PLOS CURRENTS



PMCID: PMC6072552 PMID: 30116653

Version 1. PLoS Curr. 2018 July 13; 10: ecurrents.tol.aba557de56be881793261f7e1565cf35. Published online 2018 July 13. doi: 10.1371/currents.tol.aba557de56be881793261f7e1565cf35: 10.1371/currents.tol.aba557de56be881793261f7e1565cf35 Tree of Life

A Post-pleistocene Calibrated Mutation Rate from Insect Museum Specimens

Gideon Ney, Katy Frederick, and Johannes Schul

Gideon Ney, Department of Biological Sciences, University of Missouri, Columbia, Missouri, United States; Contributor Information.

Copyright © 2018 Ney, Frederick, Schul, et al

This is an open access article distributed under the terms of the <u>Creative Commons Attribution License</u>, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Quantifying the age of recent species divergence events can be challenging in the absence of calibration points within many groups. The katydid species *Neoconocephalus lyristes* provides the opportunity to calibrate a post-Pleistocene, taxa specific mutation rate using a known biogeographic event, the Mohawk-Hudson Divide. DNA was extracted from pinned museum specimens of *N. lyristes* from both Midwest and Atlantic populations and the mitochondrial gene COI sequenced using primers designed from extant specimens. Coalescent analyses using both strict and relaxed molecular clock models were performed in BEAST v1.8.2. The assumption of a strict molecular clock could not be rejected in favor of the relaxed clock model as the distribution of the standard deviation of the clock rate strongly abutted zero. The strict molecular clock model resulted in an intraspecific calculated mutation rate of 14.4-17.3 %/myr, a rate substantially higher than the common rates of sequence evolution observed for insect mitochondrial DNA sequences. The rate, however, aligns closely with mutation rates estimated from other taxa with similarly recent lineage divergence times.

Introduction

In recent years, many examples of rapid speciation and diversification occurring during the last glacial cycle (i.e., within 500 kyr BP $\frac{1}{2}$), or even after the last glacial maximum (LGM, 19 kyr BP $\frac{2}{2}$) have been described. Arguably, the most impressive examples of rapid diversification are the cichlid radia-

tion events within the African Rift Valley, where a small number of founding species diversified into hundreds of species after the LGM ^{3,4,5}. Other examples include the old world pea aphids ⁶, North American songbirds ^{7,8,9}, and the threespine sticklebacks of British Columbia ¹⁰; in some cases, significant diversification arose in as little as 50 years ¹¹.

The accurate timing of diversification events allows us to better understand the mechanisms leading to phenotypic diversification and/or speciation. Molecular clock techniques allow the timing of diversification events based on estimates of the rate of genetic mutations per unit time ^{12,13}. Mutation rates are gene specific and can vary between lineages and through time within a lineage ^{14,15}. Therefore, accurate dating using a molecular clock requires reliable calibration of the rate of sequence evolution for that particular group, time interval, and gene. Rates can be calibrated using nodes dated from fossils ¹⁶ and from biogeographic vicariance events ^{17,18}.

Estimates of nucleotide evolution vary greatly dependent with the age of the calibrating point, with younger calibration points resulting in higher rate estimates 19,20,21. Fossils and most biogeographic events are ancient (millions of years old) and are appropriate for the dating of similarly ancient events. The few available rate estimates using very young age calibration points (<200kyrs 22,23,24) suggest an exponential increase of estimated rates 25,26,27; additional data is needed to support this pattern. The exponential pattern of estimates is likely an artifact of the estimation methods and does not reflect true differences in rates on nucleotide evolution 19. One reason for the small number of estimates for recent lineage divergences is that suitably recent calibration points are scarce 25, since these events are too recent to use fossil evidence.

Here we use a postglacial vicariance event to calibrate a lineage specific mutation rate for North American *Neoconocephalus* katydids. At a time following the LGM, water from the North American Great Lakes drained through the Mohawk-Hudson Outlet to the Atlantic coast ²⁸. Wetland habitats formed within the Hudson and Mohawk Valleys, which allowed coastal plain species to expand their ranges into the wetlands surrounding the Great Lakes ²⁹. The opening of the St. Lawrence Seaway (10,750-10,600 14C yr BP ³⁰), diverted melt water and led to the drying of the wetlands in the Mohawk-Hudson outlet. This vicariance event left disjunct wetland habitats in the Midwest (mainly bogs and fens) and along the Atlantic Coast (bogs and marsh habitat). Such disjunct ranges matching this pattern are found in plant, reptile, amphibian, and insect species possessing a coastal plain affinity ^{29,31,32}. *Neoconocephalus lyristes* is an example of such a habitat specialist, limited to bog and fen wetlands. The species' described range follows the pattern of the Mohawk-Hudson Divide, with isolated populations in the Great Lakes area ³¹ and North Atlantic Coast (³³ Fig. 1).

The eleven North American *Neoconocephalus* katydid species possess markedly little genetic variation despite their high diversity of species-specific call patterns and may be an example of a recent species radiation ^{36,37}. The accurate timing of this radiation will help identify the evolutionary mechanisms leading to rapid species diversification observed in this group. *Neoconocephalus lyristes* provides a unique opportunity among species of *Neoconocephalus* for the calibration of a post-Pleistocene mutation rate as gene flow between these two disjunct ranges likely ceased with the draining of the Mohawk-Hudson Outlet (10,750-10,600 14C yr BP ³⁰). Here we sequenced mtDNA from museum specimens representing both populations and estimated an intraspecific mutation rate using a coalescent Bayesian method.

