Has past climate change affected cold-specialized species differentially through space and time?

Journal:	*Systematic Entomology*
Manuscript ID	SYEN-2018-06-059.R2
Wiley - Manuscript type:	Original Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Schoville, Sean; University of Wisconsin-Madison, Entomology Bougie, Tierney; University of Wisconsin-Madison, Entomology; San Diego State University, Department of Biology; University of California Riverside, Department of Evolution, Ecology, and Organismal Biology Dudko, Roman; Institute of Systematics and Ecology of Animals, Siberian Branch of the Russian Academy of Sciences Medeiros, Matthew; Urban School of San Francisco; University of California, Berkeley, Department of Integrative Biology
Taxon:	Small orders, Insecta
Topic:	Phylogenomics, Biogeography, Diversity
Abstract:	Quaternary climate change has been strongly linked to distributional shifts and recent species diversification. Montane species, in particular, have experienced enhanced isolation and rapid genetic divergence during glacial fluctuations, and these processes have resulted in a disproportionate number of neo-endemic species forming in high elevation habitats. In temperate montane environments, a general model of alpine population history is well supported, where cold-specialized species track favorable climate conditions downslope during glacial episodes and upslope during warmer interglacial periods, which leads to a climate-driven population or species diversification pump. However, it remains unclear how geography mediates distributional changes and whether certain episodes of glacial history have differentially impacted rates of diversification. We address these questions by examining phylogenomic data in a North American clade of flightless, cold-specialized insects, the ice crawlers (Insecta: Grylloblattodea: Grylloblattidae: Grylloblatta). These low vagility organisms have the potential to reveal highly localized refugial areas and patterns of spatial recolonization, as well as a longer history of in situ diversification. Using continuous phylogeographic analysis of species groups, we show that all species tend to retreat to nearby low elevation habitats across western North America during episodes of glaciation, but species at high latitude exhibit larger distributional shifts. Lineage diversification was examined over the course of the Neogene and Quaternary periods, with statistical analysis supporting a direct association between climate variation and diversification rate. Major increases in lineage diversification appear correlated to warm and dry periods, rather than extreme glacial events. Finally, we identify

	substantial cryptic diversity among ice crawlers, leading to high endemism across their range. This diversity provides new insights into highly localized glacial refugia for cold-specialized species across western North America.

Has past climate change affected cold-specialized species differentially through space and time?

Sean D. Schoville[1]*, Tierney C. Bougie[1,2,3], Roman Y. Dudko[4], Matthew J. Medeiros[5,6]

1. University of Wisconsin-Madison, Department of Entomology, Madison, WI, U.S.A.
2. San Diego State University, Department of Biology, San Diego, CA, U.S.A.
3. University of California Riverside, Department of Evolution, Ecology, and Organismal Biology, Riverside, CA, U.S.A.
4. Institute of Systematics and Ecology of Animals, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
5. Urban School of San Francisco, San Francisco, CA, USA
6. Department of Integrative Biology, University of California, Berkeley, CA, USA

*Corresponding author: Sean Schoville

Postal address: University of Wisconsin-Madison, Department of Entomology, 1630 Linden Drive, 637 Russell Labs, Madison, WI 53706, U.S.A.

Phone: 608-262-2956 Fax: 608-262-3322

E-mail: sean.schoville@wisc.edu

Abstract

Quaternary climate change has been strongly linked to distributional shifts and recent species diversification. Montane species, in particular, have experienced enhanced isolation and rapid genetic divergence during glacial fluctuations, and these processes have resulted in a disproportionate number of neo-endemic species forming in high elevation habitats. In temperate montane environments, a general model of alpine population history is well supported, where cold-specialized species track favorable climate conditions downslope during glacial episodes and upslope during warmer interglacial periods, which leads to a climate-driven population or species diversification pump. However, it remains unclear how geography mediates distributional changes and whether certain episodes of glacial history have differentially impacted rates of diversification. We address these questions by examining phylogenomic data in a North American clade of flightless, cold-specialized insects, the ice crawlers (Insecta: Grylloblattodea: Grylloblattidae: *Grylloblatta*). These low vagility organisms have the potential to reveal highly localized refugia and patterns of spatial recolonization, as well as a longer history of in situ diversification. Using continuous phylogeographic analysis of species groups, we show that all species tend to retreat to nearby low elevation habitats across western North America during episodes of glaciation, but species at high latitude exhibit larger distributional shifts. Lineage diversification was examined over the course of the Neogene and Quaternary periods, with statistical analysis supporting a direct association between climate variation and diversification rate. Major increases in lineage diversification appear correlated to warm and dry periods, rather than extreme glacial events. Finally, we identify substantial cryptic diversity among ice crawlers, leading to

high endemism across their range. This diversity provides new insights into highly localized glacial refugia for cold-specialized species across western North America.

Keywords: alpine; phylogenomics; species-pump; glaciation; cryptic species; endemism

Introduction

Climate-driven shifts in species' distributions, in conjunction with the highly variable topography found within mountain ranges and the large geographical distances among mountain ranges, are hypothesized to have increased isolation, lead to rapid genetic divergence, and given rise to the disproportionate number of endemic species currently found at high elevation (Steinbauer et al., 2016). There is widespread support for a general model of alpine population history in temperate regions (Schoville and Roderick, 2009), from both genetic and paleontological data (Birks, 2008; DeChaine and Martin, 2005; Frenzel, 2005; Galbreath et al., 2009; Knowles and Carstens, 2007), wherein cold-specialized species track favorable climate conditions downslope during glacial episodes and upslope during warmer interglacial periods. Across a wide variety of taxa and in alpine habitats globally, past climate change has led to rapid lineage diversification and the formation of new species (Buckley and Simon, 2007; Comes and Kadereit, 2003; DeChaine et al., 2014; Hedin et al., 2015; Knowles, 2000; Schoville and Roderick, 2010), exemplifying the so-called “Pleistocene species pump” (Schoville et al., 2012a; Terborgh, 1992). Speciation is not the only outcome, however, as highly dispersive species experience secondary contact during climate fluctuations, leading to admixture of divergent genomes that formed during periods of isolation (e.g. Galbreath et al., 2010; Gompert et al., 2014; Schneeweiss et al., 2017; Sim et al., 2016), or the complete genetic homogenization of populations (Schönswetter et al., 2008). Despite advances in understanding how past climate has driven population and species diversification, very little is known about the dynamics of cold-specialized species movements at fine geographical scales or over multiple glacial cycles of the Pliocene-

Pleistocene epochs. It is generally held that cold-specialized species expand their distribution during glacial maxima (Stewart et al., 2010b), but this is likely highly dependent on species biology and local geography. Are population contraction and expansion patterns similar across montane regions, and can we infer the extent of geographical movement? Is climate-driven diversification constant through the Pliocene-Pleistocene, or dependent on extreme events during this period? A comparative framework for assessing these questions would provide important insight into how climate change alters species endemism, population genetic diversity and geographical distributions over time.

In western North America, glacial ice associated with the Cordilleran ice sheet and alpine glaciers displaced most high elevation habitat (Kaufman et al., 2004), pushing cold-specialized taxa into lower elevations and to the edges of periglacial habitat (Hewitt, 1996; Pielou, 2008). While some species expanded their geographical range during these events (Schoville and Roderick, 2009), patterns of genetic structure often suggest populations remained highly localized due to the presence of geographical barriers. Comparisons of geological and biological evidence, including phylogeographic patterns across taxa, indicate that multiple ice-free areas (see **Fig. 1**) existed in the Cascade/Coast Range, northern Great Basin, Rocky Mountains, Sierra Nevada, and Klamath-Trinity-Siskiyou Mountains (Brunsfeld et al., 2007; Gavin et al., 2014; Loehr et al., 2006; Roberts and Hamann, 2015; Shafer et al., 2010; Soltis et al., 1997). However, these areas are geographically complex and many species studied are widely distributed with only shallow genetic differentiation (Roberts and Hamann, 2015; Shafer et al., 2010). Low vagility organisms have the potential to increase the resolution of where refugia are

located because of their in situ persistence in such areas (Ashcroft, 2010; Martínez-Solano et al., 2007), thereby increasing our understanding of how local topography influences patterns of spatial recolonization and providing a deeper history of evolutionary responses to climate change given their more limited opportunity for genetic admixture (Rovito and Schoville, 2017; Rovito, 2010; Schoville and Roderick, 2010). Here we focus on a flightless, cold-specialized insect taxon, sampled extensively across montane western North America, to compare regional and temporal responses of lineages to glaciation.

Grylloblattids (Insecta: Grylloblattodea: Grylloblattidae) are montane insects (see **Fig. 2**) that inhabit parts of eastern Asia, the United States, and Canada (Schoville, 2014). Species in the genus *Grylloblatta* Walker (1914), commonly called "ice crawlers," are confined to western North America where they are found associated with snow in alpine talus fields, or at low elevations in caves and lava tubes that retain seasonal ice (Kamp, 1970; Schoville and Graening, 2013). Ice crawlers prefer to forage at cold temperatures (near 0° Celsius), and have a narrow thermal tolerance that is highly conserved across lineages (Schoville et al., 2015). Due to their rarity in museum collections, and the lack of adult males for descriptive taxonomy, species descriptions have been slow (Schoville, 2014). However, genetic studies have found that there are many locally endemic, undescribed *Grylloblatta* lineages, which are recognizable based on high levels of genetic divergence and reciprocal monophyly (Jarvis and Whiting, 2006; Marshall and Lytle, 2015; Schoville and Roderick, 2010). Local populations tend to be small and have limited genetic diversity, yet there is extensive genetic differentiation among nearby populations. Furthermore, divergence times of California taxa were found to be

compatible with the timing of previous glacial episodes and to support the species pump model (Schoville and Roderick, 2010). Thus, the biogeography of *Grylloblatta* serves as an example of cold-specialized species' responses to glacial advances and retreats, but at the fine-scale resolution of major drainage basins rather than the coarse scale of an entire mountain range.

Building upon prior knowledge, the goal of this study is to test 1) whether spatial responses to glacial events are similar and shared among lineages of *Grylloblatta* in diverse regions across its range in western North America and 2) if the timing of lineage diversification covaries with climate change throughout the Pliocene-Pleistocene periods. In addition, we 3) assess genomic data for evidence of cryptic ice crawler species and geographical patterns of endemism, in order to identify putative glacial refugia. To accomplish these goals, we comprehensively sampled populations throughout the known range of *Grylloblatta* and analyzed sequence data from 322 nuclear loci and one mitochondrial gene to develop a unified phylogenomic analysis of the genus.

