitude of data for further study beyond the original intended use.

Parasitologists are already taking advantage of next-generation approaches, with many recent studies examining a variety of questions using these techniques in medically important parasite taxa (e.g., Cantacessi et al., 2010, 2012; Webb and Rosenthal, 2011; Ascunce et al., 2013; Monzón et al., 2016). With careful planning (Bild et al., 2014), anyone working with non-model organisms, including parasites of wildlife, can incorporate next-generation sequencing methods into their research and potentially address a variety of questions with the molecular data they obtain.

For assistance in the laboratory at the University of Northern Iowa, we thank E. Ament, C. Freese, J. Hill, S. Huebner, A. Miller, L. Pietan, B. Ross, and M. Stueck. We also thank J. Brown, J. Henningsen, N. Heunis, W. Preisser, and S. Stephens for assistance in the laboratory at Texas A&M University. This research was funded by the National Science Foundation (DEB 1445708 and DEB 1239788).

LITERATURE CITED

Adamack, A. T., and B. Gruber. 2014. PopGenReport: Simplifying basic population genetic analyses in R. Methods in Ecology and Evolution **5:** 384–387.

Allen, J. M., B. M. Boyd, N. Nguyen, P. Vachaspati, T. Warnow, D. I. Huang, P. G. S. Grady, K. C. Bell, Q. C. B. Cronk, L. Mugisha, et al. 2017. Phylogenomics from whole genome sequences using aTRAM. Systematic Biology **66:** 786–798.

Ascunce, M. S., M. A. Toups, G. Kassu, J. Fane, K. Scholl, and D. L. Reed. 2013. Nuclear genetic diversity in human lice (*Pediculus humanus*) reveals continental differences and high inbreeding among worldwide populations. PLoS One **8:** e57619. doi:https://doi.org/10-1371/journal.pone.0057619.

Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B (Methodological) **57:** 289–300.

Bild, A. H., J. T. Chang, W. E. Johnson, and S. R. Piccolo. 2014. A field guide to genomics research. PLoS Biology **12:** e1001744. doi:https://doi.org/10.1371/journal.pbio.1001744.

Boutin-Ganache, I., M. Raposo, M. Raymond, and C. E. Deschepper. 2001. M13-tailed primers improve the readability and usability of microsatellite analyses performed with two different allele-sizing methods. BioTechniques **31:** 24–27.

Boyd, B. M., J. M. Allen, V. de Crecy-Legard, and D. L. Reed. 2014. Genome sequencing of *Candidatus* Riesia pediculisschaeffi, endosymbiont of chimpanzee lice, and genomic comparison of recently aquired endosymbionts from human and chimpanzee lice. G3 (Bethesda) **4:** 2189–2195. doi:10.1534/g3.114.012567.

Boyd, B. M., J. M. Allen, N. Nguyen, A. D. Sweet, T. Warnow, M. D. Shapiro, S. M. Villa, S. E. Bush, D. H. Clayton, and K. P. Johnson. 2017. Phylogenomics using target-restricted assembly resolves intra-generic relation-

ships of parasitic lice (Phthiraptera: *Columbicola*). Systematic Biology **66:** 786–798.

Brookfield, J. F. Y. 1996. A simple new method for estimating null allele frequency from heterozygote deficiency. Molecular Ecology **5:** 453–455.

Brownstein, M. J., J. D. Carpten, and J. R. Smith. 1996. Modulation of non-templated nucleotide addition by Taq DNA polymerase: Primer modifications that facilitate genotyping. BioTechniques **20:** 1004–1010.

Cantacessi, C., B. E. Campbell, and R. B. Gasser. 2012. Key strongylid nematodes of animals—Impact of next-generation transcriptomics on systems biology and biotechnology. Biotechnology Advances **30:** 469–488.

Cantacessi, C., M. Mitreva, B. E. Campbell, R. S. Hall, N. D. Young, A. R. Jex, S. Ranganathan, and R. B. Gasser. 2010. First transcriptomic analysis of the economically important parasitic nematode, *Trichostrongylus colubriformis*, using a next-generation sequencing approach. Infection, Genetics and Evolution **10:** 1199–1207.

Cox, M. P., D. A. Peterson, and P. J. Biggs. 2010. SolexaOA: At-a-glance quality assessment of Illumina second-generation sequencing data. BMC Bioinformatics **11:** 485. doi:10.1186/1471-2105-11-485.

