Target enrichment of thousands of ultraconserved elements sheds new light on early relationships within New World sparrows (Aves: Passerellidae)

Robert W. Bryson, Jr., 1* Brant C. Faircloth, 2 Whitney L. E. Tsai, 3 John E. McCormack, 3 and John Klicka 1

1 Department of Biology and Burke Museum of Natural History and Culture, University of Washington, Seattle, Washington, USA
2 Department of Biological Sciences and Museum of Natural Science, Louisiana State University, Baton Rouge, Louisiana, USA
3 Moore Laboratory of Zoology, Occidental College, Los Angeles, California, USA
* Corresponding author: brysonjr@uw.edu
Submitted January 28, 2016; Accepted March 16, 2016; Published May 25, 2016

ABSTRACT
Sparrows in the nine-primaried oscine family Passerellidae represent an attractive model for studying avian diversification across North and South America. However, the lack of phylogenetic resolution at the base of the New World sparrow tree has hampered the use of the existing sparrow phylogeny to test questions about the evolution of sparrow traits. We generated phylogenomic data from 1,063 ultraconserved elements to estimate phylogenetic relationships among the major clades of New World sparrows. Concatenated and species-tree analyses of 271,830 base pairs of sequence data converged on a well-supported phylogeny that differs from previous estimates. The resolved backbone of the sparrow phylogeny provides new insight into the biogeography of this radiation by suggesting both a tumultuous biogeographic history, with many colonizations of South America, and several independent ecological transitions to different habitat types.

Keywords: nine-primaried oscines, phylogenetics, species tree, UCEs

INTRODUCTION
Sparrows in the nine-primaried oscine family Passerellidae are one of the largest songbird radiations in the New World. The family consists of 26 genera and 129 species distributed across North and South America (Barker et al. 2013, Klicka et al. 2014, Slager and Klicka 2014). Although most species inhabit temperate grasslands in the Americas, New World sparrows also occupy arid tropical deciduous thornscrub, montane forests, and a variety of other habitat types (Klicka and Spellman 2007, Rising et al. 2011). Many temperate species are long-distance migrants that winter in the Neotropics, and the repeated evolution of seasonal shifts in distribution among Passerellidae may have led to the high species diversity within the family (Winger et al. 2014).

The broad distribution of New World sparrows makes them an attractive model for studying the relative impacts of geological events, paleoclimate change, migration, and ecological shifts on avian diversification across North and South America. However, the lack of phylogenetic resolution at the base of the New World sparrow tree has hampered these types of investigations (Klicka et al. 2014).
Phylogenetic analyses of multilocus data from all described species suggest that nearly all sparrows can be placed within 8 well-supported clades (Klicka et al. 2014; Figure 1). The relationships among these clades appear to be well resolved on the basis of mitochondrial DNA (mtDNA; Figure 1A). In contrast, multilocus data revealed poorly supported relationships, especially among the older sparrow lineages (Figure 1B, 1C). From these results, Klicka et al. (2014) concluded that the handful of genes used (2 mtDNA and 5 nuclear) were insufficiently informative to disentangle questions of sparrow relationships.

The advent of large-scale phylogenomic datasets has transformed the field of phylogenetics. It is now common for researchers to sequence hundreds to thousands of loci in nonmodel taxa. Target enrichment of ultraconserved elements is one of several next-generation sequencing approaches (Faircloth et al. 2012) that can produce large DNA datasets. Ultraconserved elements (UCEs) are a class of highly conserved and abundant nuclear markers found throughout the genomes of a variety of animals (Bejerano et al. 2004, Stephen et al. 2008, Faircloth et al. 2012, 2013, 2015). Sequence data from UCEs are relatively easy to obtain and, together with DNA adjacent to UCE locations (flanking DNA), can resolve both shallow-level (<5 mya; e.g., Smith et al. 2014) and deep-level (>100 mya; e.g., McCormack et al. 2013, Faircloth et al. 2015) relationships.

Here, we use phylogenomic data from >1,000 UCEs to infer phylogenetic relationships among the major clades of New World sparrows. We expected that the 100-fold increase in loci over the 7 genes used in Klicka et al. (2014)
would better resolve relationships within the New World sparrow tree and, in doing so, provide researchers with the evolutionary framework needed to understand the evolution of these songbirds across North and South America.