Material and methods

Collection of samples

A total of 376 *Grylloblatta* individuals were collected across 108 different sites ranging from the west coast of the United States in the Cascade Range and Sierra Nevada Mountains to the Central and Northern Rocky Mountains (see **Fig. 3A** and **Supplementary Table S1**). In addition, samples of Asian out-group species were included from South Korea (*Namkuniga* sp. and *Galloisiana* sp.), Japan (*Galloisiana* sp.), and Russia (*Grylloblattella pravdini*, *Grylloblattella sayanensis*, and *Grylloblattella* sp.).

Most of these sites were sampled from May-August of 2012-2015, with additional specimens provided by collectors, and an extensive survey effort by the Royal British Columbia Museum. Sample size per site ranged from 1-35 individuals. The samples were collected in glass vials with 100% ethanol and labeled according to site of retrieval and year of collection. All samples were obtained with permission: Banff National Park (LL-2014-16218), California Department of Fish and Wildlife (006977), Glacier National Park (GLAC-00210), Lassen Volcanic National Park (LAVO-01880), Lava Beds National Monument (LABE-00053), Mt. Rainier National Park (MORA-00185), North Cascade National Park (NOCA-00005), Olympic National Park (OLYM-00264), Oregon Caves National Monument (ORCA-00030), Yosemite National Park (YOSE-00234), Sequoia and Kings National Park (SEKI-00091), Stone Mountain Provincial Park (107207), and Washington State Department of Fish and Wildlife (14-116, Schoville 154-184). In addition, mitochondrial data from previous studies (Jarvis and Whiting, 2006; Marshall and Lytle, 2015; Schoville, 2012; Schoville and Roderick, 2010) were downloaded from GenBank and included in this analysis (**Table S1**).

Genomic data collection and mitochondrial sequencing

Genomic DNA was extracted from adult individuals by using one hind leg, multiple legs for late-stage nymphs, and the full body of 1st-3rd instar nymphs. DNeasy Blood and Tissue Kits (Qiagen, Valencia, CA, USA) were used to complete the extraction according to the manufacturer's protocol, except with reduced elution volume. Using PCR, a region of the mitochondrial cytochrome oxidase subunit II (COII) gene was amplified using primers COII-F-Leucine and COII-R-Lysine (Svenson and Whiting,

2004). A subset of individuals yielded nuclear pseudogene (numt) sequences with internal stop codons. Following standard protocols (Calvignac et al., 2011), diluted DNA and a much longer region was amplified using primers C1-J2756 and C2-N3665 (Simon et al., 2006) for this subset of individuals, which encompasses part of the adjacent COI locus and all of the tRNA-Leucine gene. Additionally, we targeted co-occurring individuals with divergent mitochondrial sequences for nuclear DNA sequencing (see below) to ensure mitochondrial variation at sites reflected nuclear variation as well. Resulting sequences were validated by amino acid translation and used to replace the numt sequences. PCR products were treated with EXOSAP-IT and sequenced in both directions by Sanger sequencing at the University of Wisconsin-Madison Biotechnology Center. The COII sequences were manually edited and aligned using Geneious v9.1.2 (Kearse et al., 2012). These data were deposited in GenBank (Accessions MG824747 - MG824860, **Table S1**).

Targeted resequencing of the nuclear genome

Transcriptomic data from 12 Grylloblattidae taxa (S. Schoville unpublished data) were screened for single-copy orthologues using HAMSTRv13.2.3 (Ebersberger et al., 2009). Custom myBaits probes (Arbor Biosciences, Ann Arbor, MI) were generated for 322 protein-coding genes (**Supplementary Table S2**), representing ~335kb of sequence data. Probes were designed as 120bp

sequences and tiled at 2x coverage, for a total of 19,107 probes. A total of 96 individuals across 86 sites (including 6 out-groups) were chosen for targeted bead capture (**Supplementary Table S3**), with several individuals sequenced from sites with divergent mitochondrial haplotypes. The 96 samples were sheared to 300bp using a Diagenode Bioruptor Standard sonication device, with a modified shearing time of 25 minutes. The size range of sheared DNA was confirmed using an Agilent Bioanalyzer 2100 DNA chip. DNA libraries were created using the NEBNext DNA Library Prep Reagent Set for Illumina (New England BioLabs) with modifications to account for lower input DNA concentration than recommended. Each sample was embedded with a unique set of dual-index primers via PCR to ensure that each individual could be differentiated when pooled in a library. Each library was assessed on the Bioanalyzer for a final quality check before proceeding to target sequence capture. The 96 libraries were grouped into 12 pools of eight samples each. Libraries were grouped in each pool to have a similar average concentration, while also maintaining the minimum total amount of DNA per pool for optimal

sequence capture. The myBaits protocol was followed for in-solution sequence capture. The last PCR step was modified to include more cycles to account for lower than recommended levels of input library DNA added to the 12 pools. Following PCR, the pools were quantified using a high sensitivity dsDNA kit on a Qubit Fluorometer (Life Technologies). To prepare for sequencing, 1-2µl of each of the 12 pools was added to a single tube so that the total volume reached 15µl with a concentration of 5ng/µl. The 96 pooled libraries were submitted to the UW Madison Biotechnology Center for sequencing on a single lane of the Illumina MiSeq platform with a sequence length of 2x250bp.

The sequences were demultiplexed and assessed for quality using FastQC v0.11.5 (Babraham Bioinformatics, 2011). FastQC was specifically used to check for overrepresented index and adapter sequences that may not have been completely removed during demultiplexing, and Trimmomatic v0.32 (Bolger et al., 2014) was used to remove these overrepresented sequences as well as low quality bases. The paired-end function was used to produce an output of four total files per individual, two for paired forward and reverse reads and two for unpaired forward and reverse reads. Reads were mapped to consensus sequences of the 322 protein coding genes using CLC Genomics Workbench 9 (Qiagen, Valencia, CA, USA). Subsequently, contigs of all individuals were aligned in Geneious and exported as FASTA files. These data were deposited in GenBank (BioProject PRJNA488740, Table S3).

Population genetic and phylogenetic analysis of mitochondrial data

To characterize the genetic diversity of local populations of *Grylloblatta*, population genetic summary statistics were calculated for the mitochondrial COII locus in a set of populations consisting of ten or more individuals. Genetic diversity measures focused on nucleotide diversity, θ_s and θ_π, while the Tajima's *D* and *Fu's* F_S neutrality tests were used to examine evidence of demographic change. These statistics were calculated in Arlequin v3.5 (Excoffier and Lischer, 2010). Phylogenetic analysis was used to characterize genetic relationships among all sampled *Grylloblatta* individuals. Unique COII haplotypes were identified, manually aligned and used to construct a Bayesian phylogeny in the software BEAST v2.4.0 (Bouckaert et al., 2014). In order to set the substitution model prior in BEAST, the best evolutionary model of nucleotide substitutions was estimated using maximum likelihood in MEGA v6.06 (Tamura et al., 2013). Model selection by MEGA was performed for the complete gene sequence, and separately for codon positions 1+2 and 3. The evolutionary model for the complete data was estimated as GTR+G+I, while codon positions 1+2 was estimated as T92+G+I and the third position was estimated as GTR+G+I. We used an iterative approach to choose the best substitution prior in BEAST, first using the Akaike information criterion (AICM) to compare likelihood values for a model with and without data partitioning (BEAST does not have the T92 model, so the HKY model was chosen for the codon 1+2 partition). While AICM has been shown to have limitations as an estimator of the marginal likelihood for model selection, it outperforms the harmonic mean estimator and has reduced computational demands relative to path sampling (Baele et al., 2012).

Similarly, to evaluate the tree prior in BEAST, likelihood values of runs with the Yule and Birth-Death tree prior for the unpartitioned dataset were compared using AICM (**Supplementary Table S4**). These initial models were run assuming a strict clock model for 1 billion steps of the MCMC chain, with samples taken every 100,000 steps. TRACER (Rambaut and Drummond, 2009a) was used to assess stationarity of the MCMC chain, ensure that the effective sample size (ESS) of the parameters was high (>200), and to compare models. The unpartitioned dataset was chosen with the Birth-Death tree prior for downstream analyses. Following these initial BEAST runs, we compared a strict clock model to a relaxed log-normal and a relaxed exponential model. In all models, the MCMC chain was set to 200 million steps to ensure enough search time and the data was stored every 200,000 steps. Models were compared using the Harmonic Mean estimator, AICM, and path sampling (Baele et al., 2012). The path sampling algorithm was run with a beta value of 0.3 for 8 steps. After selecting the best performing model, a second run was conducted with the same length and thinning interval to ensure convergence of the MCMC chain and the runs were combined to produce

a final dataset. A maximum clade credibility tree was estimated and a tree image was generated in FigTree v1.4.2 (http://tree.bio.ed.ac.uk/software/figtree).

Phylogenomic analysis of nuclear data

We used both a concatenation and a species tree approach to generate a phylogeny based on the 322 nuclear genes. As missing data can introduce bias in phylogenetic reconstruction (Wiens, 2006), we resampled the full dataset holding the number of individuals constant, allowing for different levels of missing genes: <30%, <20%, <10%, and less than 5%. BEAST v2.4.0 was used to analyze concatenated datasets for each level of missing data, using the HKY substitution model for each gene partition, a strict clock model and the Yule tree prior. The MCMC chain length was set to 10 million steps to ensure enough search time and the data was stored every 100,000 chains. The output file was examined to ensure MCMC chain stabilization and acceptable ESS values (>150) of most parameters. A maximum clade credibility tree was estimated for each dataset in the BEAST TREEANNOTATOR module, with a 20% burnin. To compare difference among trees, we used the Kendall Colijn metric (Kendall and Colijn, 2016) of pairwise differences among tree topologies. Trees can be compared by counting the number of edges between

two tips and their most recent common ancestor, or based on the summed branch lengths between the two tips. The R package *treespace* (Jombart et al., 2017) was used to compare trees based on both topology and branch lengths using the multidist function. Additionally, membership of the tips in major clades was assessed by comparing the distance to most recent common ancestor of major clades among trees using the relatedTreeDist function (Kendall et al., 2018).

Additionally, a species tree approach was used to account for the independent history of each gene and the possibility of incomplete lineage sorting leading to statistical inconsistency in the concatenated dataset. Individual gene trees were estimated in BEAST for the 80% dataset (191 genes) and used to estimate a species tree in the program ASTRAL-II v4.10.12 (Mirarab and Warnow, 2015). This method applies a multi-species coalescent framework to infer a phylogeny, given the input gene tree bipartitions found in the gene set. Support for each quartet is measured as a local posterior probability value and branch lengths are measured at internal nodes in terms of coalescent units

(small values indicate a large amount of gene tree discordance).