Methods

Over three years of searching previous collection sites we found only a single extant population of *N. lyristes*, in Cedar Bog Nature Preserve, Urbane, OH, USA. Due to apparent local extinction of *N. lyristes* from most of its Midwest and its entire Atlantic range, we used museum samples collected in the first half of the 20th century. We selected 18 dried *N. lyristes* specimens, from the Hebard Collection at the Academy of Natural Sciences of Drexel University for DNA extraction and analysis. Specimens represent samples from both Atlantic Coastal and Midwest populations (Fig. 1), with collection dates ranging from 1905-1932. We used a non-destructive method for DNA extraction (modified from ³⁸). A hind leg was removed and placed in a 1.5 ml microcentrifuge tube fully submerged in one ml of digestion buffer: 3 mM CaCl2, 2% sodium dodecyl sulphate (SDS), 40 mM dithiothreitol (DTT), 250 mg/ml proteinase K, 100 mM Tris buffer pH 8 and 100 mM NaCl (quantities represent molarity of final concentrations). Hind legs were incubated overnight (17-19 hrs.) at 55°C. Following digestion we removed the hind legs from buffer and placed them in 100% EtOH for two hours to stop enzymatic activity. Extraction of DNA contained in the buffer was completed using the standard Qiagen DNeasy Blood + Tissue Kit (Qiagen Inc., Valencia, CA, USA) extraction method.

Amplification took place in a laboratory without prior exposure to DNA that could be amplified by primers used in this study. Polymerase chain reaction (PCR) prep was performed in a UV hood. All equipment and surfaces were sanitized with a 10% bleach solution and tools were sanitized in a UV Stratalinker 1800. For this study we designed six overlapping primer pairs (Appendix: Supplemental Table 1) around non-variable regions of the mitochondrial gene cytochrome oxidase I (COI). These primers were based upon extant *N. lyristes*, *N. robustus*, and *N. bivocatus* COI sequences and designed using the Primer3³⁹ plugin in Geneious v6.0.5⁴⁰. Each primer pair amplified approximately 150 bp; combined, they provide complete coverage of the 743 bp target region.

PCR amplification was performed on an Eppendorf Mastercycler gradient (Eppendorf-Brinkman Instruments Inc., Westbury, NY, USA) using Taq DNA polymerase (Platinum Taq, Invitrogen Inc., Carlsbad, CA, USA). All primers were used at a concentration of 10 mM. Thermocycling conditions for all six primer-sets are as follows: Hot start at 94°C 2 min, denaturation at 94°C 30 sec, annealing at 56°C 30 sec, extension 72°C 40 sec, repeated 40x, with a final 72°C extension for 7 min. Amplified PCR products were prepared for sequencing using a ExoI/SAP enzymatic cleanup (2.75 μ l 10x SAP buffer, 0.5 μ l SAP, 0.25 μ l ExoI per 20 μ l of PCR product) incubated at 37°C for 30 min, followed by 80°C for 15 min to inactivate enzymes. Sequencing was performed at the DNA Core Facility, University of Missouri, Columbia, MO, USA on an ABI 3730 DNA Analyzer, using standard Big Dye Terminator cycle sequencing chemistry (Applied Biosystems, Foster City, CA, USA). Sequences were edited, aligned and trimmed in Geneious v6.0.5 ⁴¹. We used a global alignment with free end gaps and 70% similarity rule. Regions of sequence with high ambiguity were labeled as missing. One individual, with greater than ten percent ambiguity, was removed from the analysis (m017). Individual m007 failed to amplify. We successfully sequenced COI from 16 individuals.

We evaluated substitution models using jModel Test v0.1.1 $\frac{42}{}$ and found GTR+G to be a suitable model. Phylogenetic analyses were conducted using a coalescent method as implemented in BEAST v1.8.2 xml $\frac{43}{}$; input files were formatted using BEAUti v1.7.4 $\frac{43}{}$. Our analysis assumed a constant

population size for the coalescent inferences ⁴⁴. We ran this analysis to convergence, performing ten runs with twenty million generations sampled every two thousand trees. We assessed convergence through visual inspection of posterior values among the ten runs in Tracer v1.5 ⁴⁵. This analysis was performed using both a strict ⁴⁶, as well as a relaxed molecular clock model ¹³. The Midwest individuals were run both unconstrained as well as constrained to monophyly. The constrained run assured that the age calibration point was assigned to the correct node in all trees ²¹. To evaluate the influence of the prior settings on the posterior samples, we repeated the analysis as above but without any sequence data.

Using the radiocarbon date of $10,750\pm150$ 14C yr BP, the end of 150-300 year period of steady melt water flow following the final large flood through the Hudson Valley at 10,900 14C yr BP ³⁰, we calibrated the calendar age of the Mohawk-Hudson Divide. We performed the radiocarbon to calendar age conversion using the IntCal13 curve in OxCal v4.2 online ⁴⁷. The age estimate was fixed to the highest likelihood value within the 95% confidence interval; yielding a calibrated date of 10,739.5 cal BP. Being a known biogeographic barrier we allowed the node age prior probability of the Midwest clade to vary along a normal distribution, with the calibrated date as the mean age and a standard deviation of one-thousand years. This allows for the possibility of lineage divergence prior to the biogeographic event, as well as the overestimation of the events age ⁴⁸. The Euclidean mean and standard deviation priors were set to exponential with mean values of 10 and 0.3 respectively. Convergence of MCMC runs was visualized using Tracer v1.5 ⁴⁵ to ensure that all runs converged. With Tracer v1.5 we ascertained the average mutation rate between populations of *N. lyristes* based on the Mohawk-Hudson calibration. Runs were combined in LogCombiner v1.8.2 ⁴⁹ and a maximum clade credibility consensus tree was formed in TreeAnnotator v1.7.4 ⁵⁰.