**METHODS**

**Taxon Sampling**

We collected tissues from 28 individuals representing all 8 major clades of New World sparrows (Klicka et al. 2014; Table 1) and the monotypic species Pezopetes capitalis. Seven clades were represented by ≥2 samples. We also included 2 samples, Geothlypis trichas and Icterus gularis, to serve as outgroups (Barker et al. 2013, Klicka et al. 2014), and we harvested UCE data from the genome sequence of Geospiza fortis (described below), which we used as an additional set of outgroup loci.

**Library Preparation, UCE Target Enrichment, and Sequencing**

We extracted genomic DNA from vouchered bird tissues using Qiagen DNeasy Blood & Tissue Kits (Qiagen, Valencia, California, USA), and we prepared sequencing libraries for Illumina sequencing using Nextera XT Library Preparation Kits (Illumina, San Diego, California; Adey et al. 2010). We generally followed the sequence-capture workflow for Nextera-prepared libraries detailed in Faircloth et al. (2012), but we modified this protocol by using one-quarter aliquots of the required Nextera kit reagents and genomic DNA, and we pooled indexed libraries into groups of 8 libraries, prior to enrichment. We enriched each pool using a set of 5,472 custom-designed probes (MYbaits; MYcroarray, Ann Arbor, Michigan, USA) targeting 5,060 UCE loci (Faircloth et al. 2012) following an open-source protocol (for the full protocol, see http://www.ultraconserved.org). Following enrichment and an 18-cycle PCR-recovery, we combined pooled libraries in equimolar ratios to a final concentration of 10 nM, and we sent the pooled libraries to the UCLA Genoseq Facility for 250 bp (base pair) paired-end sequencing using 2 runs of an Illumina MiSeq.

**Data Extraction from Geospiza fortis**

In addition to the data we collected from Geothlypis trichas and Icterus gularis, we harvested UCE loci from the existing

---

**TABLE 1.** Taxa used in this study, with specimen source, locality information, and clade membership. Clade designations follow Klicka et al. (2014), as shown in Figure 1. Institution abbreviations: JFBM = James Ford Bell Museum of Natural History; LSUMNS = Louisiana State University Museum of Natural Science; MVZ = Museum of Vertebrate Zoology, University of California, Berkeley; and UWBM = University of Washington, Burke Museum.