Spatial reconstruction of responses to glaciation

In order to reconstruct spatial responses to glacial events, we estimated the ancestral ranges for well-supported genetic lineages (both known species and putative unidentified species) using geographical data and the COII gene sequences. The COII dataset, while based on a single gene, provides the most extensive geographical sampling within each lineage. Only a subset of lineages were analyzed, because the analysis requires samples from multiple geographical locations. Using the continuous phylogeographic modeling approach implemented in BEAST (Lemey et al., 2010), samples in a coalescent tree disperse from sample localities backwards in time under a relaxed random walk model, and the probable ancestral range is reconstructed for the lineage. BEAST was run with the HKY gamma site model, a strict clock rate of 1.5% (log-normal prior and SD 0.061), and coalescent skyline prior. Similar values for the molecular clock were supported in prior analyses of *Grylloblatta* species (Schoville and Roderick, 2010; Schoville et al., 2013). Two chains were run for each lineage for 10 million steps each, and the output was examined in TRACER v1.6 (Rambaut and Drummond, 2009b) to ensure MCMC chain stabilization and high ESS values. The first 10% of the trees were discarded as burnin and the two chains were combined to estimate the tree topology, the ancestral area at each node, and the root age of each lineage. SpreaD3 (Bielejec et al., 2016) was used to render the

ancestral area reconstruction into graphical form. To estimate changes in effective size through time, a Bayesian skyline plot was generated for each lineage assuming a stepwise constant coalescent model (Drummond et al., 2005).

Lineage diversification and climate change

In order to estimate rates of diversification through time, we first estimated branch lengths on the fixed species tree topology in terms of nucleotide substitution rates and rescaled the branches relative to absolute divergence time. Branch lengths were estimated in RAxMLv8.2.10 (Stamatakis, 2014) based on the joint estimate of nucleotide substitution rates across gene partitions. The program r8sv1.70 (Sanderson, 2003) was used to estimate a scaled ultrametric tree (in millions of years, MY), assuming a fixed-age constraint of 25 MY for the divergence between *Galloisiana* lineages sampled from Shikoku, Japan and South Korea. This assumption is supported by geological evidence underlying the separation of the Japanese islands from the Korean peninsula and is also consistent with the insect COI-COII substitution rate of 1.5% used above (Gaunt and Miles, 2002; Schoville and Roderick, 2010; Schoville et al., 2013).

The species tree was pruned to 50 well-supported lineages within *Grylloblatta* and six known missing lineages were included in the total, for a total of 56 lineages. Diversification rates were estimated in six competing models of speciation (birth, or λ) and extinction (death, or μ) in RPANDA (Morlon et al., 2016). The first four models

assumed time-dependent, exponential rates: 1) constant birth and death, 2) a variable rate of birth and a constant rate of death, 3) a constant rate of birth and variable rate of death, and 4) a variable rate of both birth and death. The last two models assumed environmentally-dependent, exponential rates: 5) a constant dependence of the speciation rate on paleotemperature with no extinction, and 6) a constant dependence of the speciation rate on paleotemperature with constant extinction. The best model was selected using corrected Akaike Information Criteria (AICc) as the model with the lowest value.

Results

Population genetic variation in mtDNA

Sequence variation in the COII mitochondrial gene was used to examine genetic diversity at 14 well-sampled populations ($n \geq 10$), which showed very low levels of within population genetic variation (**Table 1**). Six populations had levels of genetic variation close to or equal to zero. The remaining populations, Adams, Jefferson, Lassen Peak, Marble Valley, Spokane, Thompson, Barker Pass, and Susie Lake exhibited low levels of within population variation, with θ_s ranging from 0.002-0.009 per base pair and θ_π ranging from 0.001-0.008 per base pair. Tests for deviation from neutrality revealed statistically significant negative Tajima's *D* values in four populations, indicative of an excess of low frequency polymorphisms relative to expectation and supporting a population expansion (following a bottleneck event) or a selective sweep event. Four populations had significant negative Fu's F_s statistics indicating an excess of alleles compared to expectation, and supporting a recent population expansion or selective

sweep. Only two populations, Mt. Jefferson and Mt. Thielsen, had significant values of both Tajima's *D* and Fu's F_s tests supporting a population expansion.

Phylogenetic relationships and species delimitation

Mitochondrial COII sequences were generated for 376 individuals at 108 different sites across western North America (**Fig. 3A** and **Supplemental Fig. S1** for detail), resulting in 114 unique haplotypes. Based on model selection procedures (**Table S4**), a mitochondrial gene tree was estimated assuming an unpartitioned dataset with the Birth Death tree prior. This yielded a phylogeny with high posterior probability support (>0.95) at most shallow nodes, but some deeper nodes remained poorly supported (<0.90, **Fig. 3B**). Nine highly genetically differentiated lineages geographically co-occur at ten sites (**Supplementary Table S1**), suggesting the presence of multiple species and cryptic taxa at these sites.

The concatenated nuclear gene phylogeny using all 322 loci (208,444 bp total length, with an average of 12% missing base pairs per individual) shows high posterior support at all nodes (**Fig. 3C** and **Fig. 4**). The mitochondrial gene tree and nuclear gene phylogeny have highly concordant groupings of individuals at shallow nodes, as well as broad-scale genetic cohesion among geographically proximal samples. Notable differences among the topologies appear in relationships at deeper nodes. To examine the effect of missing data on the concatenated nuclear phylogeny, we analyzed the full data and data subsets allowing for <30%, <20%, <10%, and less than 5% missing genes per individual, resulting in 322, 266, 191, 77, and 26 loci, respectively. Posterior probabilities remained at >0.95 for the majority of nodes (**Supplemental Fig. S2-S6**), until the

smallest dataset of 26 genes, where 15 nodes had less than 0.95 posterior support. Notably, topological relationships among clades remained stable, although relationships at shallower levels within clades differed among the datasets, suggesting that the decrease in data reduced the ability to delimit monophyletic subclades (that might represent species) and determine their relationships. Pairwise distance among trees using the Kendall Colijn metric suggests that the 5% missing data tree differs substantially (**Supplementary Table S5**), while tip membership in the major clades remains the same across trees (pairwise distance among trees measured from the most recent common ancestors of the major clades is zero). Relationships within the *G. campodeiformis* clade (**Fig. 4**) vary the most among the datasets. The phylogeny with less than 5% missing data had additional changes at deeper nodes in groupings of clades in California, where branch lengths separating major clades are short.

The species tree (**Fig. 5** and **Supplemental Fig. S7** for detail) showed high similarity in the membership of tips in major clades to the full concatenated nuclear phylogeny (pairwise distance among trees measured from the most recent common ancestors of the major clades is zero). However, pairwise distance among trees using the Kendall Colijn metric suggests there are substantial differences in the species tree topology (**Supplementary Table S5**). This appears to be driven largely by rearrangements in the relationships among major clades and outgroups at deeper nodes in the tree. Local posterior probabilities provide strong support (>0.95) for major clades in the species tree, as well as nested clades and some independent lineages. Here, branch lengths are proportional to coalescent units, revealing that many of these nested clades and lineages have been diverging for substantial time relative to their population size. For

example, samples of a group from Washington and southwestern British Columbia are highly differentiated from other *G. campodeiformis* samples and comprise a well-supported reciprocally monophyletic group (**Supplemental Fig. S7)**. Similarly, multiple lineages in the Sierra Nevada suggest the presence of highly differentiated lineages at a relatively small geographical scale (~50 km), and notably five lineages can be found in sympatry with another lineage (**Supplementary Table S3**).

Spatial responses to glacial events

Ancestral ranges were estimated using the COII gene sequences, for 13 well-supported lineages comprising samples from multiple geographical sites. Nearly all major clades in the mitochondrial tree were included in this analysis. The estimated ancestral ranges were highly localized geographically, suggesting a high degree of endemism (shown for ten lineages in **Fig. 6A**, all available separately in **Supplementary Fig. S8-S20**). Ancestral ranges were situated in and around mountain ranges, with one exception. The *G. campodeiformis* clade was broadly distributed across the Northern and Central Rocky Mountains, suggesting a substantial geographical expansion of this clade. There is a general trend of ancestral range size increasing with latitude, evident in those clades from central Oregon (**Supplementary Fig. S15**) through Washington State (**Supplementary Fig. S10, S16 and S18**). Root age estimates ranged from median values of 120,400 years to 1.3 million years, but overlapped in the 95% highest posterior density values (**Supplementary Table S6)**. Bayesian skyline plots showed recent declines in estimated effective population size for all lineages considered (shown for ten lineages in **Fig. 6B**, all available separately in **Supplementary Fig. S8-S20)**. Only two lineages

showed an increase prior to this decline, *G. chandleri* and *G.* sp. Central Sierra Nevada, both from California.

Divergence dating and lineage diversification rates

Diversification rates were fit to six competing models of speciation and extinction rates. The temperature-dependent, no extinction model was favored with the lowest AICc score (**Table 2**). The speciation rate under this model was low ($\lambda = 0.0265$ per million years) and declines towards the present (at a rate of $\Delta\lambda = 0.262$ per million years). A plot of paleotemperature change and the speciation rate through time (**Fig. 7**) shows a large increase in speciation rate around 16.5 to 14 million years ago, and a smaller increase around 7.5 to 3 million years ago.

Discussion

Glacial cycles of the Pliocene-Pleistocene epochs not only impacted the distributions of species (Davis and Shaw, 2001), but caused dramatic changes in genetic diversity, population genetic structure, and rates of lineage diversification (Hewitt, 1996). While species responses to past climate change vary (Stewart et al., 2010a), generalizations are often possible (Avise, 2009), especially within ecotypes (e.g. alpine species; Schmitt, 2007). For cold-specialized montane taxa, glacier expansion caused downslope shifts and enabled range expansion among mountain ranges, while glacier retreat led to upslope shifts and isolation in 'sky islands' (Elias, 1996; Hafner and Sullivan, 1995). Beyond these generalizations, less is known about how local geographical features or temporal variation in the extremity of glacial events has

impacted cold-specialized species' responses to climate change. The ice crawlers (*Grylloblatta* spp.) provide a good model for this question because of their low vagility and deep evolutionary history (Schoville and Roderick, 2010), but also because divergent lineages remain highly conserved in ecology and physiology (Schoville et al., 2015). We provide evidence that responses to glacial events were similar among lineages of *Grylloblatta* throughout western North America, despite variation in local topography and the extent and timing of glaciation. All species have a clear pattern of retreating to nearby low elevation habitats, although species at high latitude do exhibit larger distributional shifts. We also show that the timing of lineage diversification varied throughout the Neogene, with distinct increases in diversification occurring at two time periods predating glacial events. Finally, we provide much finer detail regarding the geographical patterns of endemism and the prevalence of cryptic ice crawler species, which provides insight into the location of montane glacial refugia.