Criscione, C. D., R. Poulin, and M. S. Blouin. 2005. Molecular ecology of parasites: Elucidating ecological and microevolutionary processes. Molecular Ecology **14:** 2247–2257.

Demastes, J. W., T. A. Spradling, M. S. Hafner, G. R. Spies, D. J. Hafner, and J. E. Light. 2012. Cophylogeny on a fine scale: *Geomydoecus* chewing lice and their pocket gopher hosts, *Pappogeomys bulleri*. Journal of Parasitology **98:** 262–270.

Detwiler, J. T., I. C. Caballero, and C. D. Criscione. 2017. Role of parasite transmission in promoting inbreeding: I. Infection intensities drive individual parasite selfing rates. Molecular Ecology **26:** 4391–4404.

Detwiler, J. T., and C. D. Criscione. 2011. Testing Mendelian inheritance from field-collected parasites: Revealing duplicated loci enables correct inference of reproductive mode and mating system. International Journal for Parasitology **41:** 1185–1195.

Detwiler, J. T., and C. D. Criscione. 2017. Role of parasite transmission in promoting inbreeding: II. Pedigree reconstruction reveals sib-transmission and consequent kin-mating. Molecular Ecology **26:** 4405–4417.

Duffy, C. W., H. Ba, S. Assefa, A. D. Ahouidi, Y. B. Deh, A. Tandia, F. C. M. Kirsebom, D. P. Kwiatkowski, and D. J. Conway. 2017. Population genetic structure and adaptation of malaria parasites on the edge of endemic distribution. Molecular Ecology **26:** 2880–2894.

Excoffier, L., and H. E. L. Lischer. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetic analyses under Linux and Windows. Molecular Ecology Resources **10:** 564–567.

Faircloth, B. 2008. MSATCOMMANDER: Detection of microsatellite repeat arrays and automated, locus-specific primer design. Molecular Ecology Resources **8:** 92–94.

Glaubitz, J. C. 2004. CONVERT: A user-friendly program to reformat diploid genotypic data for commonly used population genetic software packages. Molecular Ecology Resources **4:** 309–310.

Hafner, M. S., and S. A. Nadler. 1988. Phylogenetic trees support the coevolution of parasites and their hosts. Nature **332:** 258–259.

Hafner, M. S., P. D. Sudman, F. X. Villablanca, T. A. Spradling, J. W. Demastes, and S. A. Nadler. 1994. Disparate rates of molecular evolution in cospeciating hosts and parasites. Science **265:** 1087–1090.

Harper, S., T. A. Spradling, J. W. Demastes, and C. S. Calhoun. 2015. Host behavior drives parasite genetics at multiple geographic scales: Population genetics of the chewing louse, *Thomomydoecus minor*. Molecular Ecology **24:** 4129–4144.

Huyse, T., R. Poulin, and A. Theron. 2005. Speciation in parasites: A population genetics approach. Trends in Parasitology **21:** 469–467.

Jirsová, D., J. Štefka, and J. Miloslav. 2017. Discordant population histories of host and its parasite: A role for ecological permeability of extreme environment. PLoS One **12:** e0175286. doi:https://doi.org/10.1371/journal.pone.0175286.

Johnson, K. P., B. L. Williams, D. M. Drown, R. J. Adams, and D. H. Clayton. 2002. The population genetics of host specificity: Genetic differentiation in dove lice (Insecta: Phthiraptera). Molecular Ecology **11:** 25–38.

Karlsson, S., M. A. Renshaw, E. Rexroad, and J. R. Gold. 2008. PCR primers for 100 microsatellites in red drum (*Sciaenops ocellatus*). Molecular Ecology Resources **8:** 393–398.

Kearse, M., R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, et al. 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics **28:** 1647–1649.

Koop, J. A. H., K. E. DeMatteo, P. G. Parker, and N. K. Whiteman. 2014. Birds are islands for parasites. Biology Letters **10:** 20140255. doi:10.1098/rsbl.2014.0255.

Leo, N. P., J. M. Hughes, X. Yang, S. K. S. Poudel, W. G. Brogdon, and S. C. Barker. 2005. The head and body lice of humans are genetically distinct (Insecta: Phthiraptera: Pediculidae): Evidence from double infestations. Heredity **95:** 34–40.

Mazé-Guilmo, E., S. Blachet, K. D. McCoy, and G. Loot. 2016. Host dispersal as the driver of parasite genetic structure: a paradigm lost? Ecology Letters **19:** 336–347.