<table>
<thead>
<tr>
<th>Species</th>
<th>Specimen [tissue]</th>
<th>Locality</th>
<th>Clade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aimophila rufescens</td>
<td>UWBM 93611 [DHB 3531]</td>
<td>Honduras: Copan C</td>
<td>B</td>
</tr>
<tr>
<td>Ammodramus lecontei</td>
<td>UWBM 80599 [CSW 6816]</td>
<td>USA: North Dakota A</td>
<td>A</td>
</tr>
<tr>
<td>A. savannarum</td>
<td>JFBM 43678 [JK 94-056]</td>
<td>USA: Montana G</td>
<td>G</td>
</tr>
<tr>
<td>Amphispiza bilineata</td>
<td>UWBM 96901 [TKA 098]</td>
<td>USA: Nevada F</td>
<td>F</td>
</tr>
<tr>
<td>Arrenon aurantirostis</td>
<td>UWBM 93635 [JK 00-053]</td>
<td>Honduras: Copan E</td>
<td>E</td>
</tr>
<tr>
<td>A. brunneinucha</td>
<td>UWBM 69340 [DAB 1706]</td>
<td>Nicaragua: Granada E</td>
<td>E</td>
</tr>
<tr>
<td>Arrenonopis rufivirgatus</td>
<td>JFBM 43282 [X6828]</td>
<td>USA: Texas G</td>
<td>G</td>
</tr>
<tr>
<td>Artemisiospiza belli</td>
<td>MVZ 170007 [20306]</td>
<td>USA: California A</td>
<td>A</td>
</tr>
<tr>
<td>Atlapetes citrinellus</td>
<td>UWBM 92260 [B8332]</td>
<td>Argentina: Tucuman C</td>
<td>C</td>
</tr>
<tr>
<td>A. gutturalis</td>
<td>UWBM 93636 [GAV 1374]</td>
<td>Honduras: Copan C</td>
<td>C</td>
</tr>
<tr>
<td>Calamospiza melanocorys</td>
<td>JFBM 43748 [JK 94-058]</td>
<td>USA: Montana F</td>
<td>F</td>
</tr>
<tr>
<td>Chlorospingus ophthalamicus</td>
<td>UWBM 114696 [JK 99-074]</td>
<td>Honduras: Copan H</td>
<td>H</td>
</tr>
<tr>
<td>Chondestes grammacus</td>
<td>UWBM 103584 [DHB 3910]</td>
<td>USA: Arizona F</td>
<td>F</td>
</tr>
<tr>
<td>Junco phaeonotus</td>
<td>UWBM 107920 [GMS 596]</td>
<td>Mexico: Jalisco D</td>
<td>D</td>
</tr>
<tr>
<td>Melospiza melodia</td>
<td>JFBM 43741 [JK 94-084]</td>
<td>USA: Montana A</td>
<td>A</td>
</tr>
<tr>
<td>Melzone aberti</td>
<td>UWBM 98275 [DHB 2352]</td>
<td>USA: Nevada B</td>
<td>B</td>
</tr>
<tr>
<td>M. leucotis</td>
<td>UWBM 109906 [JK 02-053]</td>
<td>Guatemala: Quetzaltenango B</td>
<td>B</td>
</tr>
<tr>
<td>Oriturus supercilius</td>
<td>FMNH 343328</td>
<td>Mexico: Morelos A</td>
<td>A</td>
</tr>
<tr>
<td>Passerculus sandwichensis</td>
<td>UWBM 110768 [JK 03-341]</td>
<td>Canada: Newfoundland A</td>
<td>A</td>
</tr>
<tr>
<td>Passerella iliaca</td>
<td>JFBM [DNA extract FOSP91]</td>
<td>USA: California D</td>
<td>D</td>
</tr>
<tr>
<td>Peucaea ruficaua</td>
<td>UWBM 69323 [DAB 1680]</td>
<td>Nicaragua: Rivas G</td>
<td>G</td>
</tr>
<tr>
<td>Pezopetes capitalis</td>
<td>LSUMZ 138633 [B-16232]</td>
<td>Costa Rica: San Jose –</td>
<td>–</td>
</tr>
<tr>
<td>Pipilo chloropus</td>
<td>UWBM 77977 [VGR 711]</td>
<td>USA: Arizona C</td>
<td>C</td>
</tr>
<tr>
<td>Poecetes gramineus</td>
<td>JFBM 43743 [JK 94-063]</td>
<td>USA: Montana A</td>
<td>A</td>
</tr>
<tr>
<td>Rhynchospiza strigiceps</td>
<td>UWBM 92416 [DHB 2425]</td>
<td>Argentina: Salta G</td>
<td>G</td>
</tr>
<tr>
<td>Spizella atrogularis</td>
<td>UWBM 109557 [JK 01-210]</td>
<td>USA: Nevada F</td>
<td>F</td>
</tr>
<tr>
<td>Spizelloides arborea</td>
<td>UWBM 66284 [GKD 126]</td>
<td>USA: Alaska D</td>
<td>D</td>
</tr>
<tr>
<td>Zonotrichia atricapilla</td>
<td>UWBM 80383 [EVL 740]</td>
<td>USA: Alaska D</td>
<td>D</td>
</tr>
<tr>
<td>Geothlypis trichas</td>
<td>UWBM 77569 [EVL 782]</td>
<td>USA: Arizona Outgroup</td>
<td>Outgroup</td>
</tr>
<tr>
<td>Icterus gularis</td>
<td>UWBM 114470 [JK 11-007]</td>
<td>Mexico: Chiapas Outgroup</td>
<td>Outgroup</td>
</tr>
</tbody>
</table>
We performed unpartitioned concatenated maximum-likelihood (ML) analyses of the 75% complete dataset using RAxML 8.0.19 (Stamatakis 2014) by conducting 20 searches of the data with the GTR GAMMA site-rate substitution model for the best ML tree. We assessed support for the best ML topology by performing nonparametric bootstrapping using the autoMRE option in RAxML with the GTR GAMMA site-rate substitution model, and we reconciled the best ML tree with the bootstrap replicates using RAxML. We performed unpartitioned concatenated Bayesian analyses of the 75% complete dataset with ExaBayes 1.2.1 (Aberer et al. 2014) using 4 independent runs of 1 million iterations with heated chains turned off. We assessed convergence of the posterior distribution using the default thinning for ExaBayes (500 iterations) after discarding 25% of the posterior sample as "burn-in" by (1) computing the standard deviation of split frequencies, (2) examining the ESS (effective sample size) and potential scale reduction factor (PSRF) values of the parameter estimates output by the "postProcParam" ExaBayes helper program, and (3) visually assessing posterior trace distributions using Tracer 1.6.0 (Rambaut et al. 2014). We computed the majority rules extended (MRE) consensus of the thinned, post-burn-in posterior distribution using the "consense" helper program that is part of ExaBayes.