Spatial and demographic responses of ice crawlers to glacial events

Despite extensive variation in local topography and the configuration of mountain ranges, lineages of ice crawlers exhibit similar patterns of recent spatial and demographic change (**Fig. 6**). In contrast, comparisons among taxonomically diverse species in the same region often reveal individualistic responses to glaciation, with varied rates of recolonization and spatial spread (Stewart et al., 2010b). Estimates of the ancestral distribution of ice crawler lineages suggest high local endemism in present-day montane habitats. This general pattern is robust across locations in western mountains, as 13 lineages show similar responses from California to northern Washington. In these focal

lineages, ancestral areas are estimated in close proximity to mountain slopes, supporting the argument that ice crawlers tracked expanding and retreating alpine glaciers (Schoville and Roderick, 2010) through multiple glacial cycles (because root age estimates span a time period that encompasses multiple glaciation events). Furthermore, five of these lineages include low elevation cave populations in their ancestral area estimates, indicating that disjunct cave and alpine populations are still inferred to have arisen from a narrow area of endemism, rather than to have spread broadly across low elevation basins. While the root ages are not the same across lineages, they remain largely consistent in reflecting lineage diversification responses of the Late Pleistocene (95% HPD ranging from 14,200-1,794,300 years before present across lineages).

One lineage deviates from a pattern of narrow endemism, *G. campodeiformis* from the Central and Northern Rocky Mountains. This clade exhibits extensive spatial expansion in a region encompassed by the continental ice sheet during the last glacial maximum. However, the ancestral age estimates reflect the coalescent history across multiple glacial events (in the case of *G. campodeiformis*, the median root age is estimated at 1.3 million years ago). Thus, while many other organisms spread widely during their recolonization of Canadian boreal and arctic habitats (Roberts and Hamann, 2015; Shafer et al., 2010), it is unlikely that poorly dispersing ice crawlers did so in a single event following the last glacial maximum. Singleton substitutions in mitochondrial haplotypes (0.02%) are evident in populations from the Cassiar and Northern Rocky Mountains (>58^{o} latitude). Kamp (1979) argued that these two populations persisted in cryptic refugia during the last glacial maximum, based on morphological differences of these populations (which he designated as *G. campodeiformis nahanni* and *G. c.*

athapaska) compared to those of *G. c. campodeiformis* from the Central Rocky Mountains. Populations of *G. c. campodeiformis* comprise one mitochondrial haplotype throughout the Central Rocky Mountains, supporting a single population expansion from ice-free regions in northern Montana (47-48^{o} latitude) following the retreat of the last glacial maximum (~19,000 years ago; Clark et al., 2009; Locke, 1990; Pierce, 2003), although notably there are two site-specific mitochondrial haplotypes in central and southern Montana (basal to all *G. c. campodeiformis*) that suggest separate glacial refugia in these regions.

We note that there are several limitations in these analyses due to our reliance on mitochondrial DNA. First, as a single non-recombining genetic marker, it is a single representation of genealogical history and might deviate from estimates derived from a larger set of independent genetic markers (Degnan, 1993). Our analyses of demographic and spatial responses should therefore be interpreted with caution. However, one positive aspect of mitochondrial data are their utility (compared to nuclear data) in reconstructing recent events in demographic history due to their high mutation rate and reduced effective population size (Eytan and Hellberg, 2010). Second, the presence of nuclear copies of mitochondrial sequences (numts) also potentially complicate our analyses, as highly divergent sequences could inflate estimates of effective size and spatial movement. When highly divergent sequences were found in particular populations, and when they had stop codons, we assessed whether they were numts. We re-amplified a longer section of the mitochondria and diluted DNA extractions for these samples, following recommendations in Calvignac et al. (2011). While this procedure does not guarantee complete removal of numt sequences (Moulton et al., 2010), our dataset does

seem robust to numt bias. Examination of nuclear genomic data shows strong phylogenetic concordance with the mitochondrial gene tree, which would indicate that numt sequences (if present) would be recently derived as the tree topology reflects lineage divergence.

Lineage diversification in relation to past climate change

Previous research supports the idea that glacial cycles displaced cold-specialized species from alpine habitats (e.g. Hafner and Sullivan, 1995; Rovito and Schoville, 2017; Varga and Schmitt, 2008), and the process of shifting up and downslope lead to demographic bottlenecks (Schoville et al., 2012b). This pattern is evident for the ice crawlers as well, as lineages now occupy previously glaciated landscapes, while demographic change indicates recent population declines in these lineages (**Fig. 6B**). Population genetic variation is low at most alpine sites, suggesting small numbers of founding individuals recolonized these habitats. However, an unresolved question is whether this spatial and demographic response to glaciation can be quantified for previous glacial cycles? Most genetic studies have focused solely on the effects of the last glacial cycle (but see: Schmitt et al., 2005), even though multiple glacial cycles occurred during the Pliocene-Pleistocene epochs and some were more extensive (Lisiecki and Raymo, 2007). One challenge with the ice crawlers is that recent bottlenecks erode the genomic signal of prior demographic events (Nei et al., 1975).

To address this question in our system, we examined whether lineage diversification, as opposed to recent population responses, was correlated with long-term climate fluctuations during the Neogene and Quaternary periods, and asked whether there

were particular episodes that drove diversification patterns. By fitting the *Grylloblatta* species tree to several competing species diversification (birth and death) models, we found the best model was one where lineage diversification was correlated with the paleotemperature record. Furthermore, based on our assumptions about a molecular clock rate, we identified two increases in ice crawler diversification rate, both prior to glaciation events (around 16.5 to 14 million years ago and around 7.5 to 3 million years ago). The first increase coincides with the Miocene climatic optimum (18 to 14 million years ago), a period of warm temperatures and increasing aridity in North America (Böhme, 2003; Eronen et al., 2012). The second increase in diversification coincides with a cooling and drying event in the Late Miocene (Herbert et al., 2016), which preceded the onset of Pliocene glaciations. Following this last episode (and during the Quaternary period), the speciation rate declined towards the present. Thus, rather than support a distinct role for the cold temperatures of glacial cycles in pulses of species diversification, we found that dry periods had a much stronger effect. For cold-specialized insects such as ice crawlers, which rely on subterranean microhabitats during unseasonal conditions, humidity constraints might be more challenging than temperature (Kamp, 1973) and lead to substantial range fragmentation and subsequent genetic isolation (Dudko, 2011).

Cryptic species diversity and endemism

While previous studies provided evidence of substantial genetic structure across the range of *Grylloblatta* (Jarvis and Whiting, 2006; Schoville and Roderick, 2010), these studies were based on limited geographical sampling and genetic data. In this study, we

have shown that phylogenomic data are consistent with extensive genetic divergence among populations at fine spatial resolution. How many of these well supported genetic lineages are cryptic species? Recent efforts (Marshall and Lytle, 2015; Schoville, 2012) to update the taxonomy of *Grylloblatta* found morphological and genetic evidence for the species status of several isolated populations. In these cases, the mitochondrial COII locus provided evidence of substantial genetic divergence (3.7-11% uncorrected pairwise distance) and/or reciprocal monophyly in each species (Schoville and Graening, 2013). If similar genetic criteria were used here, 11 lineages could be considered as cryptic lineages from the mitochondrial data alone. Known species remain well-supported when phylogenomic data are considered (>0.4% average sequence divergence and coalescent branch lengths >2.0). Using these criteria to evaluate the number of cryptic lineages supported in the nuclear phylogenomic dataset, 20 lineages are recognized as possible cryptic species. Combining genetic and geographical data, nine lineages are sympatric at 10 sites (**Supplementary Table S1**), further supporting the presence of cryptic, reproductively-isolated species. Further work on morphological and ecological characteristics should be employed to refine these estimates and provide an updated taxonomy of *Grylloblatta*.

The geographical distribution of genetic lineages is quite limited among ice-crawlers (**Fig. 3A** and **5B**), reflecting a low dispersal ability and a long history in different regions across their range (**Fig. 6A**). What new regions appear to be glacial refuges for cold-specialized species? Multiple refugia in Montana (northern, central and southern), as well as a refugia near Mt. Spokane in eastern Washington are supported by our data. Additionally, some known refugia are subdivided by multiple lineages,

suggesting that there are smaller microrefugia during glacial periods. For example, several lineages have non-overlapping or partially overlapping ranges within the Trinity-Siskiyou Mountains of northern California and southern Oregon, the eastern slope of the Oregon Cascades, and the western slopes of the Sierra Nevada Mountains in California. This suggests substantially more refugia existed in western North American mountains (Brunsfeld et al., 2007; Gavin et al., 2014; Loehr et al., 2006; Roberts and Hamann, 2015; Shafer et al., 2010; Soltis et al., 1997) and provides an important source of knowledge for conservation planning and biogeographical analysis.

Conclusion

Using a large phylogenomic dataset, we demonstrated population responses to glaciation were highly similar across *Grylloblatta* species in North America. Ice crawler species exhibit a clear spatial pattern of retreating to nearby low elevation habitats and recolonizing alpine sites following glacial maxima, despite variation in local topography across this region. We also demonstrated that the timing of lineage diversification in the genus *Grylloblatta* covaries with long-term climate change, but interestingly, speciation increases during global aridity events, rather than occurring during glacial episodes. Finally, we provide genomic evidence that ice crawlers are much more diverse than currently appreciated, with as many as 11-20 lineages representing cryptic species (nine of these occur in sympatry and are genetically distinct from co-occurring lineages). This high diversity reflects persistent narrow endemism across western North America despite repeated glacial events, and thus the geographical patterns of *Grylloblatta* lineages provide novel evidence of glacial microrefugia in Montana, Oregon and California.

Integrating these results with evidence of strong physiological conservatism in thermal tolerance of ice crawlers (Schoville et al., 2015), as well as ongoing changes to ice crawler habitats, suggests genetically isolated and highly divergent lineages are at risk of local extinction. Most notably, ongoing declines in alpine snowpack (Minder, 2010; Mote et al., 2005) and reductions in the mass balance of ice in seasonal ice caves where *Grylloblatta* occur (Kern and Thomas, 2014), suggests both rising temperatures and increasing aridity are threatening populations of ice crawlers. Ice crawler species are found in (and often endemic to) U.S. and Canadian National Parks, where conservation management and stewardship of biological diversity is a mandate. To meet these goals, there is an urgent need to expand our understanding of ice crawler ecology so that we can develop better predictive modeling of climate change impacts and develop conservation management strategies that are responsive to ongoing environmental change.