McCoy, K. D. 2009. Host–parasite determinants of parasite population structure: Lessons from bats and mites on the importance of time. Molecular Ecology **18:** 3545–3547.

Monzón, J. D., E. G. Atkinson, B. M. Henn, and J. L. Benach. 2016. Population and evolutionary genomics of *Amblyomma americanum*, an expanding arthropod disease vector. Genome Biology and Evolution **8:** 1351–1360.

Nadler, S. A., M. S. Hafner, J. C. Hafner, and D. J. Hafner. 1990. Genetic differentiation among chewing louse populations (Mallophaga: Trichodectidae) in a pocket gopher contact zone (Rodentia: Geomyidae). Evolution **44:** 942–951.

Nessner, C. E., J. J. Andersen, M. A. Renshaw, M. M. Giresi, and J. E. Light. 2014. Characterization of 17 novel polymorphic microsatellite loci in the mammal chewing louse *Geomydoecus ewingi* (Insecta: Phthiraptera) for population genetic analyses. Journal of Parasitology **100:** 873–877.

Opiro, R., N. P. Saarman, R. Echodu, E. A. Opiyo, K. Dion, A. Halyard, A. W. Dunn, S. Aksoy, and A. Caccone. 2017. Genetic diversity and population structure of the tsetse fly *Glossina fuscipes fuscipes* (Diptera: Glossinidae) in Northern Uganda: Implications for vector control. PLoS Neglected Tropical Diseases **11:** e101291488. doi:https://doi.org/10.1371/journal.pntd.0005485.

Peterson, B. K., J. N. Weber, E. H. Kay, H. S. Fisher, and H. E. Hoekstra. 2012. Double digest RADseq: An inexpensive model for *De Nova* SNP discovery and genoyping in model and non-model species. PLoS One **7:** e32253. doi:https://doi.org/10.1371/journal.pone.0037135.

Puritz, J. B., J. A. Addison, and R. J. Toonen. 2012. Next-generation phylogeography: a targeted approach for multilocus sequencing of non-model organisms. PLoS One **7:** e34241. doi:https://doi.org/10.1371/journal.pone.0034241.

R Core Team. 2013. R: a language and environment for statistical computer. R Foundation for Statistical Computing, Vienna, Austria. Available at: http://www.R-project.org/. Accessed 1 June 2017.

Riesgo, A., S. C. S. Andrade, P. P. Sharma, M. Novo, A. R. Pérez-Porro, V. Vahtera, V. L. González, G. Y. Kawauchi, and G. Giribet. 2012. Comparative description of ten transcriptomes of newly sequenced invertebrates and efficiency estimation of genomic sampling in non-model taxa. Frontiers in Zoology **9:** 33. doi:https://doi.org/10.1186/1742-9994-9-33.

Rousset, F. 2008. GENEPOP'007: A complete re-implementation of the GENEPOP software for Windows and Linux. Molecular Ecology Resources **8:** 103–106.

Rozen, S., and H. J. Skaletsky. 2000. Primer3 on the WWW for general users and for biologist programmers. *In* Bioinformatics methods and protocols: Methods in molecular biology, S. Krawetz and S. Misener (eds.). Humana Press, Totowa, New Jersey, p. 365–386.

RStudio Team. 2016. RStudio: Integrated development for R. RStudio, Inc., Boston, Massachusetts. Available at: http://www.rstudio.com. Accessed 1 June 2017.

Schuelke, M. 2000. An economic method for the fluorescent labeling of PCR fragments. Nature Biotechnology **18:** 233–234.

Shao, R., E. F. Kirkness, and S. C. Barker. 2009. The single mitochondrial chromosome typical of animals has evolved into 18 minichromosomes in the human body louse, *Pediculus humanus*. Genome Research **19:** 904–912.

Sikes, R. S., and the Animal Care and Use Committee of the American Society of Mammalogists. 2016. 2016 Guidelines of the American Society of Mammalogists for the use of wild mammals in research and education. Journal of Mammalogy **97:** 663–688.

Thompson, J. N. 1994. The coevolutionary process. University of Chicago Press, Chicago, Illinois, 376 p.

Webb, K. M., and B. M. Rosenthal. 2011. Next-generation sequencing of the *Trichinella murrelli* mitochondrial genome allows comprehensive comparison of its divergence from the principal agent of human trichinellosis, *Trichinella spiralis*. Infection, Genetics and Evolution **11:** 116–123.

Zerbino, D. R., and E. Birney. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Research **18:** 821–829.