Results

We successfully sequenced 743 bp of the mitochondrial gene COI from sixteen individuals (5 from Midwest and 11 from Atlantic Coast populations, Table 1). Sequence similarity among the 16 samples ranged from 92.0% to 99.8%. We found the greatest diversity within the Atlantic population. The Midwest clade fell within the larger clade of Atlantic Coast *N. lyristes* (Fig. 2). This observation is congruent with the hypothesized biogeographic history of the species where the Midwest populations diverging from the ancestral Atlantic population.

Using the unconstrained coalescence model, four out of five Midwestern individuals formed a clade within the larger clade of Atlantic Coast *N. lyristes*. One Midwest individual (m010) grouped among Atlantic individuals (Appendix: Supplemental Fig. 1). In order to prevent the age calibration point from being assigned to the wrong nodes in some trees we constrained the Midwest clade to monophyly in further analyses ²¹. The resulting constrained consensus tree (Fig. 2) is congruent with the hypothesized biogeographic history of the species, with the Midwest population diverging from the ancestral Atlantic population.

Using a relaxed clock model, we obtained branch specific mutation rates between 14.4 and 37.5 %/myr from the consensus of the ten runs. The average rate of mutation among branches was 15.8 %/myr, ranging from 15.7-15.9 %/myr between the ten independent runs. The distribution of the stan-

dard deviation of the clock rate strongly abutted zero when the relaxed molecular clock was used (Fig. 3). This indicates support for a constant rate of substitution and a strict molecular clock was used $\frac{51}{1}$. The strict molecular clock analysis produced a tree (Fig. 2) with a similar, but not identical, topology to the relaxed clock's consensus tree. The relationship between Midwest animals and their relationship to the Atlantic clade remained unchanged, with minor changes in the relationships between Atlantic individuals. The strict consensus tree, with Midwest clade constrained to monophyly possessed an average mutation rate of 17.3 %/myr, with mutation rates between the ten runs. Predictably, a slower rate of 14.4 %/myr was obtained when the same analysis was run with individual M010 removed. These two rates, while diverging slightly, both indicate a rate of mutation significantly faster than most reported in the literature $\frac{19,20}{2}$.

Discussion

Here, we focused on the calibration of an intraspecific mutation rate at a very recent timescale. Evolutionary rates calibrated across divergent timescales can be markedly different 25 , with younger calibration dates (<1 Mya) showing substantially higher estimates of rates divergence than older lineages 19,20 . In mammals, for example, the age of the calibration dates shows a negative relationship with estimates for molecular evolutionary rates 19,52 . Metastudies utilizing insect mtDNA rates estimated from both inter- and intraspecific calibrations show a similar pattern to that observed in mammals 20,21 . Available data suggest an exponential increase of estimated rates 22,23,24 with decreasing calibration age (Fig. 4). The exponential pattern of estimates is likely an artifact of the estimation methods and does not reflect true differences in rates on nucleotide evolution 19 .

The sequence variation among populations has two components, fixed substitutions between them that have accumulated since divergence and current within population variation ¹⁹. The fixed substitutions among lineages represent the actual evolutionary divergence. Most of the within population genetic variation will be removed over time by genetic drift and selection and therefore only a small fraction will ultimately contribute to lineage divergence ^{26,53}. For young divergence times, the within species variation will contribute a much larger fraction of the total nucleotide differences, as only few fixed substitutions have accumulated. For ancient divergence times, in contrast, the same amount of within population variation would be dwarfed by fixed substitutions accumulated since divergence ^{27,54}. Thus, short calibration times should lead to gross overestimations of evolutionary divergence rates, while ancient calibration times (>1 Mya ⁵⁴) should provide much more realistic estimates.

Insect mtDNA rates of mutation:

We estimated a mutation rate for COI at 14.4-17.3 %/myr, using the strict molecular clock model and a very recent calibration time. Our estimate is significantly higher than the commonly assumed mtDNA mutation rate of 1.15 %/myr ^{55,56}, which were based on much older divergence times. Estimates of substitution rates calibrated from the age of the Mid-Aegean Trench (9-12 Mya), for example, within an insect model range from 1.0-2.7 %/myr dependent upon application of various substitution and clock models ²¹. Our estimated rate of 14.4-17.3 %/myr, on the other hand, aligns with estimates found using similarly recent calibration dates (Fig. 4). A mutation rate of 19.2 %/myr was estimated for the European butterfly *Parnassius mnemosyne*, calibrated with a vicariance event at

10,000 years BP ²³. Intermediate calibration dates resulted in an intermediate estimate of evolutionary rates. The mutation rate for the North American ground beetle (*Nebria*) was estimated at 5.7 %/myr, using a vicariance event dated to 150,000 years BP ²². Our estimates for *N. lyristes* fit into the exponential pattern previously described (Fig. 4). Thus, this study agrees with the slower estimates of Orthopteran mtDNA sequence evolution and may serve as an internal calibration point for *Neoconocephalus* diversification.

The high mutation rate inferred from our data set could be due to problems in the mathematical models underlying the molecular clock. This seems unlikely, since both fixed clock and relaxed clock models lead to nearly identical results. Furthermore, the close fit of our data point to data from previous work conducted with a variety of methods ^{20,21} suggests that our particular methods were not responsible for the high estimated mutation rate.