Acknowledgements

We would like to thank a number of colleagues for contributing specimens for this project: Ming Bai, Robb Bennett, James Bergdahl, Claudia and Darren Copley, Joe Giersch, G.O. Graening, Byung-Woo Kim, Neil Marchington, and Brent McGregor. We are also grateful to the National Park system of the United States and Canada, as well as the Canadian First Nations, for providing permission to conduct our research.

Funding: This research was supported by the National Science Foundation DEB program (1655615), National Geographical Society (8993-11), and Seattle City Lights Wildlife Research Program (2014-03). R. Dudko was supported by the Federal Fundamental

Scientific Research Program for 2013-2020, grant № VI.51.1.5 (AAAA-A16-116121410121-7).

References

Ashcroft, M.B., 2010. Identifying refugia from climate change. J. Biogeogr. 37, 1407-1413.

Avise, J.C., 2009. Phylogeography: retrospect and prospect. J. Biogeogr. 36, 3-15.

Babraham Bioinformatics, 2011. FastQC: a quality control tool for high throughput sequence data. Babraham Institute, Cambridge, United Kingdom.

Baele, G., Lemey, P., Bedford, T., Rambaut, A., Suchard, M.A., Alekseyenko, A.V., 2012. Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. Mol. Biol. Evol. 29, 2157-2167.

Bielejec, F., Baele, G., Vrancken, B., Suchard, M.A., Rambaut, A., Lemey, P., 2016. Spread3: Interactive visualization of spatiotemporal history and trait evolutionary processes. Mol. Biol. Evol. 33, 2167-2169.

Birks, H.H., 2008. The Late-Quaternary history of arctic and alpine plants. Plant Ecology & Diversity 1, 135-146.

Böhme, M., 2003. The Miocene climatic optimum: evidence from ectothermic vertebrates of Central Europe. Palaeogeogr., Palaeoclimatol., Palaeoecol. 195, 389-401.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics, btu170.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., Suchard, M.A., Rambaut, A., Drummond, A.J., 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. PLoS Comput Biol 10, e1003537.

Brunsfeld, S.J., Miller, T.R., Carstens, B.C., 2007. Insights into the biogeography of the Pacific Northwest of North America: evidence from the phylogeography of *Salix melanopsis*. Syst. Bot. 32, 129-139.

Buckley, T.R., Simon, C., 2007. Evolutionary radiation of the cicada genus *Maoricicada* Dugdale (Hemiptera: Cicadoidea) and the origins of the New Zealand alpine biota. Biol. J. Linn. Soc. 91, 419-435.

Calvignac, S., Konecny, L., Malard, F., Douady, C.J., 2011. Preventing the pollution of mitochondrial datasets with nuclear mitochondrial paralogs (numts). Mitochondrion 11, 246-254.

Clark, P.U., Dyke, A.S., Shakun, J.D., Carlson, A.E., Clark, J., Wohlfarth, B., Mitrovica, J.X., Hostetler, S.W., McCabe, A.M., 2009. The last glacial maximum. science 325, 710-714.

Comes, H.P., Kadereit, J.W., 2003. Spatial and temporal patterns in the evolution of the flora of the European alpine system. Taxon 52, 451-462.

Davis, M.B., Shaw, R.G., 2001. Range shifts and adaptive responses to Quaternary climate change. Science 292, 673-679.

DeChaine, E.G., Martin, A.P., 2005. Historical biogeography of two alpine butterflies in the Rocky Mountains: broad-scale concordance and local-scale discordance. J. Biogeogr. 32, 1943-1956.

DeChaine, E.G., Wendling, B.M., Forester, B.R., 2014. Integrating environmental, molecular, and morphological data to unravel an ice-age radiation of arctic-alpine *Campanula* in western North America. Ecology and evolution 4, 3940-3959.

Degnan, S., 1993. The perils of single gene trees—mitochondrial versus single-copy nuclear DNA variation in white-eyes (Aves: Zosteropidae). Mol. Ecol. 2, 219-225.

Drummond, A.J., Rambaut, A., Shapiro, B., Pybus, O.G., 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. Mol. Biol. Evol. 22, 1185-1192.

Dudko, R.Y., 2011. Relict beetles (Coleoptera: Carabidae, Agyrtidae) with Altai-East Asian disjunctive range. Euroasian Entomological Journal 10, 349-360+348+VI.

Ebersberger, I., Strauss, S., von Haeseler, A., 2009. HaMStR: profile hidden markov model based search for orthologs in ESTs. BMC Evol. Biol. 9, 157.

Elias, S.A., 1996. The Ice Age History of National Parks in the Rocky Mountains. Smithsonian Press, Washington D.C.

Eronen, J.T., Fortelius, M., Micheels, A., Portmann, F., Puolamäki, K., Janis, C.M., 2012. Neogene aridification of the Northern Hemisphere. Geology 40, 823-826.

Excoffier, L., Lischer, H.E.L., 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Res. 10, 564-567.

Eytan, R.I., Hellberg, M.E., 2010. Nuclear and mitochondrial sequence data reveal and conceal different demographic histories and population genetic processes in Caribbean reef fishes. Evolution: International Journal of Organic Evolution 64, 3380-3397.

Frenzel, B., 2005. History of flora and vegetation during the Quaternary North America. Progress in Botany 66, 409-440.

Galbreath, K.E., Hafner, D.J., Zamudio, K.R., 2009. When cold is better: climate-driven elevation shifts yield complex patterns of diversificaiton and demography in an alpine specialist (American Pika, *Ochotona princeps*). Evolution 63, 2848-2863.

Galbreath, K.E., Hafner, D.J., Zamudio, K.R., Agnew, K., 2010. Isolation and introgression in the Intermountain West: contrasting gene genealogies reveal the complex biogeographic history of the American pika (*Ochotona princeps*). J. Biogeogr. 37, 344-362.

Gaunt, M.W., Miles, M.A., 2002. An insect molecular clock dates the origin fo the insects and accords with paleontological and biogeographic landmarks. Mol. Biol. Evol. 19, 748-761.

Gavin, D.G., Fitzpatrick, M.C., Gugger, P.F., Heath, K.D., Rodríguez-Sánchez, F., Dobrowski, S.Z., Hampe, A., Hu, F.S., Ashcroft, M.B., Bartlein, P.J., 2014. Climate refugia: joint inference from fossil records, species distribution models and phylogeography. New Phytol. 204, 37-54.

Gompert, Z., Lucas, L.K., Buerkle, C.A., Forister, M.L., Fordyce, J.A., Nice, C.C., 2014. Admixture and the organization of genetic diversity in a butterfly species complex revealed through common and rare genetic variants. Mol. Ecol. 23, 4555-4573.

Hafner, D.J., Sullivan, R.M., 1995. Historical and ecological biogeography of Nearctic pikas (Lagomorpha: Ochotonidae). Journal of Mammology 76, 302-321.

Hedin, M., Carlson, D., Coyle, F., 2015. Sky island diversification meets the multispecies coalescent–divergence in the spruce-fir moss spider (*Microhexura montivaga*, Araneae, Mygalomorphae) on the highest peaks of southern Appalachia. Mol. Ecol. 24, 3467-3484.

Herbert, T.D., Lawrence, K.T., Tzanova, A., Peterson, L.C., Caballero-Gill, R., Kelly, C.S., 2016. Late Miocene global cooling and the rise of modern ecosystems. Nature Geoscience 9, 843.

Hewitt, G.M., 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. Biol. J. Linn. Soc. 58, 247-276.

Jarvis, K.J., Whiting, M.F., 2006. Phylogeny and biogeography of ice crawlers (Insecta: Grylloblattodea) based on six molecular loci: Designating conservation status for Grylloblattodea species. Mol. Phylogen. Evol. 41, 222-237.

Jombart, T., Kendall, M., Almagro-Garcia, J., Colijn, C., 2017. treespace: Statistical exploration of landscapes of phylogenetic trees. Mol. Ecol. Res. 17, 1385-1392.

Kamp, J.W., 1970. The cavernicolous Grylloblattodea of the western United States. Annales de Speleologie 25, 223-230.

Kamp, J.W., 1973. Biosystematics of the Grylloblattodea. Department of Zoology. University of British Columbia, Vancouver, p. 275.

Kamp, J.W., 1979. Taxonomy, distribution, and zoogeographic evolution of *Grylloblatta* in Canada (Insecta: Notoptera). The Canadian Entomologist 111, 27-38.

Kaufman, D.S., Porter, S.C., Gillespie, A.R., 2004. Quaternary alpine glaciation in Alaska, the Pacific Northwest, Sierra Nevada, and Hawaii. In: Gillespie, A.R., Porter, S.C., Atwater, B.F. (Eds.), The Quaternary Period in the United States, 1. Elsevier, Amsterdam, The Netherlands, pp. 77-103.

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

Kendall, M., Colijn, C., 2016. Mapping phylogenetic trees to reveal distinct patterns of evolution. Mol. Biol. Evol. 33, 2735-2743.

Kendall, M., Eldholm, V., Colijn, C., 2018. Comparing phylogenetic trees according to tip label categories. bioRxiv, 251710.

Knowles, L.L., 2000. Tests of Pleistocene speciation in montane grasshoppers (genus *Melanoplus*) from the sky islands of western North America. Evolution 54, 1337-1348.

Knowles, L.L., Carstens, B.C., 2007. Estimating a geographically explicit model of population divergence Evolution 61, 477-493.

Lemey, P., Rambaut, A., Welch, J., Suchard, M., 2010. Phylogeography takes a relaxed random walk in continuous space and time. Mol. Biol. Evol. 27, 1877-1885.

Lisiecki, L.E., Raymo, M.E., 2007. Plio-Pleistocene climate evolution: trends and transitions in glacial cycle dynamics. Quat. Sci. Rev. 26, 58-69.

Locke, W.W., 1990. Late Pleistocene glaciers and the climate of western Montana, U.S.A. Arct Alp Res 22, 1-13.

Loehr, J., Worley, K., Grapputo, A., Carey, J., Veitch, A., Coltman, D., 2006. Evidence for cryptic glacial refugia from North American mountain sheep mitochondrial DNA. J. Evol. Biol. 19, 419-430.