We also analyzed a subset of the loci in the 75% complete dataset using a gene-tree-based coalescent approach. Although coalescent approaches offer a promising alternative to concatenated analyses, which can mislead inferences under certain conditions (Kubatko and Degnan 2007, Liu and Pearl 2007, Linkem et al. 2016), coalescent methods that are applicable to phylogenomic datasets can be negatively affected by poorly resolved gene trees (Gatesy and Springer 2014, Lanier et al. 2014, Xi et al. 2015, Meiklejohn et al. 2016). Poorly resolved gene trees often result when alignments input to tree-inference programs contain reduced phylogenetic signal—one characteristic of conserved loci, particularly when the phylogenetic analysis spans relatively recent divergence times. To address this potential problem, several filtering approaches have been proposed to exclude loci that may bias species-tree analyses (e.g., Harris et al. 2014, Hosner et al. 2016, Meiklejohn et al. 2016). Here, we computed the number of parsimony-informative sites in each locus of our 75% complete dataset, and we created a subset of loci having a number of parsimony-informative sites in the upper quartile of the range. We sampled this reduced dataset by loci, with replacement, to create 100 resampled datasets, and we performed nonparametric bootstrapping by sites of each locus in each dataset using RAXML with a GTR GAMMA model of site-rate substitution. After running the gene-tree estimation procedure, we sorted the bootstrapped gene trees to ensure that each resampled dataset was resampled by loci and by sites (Seo 2008). We input the 100 resampled datasets to ASTRAL-II 4.7.8 (Mirarab and Warnow 2015), which is a summary-based species-tree-inference program that is reasonably robust to the effects of incorrectly estimated gene trees (Simmons and Gatesy 2015, Meiklejohn et al. 2016). We computed the extended majority-rule consensus of the resulting species trees using RAXML, and we collapsed branches with bootstrap support values <70%.

**RESULTS**

**Sequence Data**

After trimming reads for adapter contamination and quality, we collected an average of 958,240 sequence reads...
(95% confidence interval [CI]: 140,934) from each library, with an average length of 136.6 bp (95% CI: 4.9 bp; Supplemental Material Appendix A), and we assembled these reads into an average of 32,245 contigs (95% CI: 4,110) with an average length of 258.5 bp (95% CI: 2.1 bp; Supplemental Material Appendix B). After UCE detection, we enriched an average of 1,403 contigs (95% CI: 95) representing UCE loci with an average length of 343 bp (95% CI: 8.5) and an average sequencing coverage of 6× (95% CI: 0.4; Supplemental Material Appendix C). After taxon-set creation, FASTA extraction, alignment, and alignment trimming, the 75% complete dataset contained 1,063 loci with an average length of 256 bp (95% CI: 2.7 bp), an average of 27 taxa locus⁻¹ (95% CI: 0.2), and an average of 2.4 parsimony-informative sites locus⁻¹ (95% CI: 0.2). The concatenated supermatrix of the 75% complete dataset contained 271,830 bp, 36,816 alignment patterns, and 2,538 parsimony-informative sites. The upper quartile of parsimony-informative loci in the 75% complete dataset were those having ≥3 parsimony-informative sites (n = 329).

**Phylogenetic Estimation**

Maximum-likelihood and Bayesian analyses of the concatenated UCE loci converged on a strongly supported topology (Figure 2A). Monophyly of the major clades (A–H) in Klicka et al. (2014) represented by ≥2 samples in our analyses was strongly supported, and only one node in the tree received <70% bootstrap and 0.95 posterior probability support. This node (collapsed in Figure 2A) weakly supported a sister relationship between *Artemisiospiza bellii* and *Poecetes gramineus*. The species tree reconstructed from the screened subset of UCE loci was less well supported (Figure 2B) but generally matched the concatenated UCE topology. Ten nodes received <70% bootstrap support, mostly involving splits within clades. Phylogenetic relationships among clades were strongly supported, with one exception near the base of the tree. Monophyly of all
but one of the major sparrow clades represented by ≥2 samples in our analyses was strongly supported.