As more evidence accumulates supporting the occurrence of postglacial species diversification, the greater the need for appropriate tools for timing these events. This will in part include the utilization of young vicariance events for molecular clock calibration. Geologically supported postglacial vicariance events within North America are lacking for many taxa groups ^{57,58}. The Mohawk-Hudson Divide provides a recent biogeographic vicariance event, with the potential for the calibration of lineage specific mutation rates for a number of plant, amphibian, reptile, and insect groups.

Use of museum samples:

The use of ancient DNA (aDNA) samples can be hindered by severe degradation $\frac{38,59}{2}$. In this study two of the eighteen samples could not be sequenced successfully. These two samples were not the oldest, nor from the same locality. Severe degradation of DNA, beyond that in the other sixteen samples, or a mismatch in primer binding sites may account for failed amplification (Table 1). In those samples that were sequenced successfully ambiguities were high, while this is likely due to the degraded nature of aDNA, the coamplification of nuclear pseudogenes could also lead to such ambiguities. The amplification of relatively short (150 bp) segments increases the likelihood of amplifying pseudogenes, not amplified when targeting longer sequences. Nuclear pseudogenes of COI, while not noted in *Neoconocephalus*, have been found in other Orthopterans ⁶⁰. We found no internal stop codons within our COI sequences. As internal stop codons are common in pseudogenes, it is unlikely that our data is affected by their presence. Our primers were developed from COI reference sequences from three extant Neoconocephalus species. Amplification would therefore not be affected by sequence degradation, as may be the case if primers are developed from the aDNA itself. One concern with the use of aDNA is sequence degradation, with post mortem C-U deamination ⁶¹, reflected in higher than expected percentage of Thymine in resulting sequences. We compared the percentages of nucleotides in sequences from our museum specimens and from live collected N. lyristes, which were almost identical (e.g., GC content 35.1% v. 36.2%). This indicates that sequence degradation has minor, if any, influence on our results.

In this study museum specimens replaced extant samples, necessitated by the rarity, or likely local extinction, of *N. lyristes* from most of its known range. Despite the additional challenges of working with museum specimens, aDNA can replace extant specimens when collection is either not possible because of extinction $\frac{61,62}{2}$ or broad resampling is untenable $\frac{63,64}{2}$. With advances in the amplification of ancient DNA ^{65,66,67}, museum collections are also opening up areas of study that are not possible with extant data alone ^{62,67,68}. Ancient DNA can be utilized in the calibration of molecular clocks through dating tip ages ⁶⁹. Samples from multiple time points, can provide additional information about the genetic and demographic changes in groups over time ⁷⁰. Ancient DNA has been used in the reconstruction and timing of many mammal groups ^{52,70,71,72}, but remains underutilized in the timing of insect lineages despite the abundance of specimens in museums. Several of the problems associated with the use of aDNA can be overcome by next generation sequencing (NGS). For example, NGS has the capability to target short and degraded DNA samples ⁷³. NGS also allows for the sequencing of whole genomes from aDNA ^{64,74,75} and less destructive sampling techniques from Museum samples ^{76,77,78}.

Data Availability

All supplementary data are available at figshare: <u>https://dx.doi.org/10.6084/m9.figshare.3100312.v2</u>

Nucleotide sequences are available at GenBank: Accession numbers KU881748 - KU881763

Competing Interests

The authors have declared that no competing interests exist.

Corresponding Author

Gideon Ney: gideon.ney@gmail.com

Editor

Due to Tree of Life Editor unavailability, a member of the *PLOS ONE* Editorial Board, Wolfgang Arthofer (University of Innsbruck, Austria), rendered the final decision on this paper.

Appendix

Acknowledgments

We thank the Academy of Natural Sciences of Drexel University, especially Jason Weintraub, for the loan of the museum specimens utilized in this study. Thanks to Cedar Bog Nature Preserve for allowing the collection of extant samples. Thanks to the Eggert and Birchler Labs for the use of laboratory space and advice on molecular techniques. Thanks to Gavin Conant, Corey Hudson, Megan Murphy, Chris Pires, and Anna Scharnagl for helpful feedback on the manuscript.

Funding Statement

This work was supported by a grant of the National Science Foundation (IOS 1146878) and through a grant from the Orthopterisits' Society.

Contributor Information

Gideon Ney, Department of Biological Sciences, University of Missouri, Columbia, Missouri, United States.

Katy Frederick, Department of Biological Sciences, University of Missouri, Columbia, Missouri, United States.

Johannes Schul, Department of Biological Sciences, University of Missouri, Columbia, Missouri, United States.

References

1. Rohling, E. J., Fenton, M., Jorissen, F. J., Bertrand, P., Ganssen, G., & Caulet, J. P. (1998). Magnitudes of sea-level lowstands of the past 500,000 years. Nature, 394(6689), 162-165.