Marshall, C.J., Lytle, D.A., 2015. Two new species of *Grylloblatta* Walker, 1914 (Grylloblattodea: Grylloblattidae) from western North America, and a neotype designation for *G. rothi* Gurney 1953. Zootaxa 3949, 408-418.

Martínez-Solano, I., Jockusch, E.L., Wake, D.B., 2007. Extreme population subdivision throughout a continuous range: phylogeography of *Batrachoseps attenuatus* (Caudata: Plethodontidae) in western North America. Mol. Ecol. 16, 4335-4355.

Mirarab, S., Warnow, T., 2015. ASTRAL-II: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. Bioinformatics 31, i44-i52.

Morlon, H., Lewitus, E., Condamine, F.L., Manceau, M., Clavel, J., Drury, J., 2016. RPANDA: an R package for macroevolutionary analyses on phylogenetic trees. Methods Ecol Evol 7, 589-597.

Moulton, M.J., Song, H., Whiting, M.F., 2010. Assessing the effects of primer specificity on eliminating numt coamplification in DNA barcoding: a case study from Orthoptera (Arthropoda: Insecta). Mol. Ecol. Res. 10, 615-627.

Nei, M., Maruyama, T., Chakraborty, R., 1975. Bottleneck effect and genetic variability in populations. Evolution 29, 1-10.

Pielou, E.C., 2008. After the ice age: the return of life to glaciated North America. University of Chicago Press.

Pierce, K.L., 2003. Pleistocene glaciations of the Rocky Mountains. Developments in Quaternary Sciences. Elsevier, pp. 63-76.

Rambaut, A., Drummond, A.J., 2009a. TRACER: MCMC Trace Analysis Tool Version v1.5.0. University of Oxford, Oxford.

Rambaut, A., Drummond, A.J., 2009b. TRACER: MCMC Trace Analysis Tool Version v1.6. University of Oxford, Oxford, http://tree.bio.ed.ac.uk/software/tracer/.

Roberts, D.R., Hamann, A., 2015. Glacial refugia and modern genetic diversity of 22 western North American tree species. Proceedings of the Royal Society B: Biological Sciences 282.

Rovito, S., Schoville, S., 2017. Testing models of refugial isolation, colonization and population connectivity in two species of montane salamanders. Heredity 119, 265-274.

Rovito, S.M., 2010. Lineage divergence and speciation in the Web-toed Salamanders (Plethodontidae: *Hydromantes*) of the Sierra Nevada, California. Mol. Ecol. 19, 4554-4571.

Sanderson, M.J., 2003. r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. Bioinformatics 19, 301-302.

Schmitt, T., 2007. Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. Frontiers in Zoology 4, 1-13.

Schmitt, T., Rober, S., Seitz, A., 2005. Is the last glaciation the only relevant event for the present genetic population structure of the meadow brown butterfly Maniola jurtina (Lepidoptera : Nymphalidae)? Biol. J. Linn. Soc. 85, 419-431.

Schneeweiss, G.M., Winkler, M., Schönswetter, P., 2017. Secondary contact after divergence in allopatry explains current lack of ecogeographical isolation in two hybridizing alpine plant species. J. Biogeogr. 44, 2575-2584.

Schönswetter, P., Elven, R., Brochmann, C., 2008. Trans-Atlantic dispersal and large-scale lack of genetic structure in the circumpolar, arctic-alpine sedge *Carex bigelowii* sl (Cyperaceae). Am. J. Bot. 95, 1006-1014.

Schoville, S.D., 2012. Three new species of *Grylloblatta* Walker (Insecta: Grylloblattodea: Grylloblattidae), from southern Oregon and northern California. Zootaxa 3412, 42-52.

Schoville, S.D., 2014. Current status of the systematics and evolutionary biology of Grylloblattidae (Grylloblattodea). Syst. Entomol. 39, 197–204.

Schoville, S.D., Graening, G.O., 2013. Updated checklist of the ice-crawlers (Insecta: Grylloblattodea: Grylloblattidae) of North America, with notes on their natural history, biogeography and conservation. Zootaxa 3737, 351-378.

Schoville, S.D., Roderick, G.K., 2009. Alpine biogeography of Parnassian butterflies during Quaternary climate cycles in North America. Mol. Ecol. 18, 3471–3485.

Schoville, S.D., Roderick, G.K., 2010. Evolutionary diversification of cryophilic *Grylloblatta* species (Grylloblattodea: Grylloblattidae) in alpine habitats of California. BMC Evol. Biol. 10, 163.

Schoville, S.D., Roderick, G.K., Kavanaugh, D.H., 2012a. Testing the 'Pleistocene species pump' in alpine habitats: lineage diversification of flightless ground beetles (Coleoptera: Carabidae: Nebria) in relation to altitudinal zonation. Biol. J. Linn. Soc. 107, 95-111.

Schoville, S.D., Roderick, G.K., Kavanaugh, D.H., 2012b. Testing the 'Pleistocene species pump'in alpine habitats: lineage diversification of flightless ground beetles (Coleoptera: Carabidae: *Nebria*) in relation to altitudinal zonation. Biol. J. Linn. Soc. 107, 95-111.

Schoville, S.D., Slatyer, R.A., Bergdahl, J.C., Valdez, G.A., 2015. Conserved and narrow temperature limits in alpine insects: Thermal tolerance and supercooling points of the ice-

crawlers, *Grylloblatta* (Insecta: Grylloblattodea: Grylloblattidae). J. Insect Physiol. 78, 55-61.

Schoville, S.D., Uchifune, T., Machida, R., 2013. Colliding fragment islands transport independent lineages of endemic rock-crawlers (Grylloblattodea: Grylloblattidae) in the Japanese archipelago. Mol. Phylogen. Evol. 66, 915-927.

Shafer, A.B.A., Cullingham, C.I., Cote, S.D., Coltman, D.W., 2010. Of glaciers and refugia: a decade of study sheds new light on the phylogeography of northwestern North America. Mol. Ecol. 19, 4589-4621.

Sim, Z., Hall, J.C., Jex, B., Hegel, T.M., Coltman, D.W., 2016. Genome-wide set of SNPs reveals evidence for two glacial refugia and admixture from postglacial recolonization in an alpine ungulate. Mol. Ecol.

Simon, C., Buckley, T.R., Frati, F., Stewart, J.B., Beckenbach, A.T., 2006. Incorporating molecular evolution into phylogenetic analysis, and a new compilation of conserved polymerase chain reaction primers for animal mitochondrial DNA. Annu Rev Ecol Evol Syst 37, 545-579.

Soltis, D.E., Gitzendanner, M.A., Strenge, D.D., Soltis, P.S., 1997. Chloroplast DNA intraspecific phylogeography of plants from the Pacific Northwest of North America. Plant Syst. Evol. 206, 353-373.

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312-1313.

Steinbauer, M.J., Field, R., Grytnes, J.-A., Trigas, P., Ah-Peng, C., Attorre, F., Birks, H.J.B., Borges, P.A.V., Cardoso, P., Chou, C.-H., De Sanctis, M., de Sequeira, M.M., Duarte, M.C., Elias, R.B., Fernández-Palacios, J.M., Gabriel, R., Gereau, R.E., Gillespie,

R.G., Greimler, J., Harter, D.E.V., Huang, T.-J., Irl, S.D.H., Jeanmonod, D., Jentsch, A., Jump, A.S., Kueffer, C., Nogué, S., Otto, R., Price, J., Romeiras, M.M., Strasberg, D., Stuessy, T., Svenning, J.-C., Vetaas, O.R., Beierkuhnlein, C., 2016. Topography-driven isolation, speciation and a global increase of endemism with elevation. Global Ecol. Biogeogr. 25, 1097-1107.

Stewart, J.R., Lister, A.M., Barnes, I., Dalen, L., 2010a. Refugia revisited: individualistic responses of species in space and time. Proceedings of the Royal Society B-Biological Sciences 277, 661-671.

Stewart, J.R., Lister, A.M., Barnes, I., Dalén, L., 2010b. Refugia revisited: individualistic responses of species in space and time. Proceedings of the Royal Society B: Biological Sciences 277, 661-671.

Svenson, G.J., Whiting, M.F., 2004. Phylogeny of Mantodea based on molecular data: evolution of a charismatic predator. Syst. Entomol. 29, 359-370.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725-2729.

Terborgh, J., 1992. Diversity and the Tropical Rain Forest. Freeman, New York.

Varga, Z.S., Schmitt, T., 2008. Types of oreal and oreotundral disjunctions in the western Palearctic. Biol. J. Linn. Soc. 93, 415-430.

Walker, E.M., 1914. A new species of Orthoptera forming a new genus and family. The Canadian Entomologist 46, 93-99.

Wiens, J.J., 2006. Missing data and the design of phylogenetic analyses. J. Biomed. Inf. 39, 34-42.

Figure Legends

Fig. 1. Distribution of core refugia for cold-specialized taxa during the Last Glacial Maximum (26,500 years before present). Location of refugia (shown as hatched polygons) modified from Shafer et al. (2010) and Roberts and Hamann (2015). The Haida-Gwai (HG) and Vancouver Island (VI) refugia are too small to see at this scale. The extent of glacial ice is shown as solid gray polygons.

Fig. 2. Photos of the family Grylloblattidae. A) *Grylloblatta* sp., Table Mountain, Washington, U.S.A. B) *Grylloblattella* sp., Krasnaya Mountain, Altai Republic, Russia. C) *Galloisiana* sp., Deogyusan National Park, Jeollabuk-Do Province, South Korea. D) *Galloisiana yuasai*, Norikura-dake, Nagano Prefecture, Japan.

Fig. 3. A) Geographical distribution of genetic samples of *Grylloblatta.* Sample sites indicated by numbered black dots. The extent of glacial ice shown in solid white polygons, overlain on a contemporary environmental basemap (ESRI). Colored lines correspond to major genetic clades. B) Bayesian phylogenetic relationships of *Grylloblatta* inferred from the mitochondrial COII locus and C) 322 concatenated nuclear genes. Nodal support values for the phylogeny are shown as Bayesian posterior probability estimates and branch lengths are in units of average substitutions per site.

Fig. 4. Bayesian phylogenetic relationships of *Grylloblatta* inferred from the 322 concatenated nuclear genes, with terminal branches labeled by site and species. Nodal

support values for the phylogeny are shown as Bayesian posterior probability estimates and branch lengths are in units of average substitutions per site.

Fig. 5. A) Species tree relationships of *Grylloblatta* lineages inferred from 322 nuclear genes. Asterisks indicate local posterior probability >0.95 and branch lengths are represented in coalescent units. Major clades (species complexes) are uniquely colored. B) The distribution of major clades suggest high geographical cohesion and local endemism.