**DISCUSSION**

Phylogenetic analysis of 1,063 UCE loci and 271,830 bp of aligned sequence data produced a strongly supported backbone phylogeny for New World sparrows. Of the internal nodes defining clades A–H, only the placement of clade G was not well resolved (Figure 2). By contrast, the most reliable estimate of sparrow relationships in Klicka et al. (2014), based on mtDNA and few nuclear markers, yielded only 2 resolved internal nodes, and none of the focal clades (A–H) could be placed with confidence (Figure 1C). In Klicka et al. (2014), the problematic clade G was also unresolved in the mtDNA, nuclear DNA, and species-tree analyses.

The UCE phylogeny suggests that prior phylogenetic hypotheses based on mtDNA, as well as those based on small numbers of nuclear genes, might each be misleading in their own ways, with biologically interesting implications. Perhaps surprisingly, in one important instance the UCE tree supported a relationship indicated by the previous mtDNA tree rather than one indicated by the nuclear DNA tree. In Klicka et al. (2014), the mtDNA tree shows clade D as sister to the clade comprising clades A, B, and C (Figure 1A). This relationship contrasts with that suggested by the nuclear gene tree from the same study, which instead places clade D as sister to clade A (Figure 1B). The relationship between clades A and D is biologically interesting because these 2 clades comprise the grassland and brushland sparrows. Thus, the prior nuclear data support a single origin for this ecological trait and a degree of phylogenetic niche conservatism, whereas the mtDNA data support more ecological transitions through time. The UCE data strongly support the latter hypothesis, with the grassland and brushland sparrows not coming together as sister clades. Rather, each is a member of a broader group containing lower-latitude species like those in the speciose genus *Atlapetes*.

In another instance, the UCE tree conflicts with the prior mtDNA tree. At their bases, the UCE tree (Figure 2) and the prior nuclear gene tree (Figure 1B) both support a sister relationship between clade F and the remaining sparrow clades. The mtDNA tree instead strongly supports clades G and H as sister to the rest of the sparrows, with clade F nested well within the New World sparrow tree (Figure 1A). This discrepancy is difficult to explain without careful testing and could be caused by analytical and biological factors that affect organellar and nuclear DNA differently. Clarifying clade F as the sister to the rest of the sparrow phylogeny, however, is an important component of any future attempts to analyze the evolution of New World sparrow traits through time.

The UCE trees generated from concatenated and species-tree analyses agreed on branching order of the deep nodes, whereas the species tree was less resolved than the concatenated tree with respect to more shallow divergences. The concatenated UCE tree showed only a single polytomy within clade A (Figure 2A). The UCE species tree, meanwhile, showed 10 collapsed nodes scattered among 5 of the 8 core clades (Figure 2B). For the remaining (uncollapsed) nodes, the concatenated tree and species tree differed at a single node. The species tree suggested a sister relationship between *Aimaphila rufescens* and *Melozone aberti*; the concatenated UCE tree placed *M. aberti* and *M. leucotis* as sister species. The lack of resolution of the UCE species tree toward the tips of the tree could reflect the relatively lower informative content of UCEs for very shallow timescales, especially because our library preparation method led to shorter locus lengths (see below) than comparable studies carried out more recently (e.g., Giara and Esselstyn 2015, McCormack et al. 2016).

A resolved hypothesis of basal relationships among sparrow clades has been lacking until now. The strong support along the “backbone” of the phylogeny allows new glimpses into the sparrow radiation in the New World. For example, the UCE phylogeny suggests a tumultuous biogeographic history. Sparrows are well represented in both North and South America. Each of the 8 core clades has representation in North America, and elements of 5 clades (C, D, E, G, and H) are represented in South America. Given a probable North American origin (Barker et al. 2015), the UCE tree suggests multiple independent colonizations of South America. As mentioned above, the UCE tree also provides insights into the origins of ecological diversity in sparrows. Grassland-dwelling sparrows, for example, do not form a clade, but instead occur in 3 phylogenetically distant clades (A, F, and G), suggesting more transitions between different habitat types (e.g., grassland, brushland, and forest) over evolutionary time than were previously suspected (e.g., Paynter 1964, Dickerman et al. 1967).