2. Yokoyama, Y., Lambeck, K., De Deckker, P., Johnston, P., & Fifield, L. K. (2000). Timing of the Last Glacial Maximum from observed sea-level minima. Nature, 406(6797), 713-716. [PubMed: 10963593]

3. Nagl, S., Tichy, H., Mayer, W. E., Takezaki, N., Takahata, N., & Klein, J. (2000). The origin and age of haplochromine fishes in Lake Victoria, East Africa. Proceedings of the Royal Society of London B: Biological Sciences, 267(1447), 1049-1061. [PMCID: PMC1690633] [PubMed: 10874756]

4. Seehausen, O. (2006). African cichlid fish: a model system in adaptive radiation research. Proceedings of the Royal Society of London B: Biological Sciences, 273(1597), 1987-1998. [PMCID: PMC1635482] [PubMed: 16846905]

5. Meyer, B. S., Indermaur, A., Ehrensperger, X., Egger, B., Banyankimbona, G., Snoeks, J., & Salzburger, W. (2015). Back to Tanganyika: a case of recent trans-species-flock dispersal in East African haplochromine cichlid fishes. Royal Society Open Science, 2(3), 140498. [PMCID: PMC4448823] [PubMed: 26064619]

6. Peccoud, J., Ollivier, A., Plantegenest, M., & Simon, J. C. (2009). A continuum of genetic divergence from sympatric host races to species in the pea aphid complex. Proceedings of the National Academy of Sciences, 106(18), 7495-7500. [PMCID: PMC2678636] [PubMed: 19380742]

7. Johnson, N. K., & Cicero, C. (2004). New mitochondrial DNA data affirm the importance of Pleistocene speciation in North American birds. Evolution, 58(5), 1122-1130. [PubMed: 15212392]

8. Milá, B., McCormack, J. E., Castañeda, G., Wayne, R. K., & Smith, T. B. (2007). Recent postglacial range expansion drives the rapid diversification of a songbird lineage in the genus Junco. Proceedings of the Royal Society of London B: Biological Sciences, 274(1626), 2653-2660. [PMCID: PMC2279216] [PubMed: 17725978]

9. Weir, J. T., & Schluter, D. (2008). Calibrating the avian molecular clock. Molecular Ecology, 17(10), 2321-2328. [PubMed: 18422932]

10. Taylor, E. B., & McPhail, J. D. (1999). Evolutionary history of an adaptive radiation in species pairs of threespine

sticklebacks (Gasterosteus): insights from mitochondrial DNA. Biological Journal of the Linnean Society, 66(3), 271-291.

11. Lescak, E. A., Bassham, S. L., Catchen, J., Gelmond, O., Sherbick, M. L., von Hippel, F. A., & Cresko, W. A. (2015). Evolution of stickleback in 50 years on earthquake-uplifted islands. Proceedings of the National Academy of Sciences, 112(52), E7204-E7212. [PMCID: PMC4702987] [PubMed: 26668399]

12. Zuckerkandl, E., & Pauling, L. (1962). Molecular disease, evolution and genetic heterogeneity.

13. Zuckerkandl, E., & Pauling, L. (1965). Evolutionary divergence and convergence in proteins. Evolving genes and proteins, 97, 97-166.

14. Drummond, A. J., Ho, S. Y., Phillips, M. J., & Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. PLoS biology, 4(5), 699. [PMCID: PMC1395354] [PubMed: 16683862]

15. Ho, S. Y., & Larson, G. (2006). Molecular clocks: when times are a-changin'. TRENDS in Genetics, 22(2), 79-83. [PubMed: 16356585]

16. Müller, J., & Reisz, R. R. (2005). Four well-constrained calibration points from the vertebrate fossil record for molecular clock estimates. BioEssays, 27(10), 1069-1075. [PubMed: 16163732]

17. Gaunt, M. W., & Miles, M. A. (2002). An insect molecular clock dates the origin of the insects and accords with palaeontological and biogeographic landmarks. Molecular Biology and Evolution, 19(5), 748-761. [PubMed: 11961108]

18. Mendelson, T. C., & Shaw, K. L. (2005). Sexual behaviour: rapid speciation in an arthropod. Nature, 433(7024), 375-376. [PubMed: 15674280]

19. Ho, S. Y., Phillips, M. J., Cooper, A., & Drummond, A. J. (2005). Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. Molecular biology and evolution, 22(7), 1561-1568. [PubMed: 15814826]

20. Ho, S. Y., & Lo, N. (2013). The insect molecular clock. Austral Entomology, 52(2), 101-105.

21. Papadopoulou, A., Anastasiou, I., & Vogler, A. P. (2010). Revisiting the insect mitochondrial molecular clock: the mid-Aegean trench calibration. Molecular Biology and Evolution, 27(7), 1659-1672. [PubMed: 20167609]

22. Clarke, T. E., Levin, D. B., Kavanaugh, D. H., & Reimchen, T. E. (2001). Rapid evolution in the Nebria gregaria group (Coleoptera: Carabidae) and the paleogeography of the Queen Charlotte Islands. Evolution, 55(7), 1408-1418. [PubMed: 11525464]

23. Gratton, P., Konopiński, M. K., & Sbordoni, V. (2008). Pleistocene evolutionary history of the Clouded Apollo (Parnassius mnemosyne): genetic signatures of climate cycles and a 'time-dependent' mitochondrial substitution rate. Molecular Ecology, 17(19), 4248-4262. [PubMed: 18986502]

24. Haag-Liautard, C., Coffey, N., Houle, D., Lynch, M., Charlesworth, B., & Keightley, P. D. (2008). Direct estimation of the mitochondrial DNA mutation rate in Drosophila melanogaster. PLoS Biol, 6(8), e204. [PMCID: PMC2517619] [PubMed: 18715119]

25. Ho, S. Y., Saarma, U., Barnett, R., Haile, J., & Shapiro, B. (2008). The effect of inappropriate calibration: three case studies in molecular ecology. [PMCID: PMC2229648] [PubMed: 18286172]

26. Ho, S. Y., Shapiro, B., Phillips, M. J., Cooper, A., & Drummond, A. J. (2007). Evidence for time dependency of molecular rate estimates. Systematic biology, 56(3), 515-522. [PubMed: 17562475]

27. Ho, S. Y., Duchêne, S., Molak, M., & Shapiro, B. (2015). Time-dependent estimates of molecular evolutionary rates: evidence and causes. Molecular Ecology, 24(24), 6007-6012. [PubMed: 26769402]

28. Teller, J. T. (2003). Controls, history, outbursts, and impact of large late-Quaternary proglacial lakes in North America. Developments in Quaternary Sciences, 1, 45-61.