Fig. 6. For well-supported *Grylloblatta* lineages, COII data were used to estimate ancestral ranges and demographic trends. A) Ancestral range estimates are shown as colored polygons with distinct haplotypes as lines coalescing into the ancestral population. B) Bayesian skyline plots of effective population size change through time. Color codes for lineages are as follows: *G. campodeiformis* (yellow), *G.* sp. 'North Cascades' (orange), *G*. sp. 'Trout Lake' (dark green), *G. rothi* (light green), *G. marmoreus* (salmon), *G. chandleri* (red), *G. washoa* (turquoise), *G*. sp. 'Central Sierra' (violet), *G. newberryensis* (teal), and *G. gurneyi* (light red). Several other clades not shown (see Supplementary Figures S8-S20).

Fig. 7. Divergence time and diversification rate analysis. A) The ultrametric species tree showing estimates of the divergence time. The statistical association between B) global temperature change and C) diversification rate is shown relative to time in millions of years.

Table 1. COII population genetic summary statistics in 14 populations of *Grylloblatta.* Values of the population genetic parameter theta based on segregating sites (θ_S) and average nucleotide diversity ($\theta\pi$) are divided by the total number of COII base pairs (bp). Significant values for Tajima's *D* and Fu's *Fs* statistics are indicated with an asterisk using $\alpha = 0.05$. Tajima's *D* and Fu's *Fs* could not be calculated for Chinook Pass or the White Mountains because only one haplotype was present.

	θ_S (per bp)	$\theta\pi$ (per bp)	Tajima's *D*	Fu's *Fs*
Mt. Adams (n = 13)	0.007	0.003	-2.18*	-0.76
Chinook Pass (n = 10)	0	0	N/A	N/A
Mt. Jefferson (n = 25)	0.004	0.001	-2.03*	-4.31*
Lassen Peak (n = 10)	0.009	0.008	-0.79	-3.23*
Marble Valley (n = 10)	0.002	0.001	-1.56*	1.22
South Sister (n = 28)	0.001	0	-1.32	-0.67
Mt. Spokane (n = 19)	0.003	0.002	-0.87	-2.27
Mt. Thielsen (n = 24)	0.001	0	-1.73*	-3.02*
Thompson Peak sp. nov. 1 (n =16)	0.003	0.002	-1.19	-4.27*
Barker Pass (n = 17)	0.004	0.006	1.73	4.31
Graveyard Lakes (n = 12)	0.001	0.001	-0.38	-0.36
Sonora Pass (n = 12)	0.001	0.001	.02	-0.62
Susie Lake (n = 22)	0.007	0.006	-0.62	-2.21
White Mountains (n = 11)	0	0	N/A	N/A

Table 2. Comparison of the diversification models using corrected Akaike Information Criteria (AICc). The temperature dependent model is the best scoring model (lowest AICc, in bold). A positive rate change indicates a decreasing diversification rate through time to the present.

Model	AICc score	Speciation rate (λ)	Extinction rate (μ)	Speciation rate change	Extinction rate change
Birth & Death Constant	320.0730	0.1071792	1.367618e-08	0	0
Birth Variable & Death Constant	313.1693	0.06157839	-2.340953e-05	0.07242739	0
Birth Constant & Death Variable	310.1239	0.1838757	813.6242	0	-745.0941
Birth & Death Variable	315.5350	0.06162372	3.550783e-08	0.07226793	-0.183418
Temperature Dependent, No Extinction	**305.9838**	0.02649174	0	0.26184010	0
Temperature Dependent, With Extinction	310.654	0.0249050	-5.032847e-08	0.2741258	4.771803e-08

Supplemental File 1

Table S1. Collection sites of *Grylloblatta* that were sampled for genetic data in this study. Location numbers correspond to numbers shown in the phylogenetic trees. Accession numbers correspond to COII haplotypes found at each site. An asterisk (*) indicates a holotype or neotype locality, while bold type indicates sympatric lineages at a site.

Table S2. List of the 322 single copy orthologs used for nuclear gene phylogeny, with length in base pairs and functional annotation.

Table S3. Collection sites in the nuclear gene tree with latitude and longitude coordinates and corresponding location numbers used on the map. Genbank accessions are available in Bioproject PRJNA488740. An asterisk (*) indicates a holotype or neotype locality, while bold type indicates sympatric lineages at a site.

Table S4. Results for phylogenetic model selection in BEAST of the COII dataset using AICM.

Table S5. Pairwise differences among trees using the Kendall Colijn metric. Below the diagonal, the pairwise distance is based on topological differences only. Above the diagonal, the pairwise distance among trees is based on branch lengths (species tree not included as branch lengths were measured in coalescent units).

Table S6. Estimated root age for well-supported *Grylloblatta* lineages examined in the ancestral range reconstruction and demographic analysis.

Fig. S1. Bayesian phylogenetic relationships of *Grylloblatta* inferred from the COII mitochondrial gene. Numbers and names at the tips refer to sample sites and known taxon names, respectively. Well-supported genetic lineages are grouped into major clades (MC), represented as colored groups and referred to in the text. Nodal support values for the phylogeny are shown as Bayesian posterior probability estimates and branch lengths are in units of substitutions per site.

Fig. S2. Bayesian phylogenetic relationships of *Grylloblatta* inferred from 322 concatenated nuclear genes (full dataset, with data exceeding 30% missing genes per individual in some cases). Numbers and names at the tips refer to sample sites and known taxon names, respectively. Well-supported genetic lineages are grouped into major clades (MC), represented as colored groups and referred to in the text. Nodal support values for the phylogeny are shown as Bayesian posterior probability estimates and branch lengths are in units of average substitutions per site.

Fig. S3. Bayesian phylogenetic relationships of *Grylloblatta* inferred from 266 concatenated nuclear genes (<30% missing genes per individual allowed). Numbers and names at the tips refer to sample sites and known taxon names, respectively. Well-supported genetic lineages are grouped into major clades (MC), represented as colored groups and referred to in the text. Nodal support values for the phylogeny are shown as Bayesian posterior probability estimates and branch lengths are in units of average substitutions per site.

Fig. S4. Bayesian phylogenetic relationships of *Grylloblatta* inferred from 191 concatenated nuclear genes (<20% missing genes per individual allowed). Numbers and names at the tips refer to sample sites and known taxon names, respectively. Well-supported genetic lineages are grouped into major clades (MC), represented as colored groups and referred to in the text. Nodal support values for the phylogeny are shown as Bayesian posterior probability estimates and branch lengths are in units of average substitutions per site.

Fig. S5. Bayesian phylogenetic relationships of *Grylloblatta* inferred from 77 concatenated nuclear genes (<10% missing genes per individual allowed). Numbers and names at the tips refer to sample sites and known taxon names, respectively. Well-supported genetic lineages are grouped into major clades (MC), represented as colored groups and referred to in the text. Nodal support values for the phylogeny are shown as Bayesian posterior probability estimates and branch lengths are in units of average substitutions per site.

Fig. S6. Bayesian phylogenetic relationships of *Grylloblatta* inferred from 26 concatenated nuclear genes (<less than 5% missing genes per individual allowed). Numbers and names at the tips refer to sample sites and known taxon names, respectively. Well-supported genetic lineages are grouped into major clades (MC), represented as colored groups and referred to in the text. Nodal support values for the phylogeny are shown as Bayesian posterior probability estimates and branch lengths are in units of average substitutions per site.

Fig. S7. Multi-species coalescent tree inferred from ASTRAL. Numbers and names at the tips refer to sample sites and known taxon names, respectively. Well-supported genetic lineages are grouped into major clades, represented as colored groups and referred to in the text. Numbers at nodes indicate local posterior probability and branch lengths are represented in coalescent units.

Fig. S8. The estimated A) ancestral ranges and B) Bayesian skyline plots of effective population size change through time for the *Grylloblatta campodeiformis* clade, based on mitochondrial COII sequence data. Ancestral range estimates are shown as a white colored polygon with distinct haplotypes as lines coalescing into the ancestral population.

Fig. S9. The estimated A) ancestral ranges and B) Bayesian skyline plots of effective population size change through time for the *Grylloblatta chandleri* clade, based on mitochondrial COII sequence data. Ancestral range estimates are shown as a white colored polygon with distinct haplotypes as lines coalescing into the ancestral population.

Fig. S10. The estimated A) ancestral ranges and B) Bayesian skyline plots of effective population size change through time for the *Grylloblatta chirurgica* clade, based on mitochondrial COII sequence data. Ancestral range estimates are shown as a white colored polygon with distinct haplotypes as lines coalescing into the ancestral population.

Fig. S11. The estimated A) ancestral ranges and B) Bayesian skyline plots of effective population size change through time for the *Grylloblatta gurneyi* clade, based on mitochondrial COII sequence data. Ancestral range estimates are shown as a white colored polygon with distinct haplotypes as lines coalescing into the ancestral population.

Fig. S12. The estimated A) ancestral ranges and B) Bayesian skyline plots of effective population size change through time for the *Grylloblatta* sp. 'Southwest Sierra Nevada' clade, based on mitochondrial COII sequence data. Ancestral range estimates are shown

as a white colored polygon with distinct haplotypes as lines coalescing into the ancestral population.

Fig. S13. The estimated A) ancestral ranges and B) Bayesian skyline plots of effective population size change through time for the *Grylloblatta marmoreus* clade, based on mitochondrial COII sequence data. Ancestral range estimates are shown as a white colored polygon with distinct haplotypes as lines coalescing into the ancestral population.

Fig. S14. The estimated A) ancestral ranges and B) Bayesian skyline plots of effective population size change through time for the *Grylloblatta newberryensis* clade, based on mitochondrial COII sequence data. Ancestral range estimates are shown as a white colored polygon with distinct haplotypes as lines coalescing into the ancestral population.

Fig. S15. The estimated A) ancestral ranges and B) Bayesian skyline plots of effective population size change through time for the *Grylloblatta rothi* clade, based on mitochondrial COII sequence data. Ancestral range estimates are shown as a white colored polygon with distinct haplotypes as lines coalescing into the ancestral population.

Fig. S16. The estimated A) ancestral ranges and B) Bayesian skyline plots of effective population size change through time for the *Grylloblatta* sp. 'North Cascades' clade, based on mitochondrial COII sequence data. Ancestral range estimates are shown as a white colored polygon with distinct haplotypes as lines coalescing into the ancestral population.