Detailed analyses that involve mapping distributional, behavioral, or phenotypic characters on the sparrow phylogeny await a tree with better taxonomic sampling. We generated phylogenomic data from only a small percentage (~22%) of sparrow species to infer relationships across the backbone of the tree. Future studies should fill in the phylogeny with additional UCE data from all passerellid species. Alternatively, our UCE data could be combined with the more densely sampled 5-locus dataset of Klicka et al. (2014) in a hybrid phylogenetic–phyloge nomic approach (Leaché et al. 2014). This hybrid approach could produce a strongly supported species-level phylogeny of New World sparrows amenable to biogeographic and character-state reconstructions without the need for further data collection.
Finally, testing, optimizing, and reducing costs of laboratory methods for target enrichment of UCEs has been a priority since we began using the technique to infer phylogeny, and library preparation is the single most expensive step, on a per sample basis, of the target enrichment process. Here, we prepared libraries using the Illumina Nextera enzyme in one-quarter reactions to reduce costs. Although we knew that the Nextera enzyme produces shorter insert sizes, on average, than random shearing (cf. Crawford et al. 2012, 2015), we found the reduced contig sizes an acceptable tradeoff, given the increased efficiency during library preparation that the Nextera system offers. However, dilution of Illumina Nextera XT transposase was inconsistent in our hands, particularly compared to earlier, similar dilutions of Epicentre Nextera transposase, which worked well (Crawford et al. 2012, Faircloth et al. 2012, McCormack et al. 2013). Dilution of the Nextera XT reagent appears to have affected this study by producing shorter contigs (expected) while also reducing the consistency of enrichment and sequencing coverage across enriched UCE loci (unexpected). Typically, we obtain longer average locus lengths (400–800 bp) and more consistent enrichment of UCE loci when we and others use alternative shearing and library preparation protocols (e.g., Giarla and Esselstyn 2015, McCormack et al. 2016). Nonetheless, we were still able to recover >1,000 UCE loci from 30 taxa in our 75% complete matrix despite the shorter contig lengths and inconsistent enrichment that we observed, demonstrating the power of target enrichment approaches over Sanger-based techniques even when the performance of the enrichment procedure is suboptimal.

ACKNOWLEDGMENTS
We thank the following curators, collection managers, and institutions for tissue samples critical for the completion of this study: F. Sheldon, R. Brumfield, and D. Dittrmann (Louisiana State University Museum of Natural Science); N. Johnson and C. Cicero (University of California, Berkeley, Museum of Vertebrate Zoology); R. Zink (James Ford Bell Museum of Natural History); and S. Rowher and S. Birks (University of Washington, Burke Museum of Natural History). We also thank D. Crowley for help in the lab, and I. Lovette and an anonymous reviewer for suggestions that improved the manuscript.

Funding statement: This work was funded in part by National Science Foundation (NSF) grants DEB-0315469 and DEB-1257785 (to J.K.) and DEB-1258205 (to J.E.M.), and computational portions were supported by DEB-1242260 (to B.C.F.). Portions of this research were also conducted with high-performance computing resources provided by Louisiana State University (http://www.hpc.lsu.edu). None of the funders had any input into the content of the manuscript, nor did they require manuscript approval before submission or publication.

Ethics statement: All federal, state, and local permits were secured prior to fieldwork.

Author contributions: (1) J.K. and J.E.M. conceived the idea; (2) R.W.B. and W.L.E.T. generated the data; (3) B.C.F. analyzed the data; (4) R.W.B., B.C.F., W.L.E.T., J.E.M., and J.K. wrote the paper; and (5) B.C.F., J.E.M., and J.K. contributed substantial materials, resources, and funding.

Data availability: Raw read data and assembled contigs are available from NCBI BioProject PRJNA318526. Aligned sequences are available at 10.6084/m9.figshare.3101359.

LITERATURE CITED


anchor thousands of genetic markers spanning multiple evolutionary timescales. Systematic Biology 61:717–726.


McCormack, J. E., M. G. Harvey, B. C. Faircloth, N. G. Crawford, T. C. Glenn, and R. T. Brumfield (2013). A phylogeny of birds based on over 1,500 loci collected by target enrichment and high-throughput sequencing. PLOS One 8:e54848. doi:10.1371/journal.pone.0054848