29. Reznicek, A. A. (1994). The disjunct coastal plain flora in the Great Lakes region. Biological Conservation, 68(3), 203-215.

30. Rayburn, J. A., Knuepfer, P. L., & Franzi, D. A. (2005). A series of large, Late Wisconsinan meltwater floods through the Champlain and Hudson Valleys, New York State, USA. Quaternary Science Reviews, 24(22), 2410-2419.

31. Thomas, E. S. (1933) Neoconocephalus lyristes (Rhen and Hebard) in the Middle West. Annals of the Entomological Society of America, XXVI: 303-308.

32. Thomas, E. S. (1951). Distribution of Ohio animals. Ohio J. Sci, 51(4), 154-167.

33. Rehn J. A. G., Hebard M. (1915) A synopsis of the species of the genus Neoconocephalus found in North America north of Mexico. Trans. Am. Ent. Soc. 40: 365-413.

34. Walker, T.J. (2000) Neoconocephalus. Singing Insects of North America. Available from http://entnemdept.ufl.edu/walker /buzz/g185a.html. Accessed November, 2015.

35. Whitesell, J.J. 1969. Biology of United States coneheaded katydids of the genus Neoconocephalus (Orthoptera: Tettigoniidae). MS thesis. Gainesville, FL: University of Florida.

36. Snyder, R. L., Frederick-Hudson, K. H., & Schul, J. (2009). Molecular phylogenetics of the genus Neoconocephalus (Orthoptera, Tettigoniidae) and the evolution of temperate life histories. PloS one, 4(9), e7203. [PMCID: PMC2745656] [PubMed: 19779617]

37. Frederick, K.H. (2013). Investigating an adaptive radiation in temperate Neoconocephalus (Orthoptera: Tettigoniidae). Ph.D. thesis, University of Missouri, Columbia, Missouri.

38. Gilbert, M. T. P., Moore, W., Melchior, L., & Worobey, M. (2007). DNA extraction from dry museum beetles without conferring external morphological damage. PLoS One, 2(3), e272. [PMCID: PMC1803022] [PubMed: 17342206]

39. Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., Rozen, S.G. (2012) Primer3 - new capabilities and interfaces. Nucleic Acids Research 40(15):e115. [PMCID: PMC3424584] [PubMed: 22730293]

40. Koressaar T, Remm M (2007) Enhancements and modifications of primer design program Primer3 Bioinformatics 23(10):1289-91. [PubMed: 17379693]

41. Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C. & Thierer, T. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data.Bioinformatics, 28(12), 1647-1649. [PMCID: PMC3371832] [PubMed: 22543367]

42. Posada, D. (2008). jModelTest: phylogenetic model averaging. Molecular biology and evolution, 25(7), 1253-1256. [PubMed: 18397919]

43. Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular biology and evolution,29(8), 1969-1973. [PMCID: PMC3408070] [PubMed: 22367748]

44. Kingman, J. F. (1982). On the genealogy of large populations. Journal of Applied Probability, 27-43.

45. Rambaut, A., & Drummond, A. J. (2007). Tracer v1.5. http.beast.bio.ed.ac.uk/Tracer.

46. Drummond, A. J., Ho, S. Y., Phillips, M. J., & Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. PLoS biology, 4(5), 699. [PMCID: PMC1395354] [PubMed: 16683862]

47. Reimer, P.J., Bard, E., Bayliss, A., Beck, J.W., Blackwell, P.G., Ramsey, C.B., Buck, C.E., Cheng, H., Edwards, R.L., Friedrich, M. & Grootes, P. M. (2013). IntCal13 and Marine13 radiocarbon age calibration curves 0–50,000 years cal BP. Radiocarbon, 55(4), 1869-1887.

48. Ho, S. Y. (2007). Calibrating molecular estimates of substitution rates and divergence times in birds. Journal of Avian Biology, 38(4), 409-414.

49. Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular biology and evolution, 29(8), 1969-1973. [PMCID: PMC3408070] [PubMed: 22367748]

50. Rambaut, A., & Drummond, A. J. (2013). TreeAnnotator v1. 7.0. University of Edinburgh, Institute of Evolutionary Biology.

51. Heath, T. A. (2015). Divergence Time Estimation using BEAST v2.

52. Shapiro, B., Drummond, A.J., Rambaut, A., Wilson, M.C., Matheus, P.E., Sher, A.V., Pybus, O.G., Gilbert, M.T.P., Barnes, I., Binladen, J. & Willerslev, E. (2004). Rise and fall of the Beringian steppe bison. Science, 306(5701), 1561-1565. [PubMed: 15567864]

53. Penny, D. (2005). Evolutionary biology: relativity for molecular clocks. Nature, 436(7048), 183-184. [PubMed: 16015312]

54. Ho, S. Y., Phillips, M. J., Cooper, A., & Drummond, A. J. (2005). Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. Molecular biology and evolution, 22(7), 1561-1568. [PubMed: 15814826]

55. Brower, A. V. (1994). Rapid morphological radiation and convergence among races of the butterfly Heliconius erato inferred from patterns of mitochondrial DNA evolution. Proceedings of the National Academy of Sciences, 91(14), 6491-6495. [PMCID: PMC44228] [PubMed: 8022810]

56. Knowles, L. L. (2000). Tests of Pleistocene speciation in montane grasshoppers (genus Melanoplus) from the sky islands of western North America. Evolution, 54(4), 1337-1348. [PubMed: 11005300]

57. Avise, J. C. (2000). Phylogeography: the history and formation of species. Harvard university press.

58. Brooks, D. R., & McLennan, D. A. (2012). The nature of diversity: an evolutionary voyage of discovery. University of Chicago Press.

59. Andersen, J. C., & Mills, N. J. (2012). DNA extraction from museum specimens of parasitic Hymenoptera. PloS one, 7(10), e45549. [PMCID: PMC3471897] [PubMed: 23077493]

60. Song, H., Buhay, J. E., Whiting, M. F., & Crandall, K. A. (2008). Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. Proceedings of the national academy of sciences, 105(36), 13486-13491. [PMCID: PMC2527351] [PubMed: 18757756]

61. Rasmussen, M., Li, Y., Lindgreen, S., Pedersen, J.S., Albrechtsen, A., Moltke, I., Metspalu, M., Metspalu, E., Kivisild, T., Gupta, R. & Bertalan, M. (2010). Ancient human genome sequence of an extinct Palaeo-Eskimo. Nature, 463(7282), 757-762. [PMCID: PMC3951495] [PubMed: 20148029]

62. Zedane, L., Hong-Wa, C., Murienne, J., Jeziorski, C., Baldwin, B. G., & Besnard, G. (2016). Museomics illuminate the history of an extinct, paleoendemic plant lineage (Hesperelaea, Oleaceae) known from an 1875 collection from Guadalupe

Island, Mexico. Biological Journal of the Linnean Society, 117(1), 44-57.

63. Dentinger, B., Gaya, E., O'Brien, H., Suz, L.M., Lachlan, R., Díaz-Valderrama, J.R., Koch, R.A. & Aime, M. C. (2016). Tales from the crypt: genome mining from fungarium specimens improves resolution of the mushroom tree of life. Biological Journal of the Linnean Society, 117(1), 11-32.

64. Staats, M., Erkens, R.H., van de Vossenberg, B., Wieringa, J.J., Kraaijeveld, K., Stielow, B., Geml, J., Richardson, J.E. & Bakker, F. T. (2013). Genomic treasure troves: complete genome sequencing of herbarium and insect museum specimens. PLoS One, 8(7), e69189. [PMCID: PMC3726723] [PubMed: 23922691]

65. Bakker, F.T., Lei, D., Yu, J., Mohammadin, S., Wei, Z., Kerke, S., Gravendeel, B., Nieuwenhuis, M., Staats, M., Alquezar-Planas, D.E. & Holmer, R. (2016). Herbarium genomics: plastome sequence assembly from a range of herbarium specimens using an Iterative Organelle Genome Assembly pipeline. Biological Journal of the Linnean Society, 117(1), 33-43.

66. Timmermans, M. J., Viberg, C., Martin, G., Hopkins, K., & Vogler, A. P. (2016). Rapid assembly of taxonomically validated mitochondrial genomes from historical insect collections. Biological Journal of the Linnean Society,117(1), 83-95.

67. Besnard, G., Christin, P. A., Malé, P. J. G., Lhuillier, E., Lauzeral, C., Coissac, E., & Vorontsova, M. S. (2014). From museums to genomics: old herbarium specimens shed light on a C3 to C4 transition. Journal of experimental botany, eru395. [PubMed: 25258360]

68. Heyduk, K., Trapnell, D. W., Barrett, C. F., & Leebens-Mack, J. (2016). Phylogenomic analyses of species relationships in the genus Sabal (Arecaceae) using targeted sequence capture. Biological Journal of the Linnean Society, 117(1), 106-120.

69. Debruyne, R., Chu, G., King, C.E., Bos, K., Kuch, M., Schwarz, C., Szpak, P., Gröcke, D.R., Matheus, P., Zazula, G. & Guthrie, D. (2008). Out of America: ancient DNA evidence for a new world origin of late quaternary woolly mammoths. Current Biology, 18(17), 1320-1326. [PubMed: 18771918]

70. Fulton, T. L., Norris, R. W., Graham, R. W., Semken, H. A., & Shapiro, B. (2013). Ancient DNA supports southern survival of Richardson's collared lemming (Dicrostonyx richardsoni) during the last glacial maximum.Molecular ecology, 22(9), 2540-2548. [PubMed: 23495672]

71. Wisely, S. M., Maldonado, J. E., & Fleische, R. C. (2004). A technique for sampling ancient DNA that minimizes damage to museum specimens. Conservation Genetics, 5(1), 105-107.

72. Seguin-Orlando, A., Hoover, C.A., Vasiliev, S.K., Ovodov, N.D., Shapiro, B., Cooper, A., Rubin, E.M., Willerslev, E. & Johnson, P. L. (2013). Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. Nature, 499(7456), 74-78. [PubMed: 23803765]

73. Burrell, A. S., Disotell, T. R., & Bergey, C. M. (2015). The use of museum specimens with high-throughput DNA sequencers. Journal of human evolution, 79, 35-44. [PMCID: PMC4312722] [PubMed: 25532801]

74. Der Sarkissian, C., Allentoft, M.E., Ávila-Arcos, M.C., Barnett, R., Campos, P.F., Cappellini, E., Ermini, L., Fernández, R., da Fonseca, R., Ginolhac, A. & Hansen, A. J. (2015). Ancient genomics. Philosophical Transactions of the Royal Society of London B: Biological Sciences, 370(1660), 20130387. [PMCID: PMC4275894] [PubMed: 25487338]

75. Dentinger, B. T., Gaya, E., O'Brien, H., Suz, L. M., Lachlan, R., Díaz-Valderrama, J. R., ... & Aime, M. C. (2015). Tales from the crypt: genome mining from fungarium specimens improves resolution of the mushroom tree of life. Biological Journal of the Linnean Society, 117(1), 11-32.

76. Thomsen, Philip Francis, Scott Elias, M. Thomas P. Gilbert, James Haile, Kasper Munch, Svetlana Kuzmina, Duane G. Froese, Andrei Sher, Richard N. Holdaway, & Willerslev, E. (2009). Non-destructive sampling of ancient insect DNA.PLoS

One, 4(4), e5048. [PMCID: PMC2660418] [PubMed: 19337382]

77. Wolinsky, H. (2010). History in a single hair. EMBO reports, 11(6), 427-430. [PMCID: PMC2892323] [PubMed: 20514042]

78. Tin, M. M. Y., Economo, E. P., & Mikheyev, A. S. (2014). Sequencing degraded DNA from non-destructively sampled museum specimens for RAD-tagging and low-coverage shotgun phylogenetics. PloS one, 9(5), e96793. [PMCID: PMC4020769] [PubMed: 24828244]



Historic collection sites for N. lyristes overlaid with hypothesized range.

Sites are modified from ^{34,35} based on literature and collection records. The collection localities of museum samples used in this study are indicated in red.

Table 1: Museum specimen list

N. lyristes pinned specimens obtained from the Hebard collection at the Academy of Natural Sciences of Drexel University. Included is all relevant data from specimen label, as well as the ambiguities present in final sequences. (*) denotes samples removed from analysis for failed amplification or excess ambiguity.

Study reference #	Locality	Collection Date	Collected/ID by	Ambiguities (#/743 bp)
m001	Cape May Court House, NJ	1914	Hebard	0
m002	Cape May Court House, NJ	1914	Hebard	1
m003	Cape May Court House, NJ	1914	Hebard	0
m004	Cape May Court House, NJ	1914	Hebard	1
m005	Cape May Court House, NJ	1914	Hebard	0
m006	Cape May Court House, NJ	1914	Hebard	24
m007	Cedar Swamp, OH	1929	Unknown	N/A*
m008	Cedar Swamp, OH	1932	Edward S. Thomas	25
m009	Cedar Swamp, OH	1932	Edward S. Thomas	0
m010	Chicago, IL (Beach IL)	1906	Unknown	2
m011	Chicago, IL (S. of Jackson Park)	1905	Unknown	18
m012	Chicago, IL (S. of Jackson Park)	1905	Unknown	15
m013	Whitesbog, NJ	1923	Det. D.C. Rentz (1974)	0
m014	Whitesbog, NJ	1923	H. Fox	0
m015	Whitesbog, NJ	1923	Unknown	0
m016	Whitesbog, NJ	1923	H. Fox	0
m017	Whitesbog, NJ	1923	Unknown	103*
m018	Whitesbog, NJ	1923	Unknown	1

Consensus tree from coalescence analysis using a strict molecular clock model and Midwest clade constrained to monophyly



Nodes possessing <0.85 posterior probabilities were collapsed. Red star represents the Mohawk-Hudson Divide, with the prior of the node age set to a normal distribution with a mean age of 10,739.5 cal BP. The Midwest specimen m010 fell outside of the Midwest clade prior to constraining the group to monophyly.

Distribution of the standard deviation rates from relaxed clock analysis



Includes data from ten combined runs (twenty million generations sampled every two thousand trees) using a relaxed molecular clock model. Units for the clock rate are in substitutions per site per million years. The distribution strongly abuts zero, indicating support for a strict molecular clock $\frac{51}{2}$.

Estimates of evolutionary rates (%/myr) plotted against calibration age (myr)



The black data points were obtained directly from ^{20,21}. The red point represents the mutation rate estimate from this study. Note both axes are in log scale.

Supplemental Table 1: Table of primers designed for amplification of N. lyristes COI sequences

Primers were designed from reference sequences of extant *N. lyristes*, *N. bivocatus*, and *N. robustus*.

Primer name	Primer sequence
lyF68 (forward)	5'-GGA ATT GCA CAT GCT GGA GC-3'
lyR197 (reverse)	5'-GTG ATA TTC CTG GGG CAC GT-3'
lyF187 (forward)	5'-ACG TGC CCC AGG AAT ATC AC-3'
lyR336 (reverse)	5'-CCG GCA GGA TCA AAG AAT GA-3'
lyF317 (forward)	5'-TCA TTC TTT GAT CCT GCC GGA-3'
lyR466 (reverse)	5'-GGC TTC CTT TTT CCC ACT TTC T-3'
lyF440 (forward)	5'-AGT CAA GAA AGT GGR AAA AAG GA-3'
lyR589 (reverse)	5'-AGC TGA AGT AAA ATA RGC TCG TG-3'
lyF545 (forward)	5'-ACA GTA GGA ATG GAT GTT GAT ACA C-3'
lyR694 (reverse)	5'-GCC TAG AGC TCA TAA AAG GGA AG-3'
lyF666 (forward)	5'-ACA GTC CTT CCC TTT TAT GAG CT-3'
lyR811 (reverse)	5'-AGA TAG AAC ATA ATG GAA ATG GGC T-3'



Supplemental Fig. 1: Consensus tree using a strict molecular clock and the Midwest clade unconstrained. Node values represent posterior probabilities calculated from eighteen million total trees. Red taxa represent Midwest samples and black taxa Atlantic samples.