Fig. S17. The estimated A) ancestral ranges and B) Bayesian skyline plots of effective population size change through time for the *Grylloblatta* sp. *'Central Sierra Nevada'* clade, based on mitochondrial COII sequence data. Ancestral range estimates are shown as a white colored polygon with distinct haplotypes as lines coalescing into the ancestral population.

Fig. S18. The estimated A) ancestral ranges and B) Bayesian skyline plots of effective population size change through time for the *Grylloblatta* sp. 'Mt. Spokane' clade, based on mitochondrial COII sequence data. Ancestral range estimates are shown as a white colored polygon with distinct haplotypes as lines coalescing into the ancestral population.

Fig. S19. The estimated A) ancestral ranges and B) Bayesian skyline plots of effective population size change through time for the *Grylloblatta* sp. 'Trout Lake' clade, based on mitochondrial COII sequence data. Ancestral range estimates are shown as a white colored polygon with distinct haplotypes as lines coalescing into the ancestral population.

Fig. S20. The estimated A) ancestral ranges and B) Bayesian skyline plots of effective population size change through time for the *Grylloblatta washoa* clade, based on mitochondrial COII sequence data. Ancestral range estimates are shown as a white colored polygon with distinct haplotypes as lines coalescing into the ancestral population.

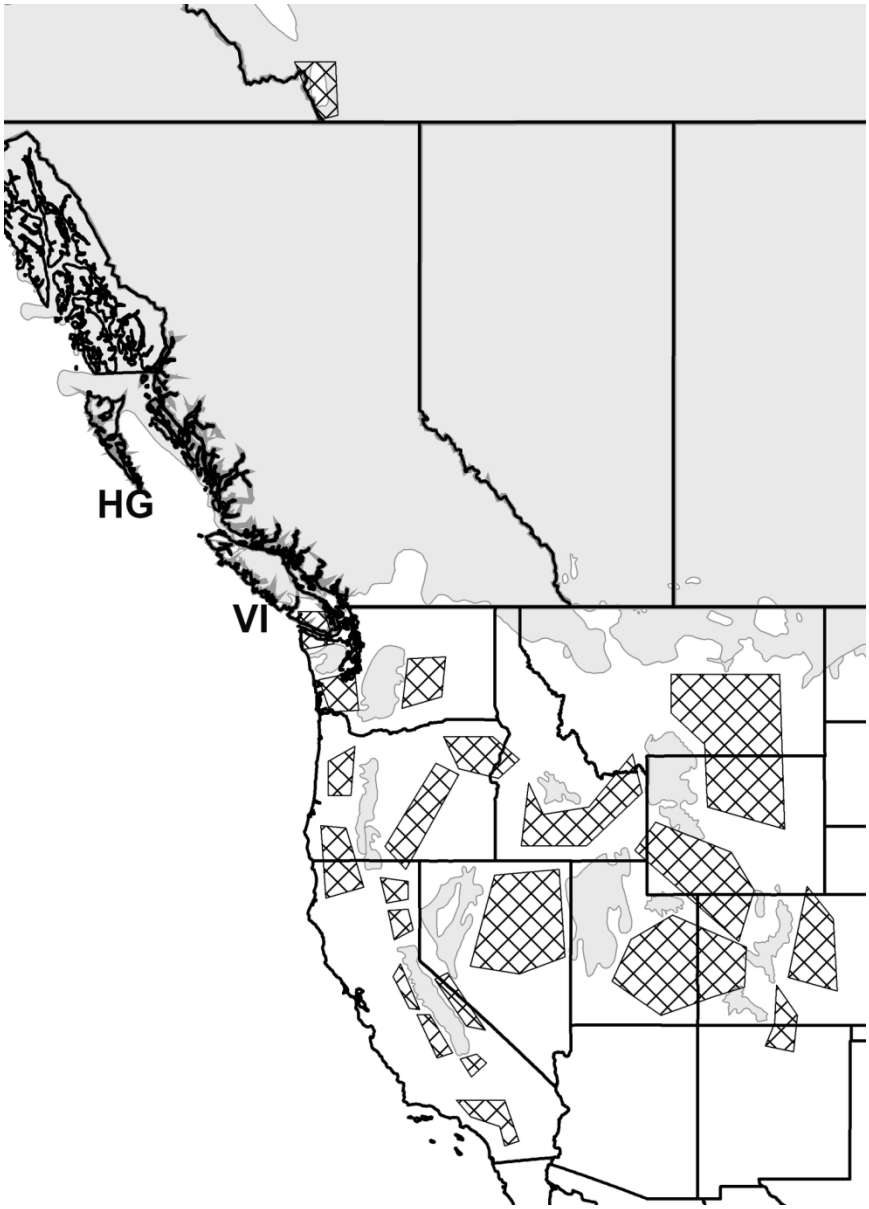

Fig. 1. Distribution of core refugia for cold-specialized taxa during the Last Glacial Maximum (26,500 years before present). Location of refugia (shown as hatched polygons) modified from Shafer et al. (2010) and Roberts and Hamann (2015). The Haida-Gwai (HG) and Vancouver Island (VI) refugia are too small to see at this scale. The extent of glacial ice is shown as solid gray polygons.

131x182mm (300 x 300 DPI)

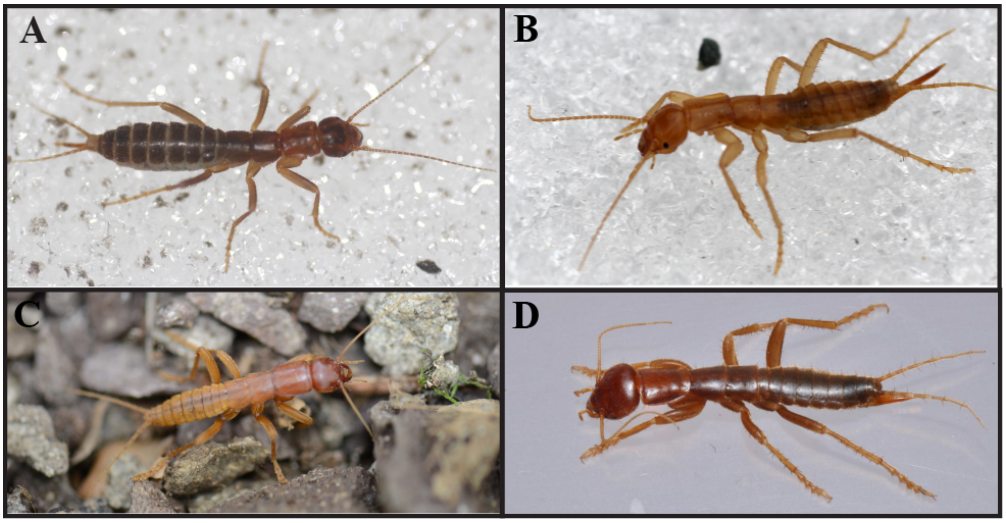

Fig. 2. Photos of the family Grylloblattidae. A) *Grylloblatta* sp., Table Mountain, Washington, U.S.A. B) *Grylloblattella* sp., Krasnaya Mountain, Altai Republic, Russia. C) *Galloisiana* sp., Deogyusan National Park, Jeollabuk-Do Province, South Korea. D) *Galloisiana yuasai*, Norikura-dake, Nagano Prefecture, Japan.

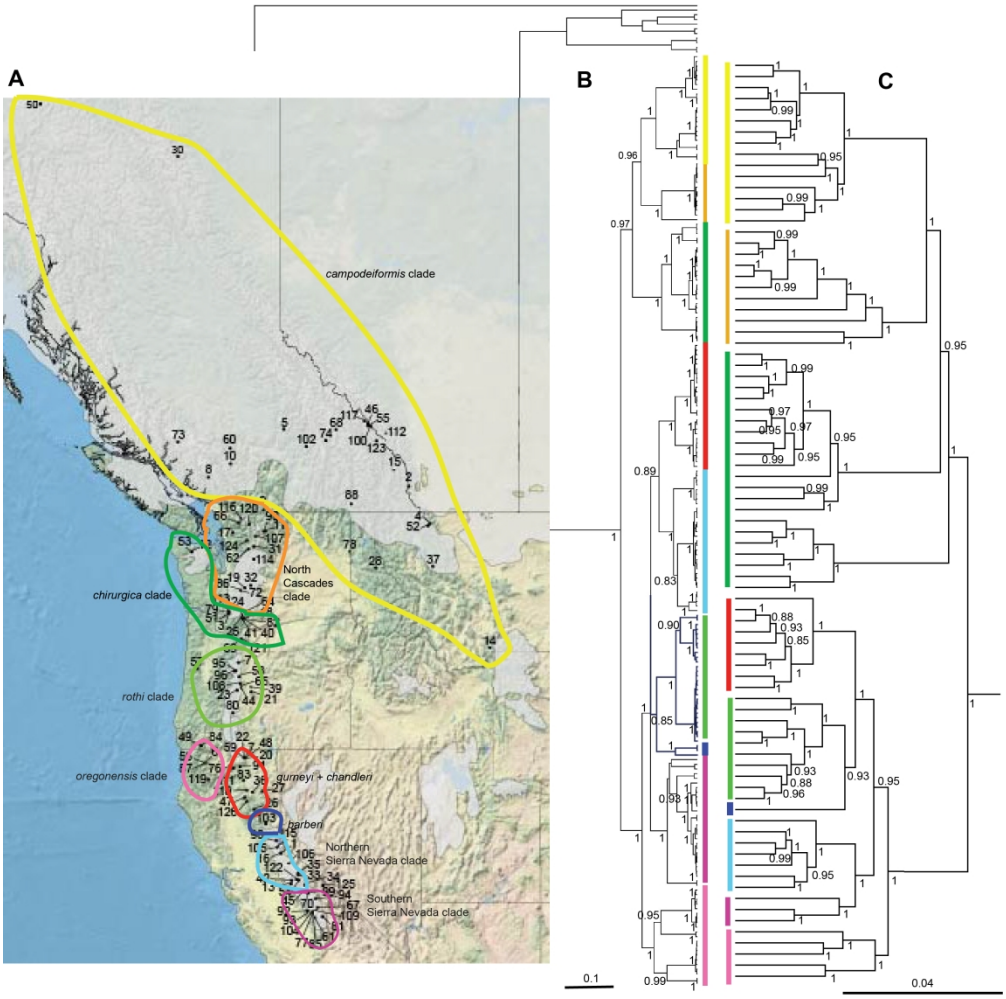

Figure 2. A) Geographical distribution of genetic samples of *Grylloblatta*. Sample sites indicated by numbered black dots. The extent of glacial ice shown in solid white polygons, overlain on a contemporary environmental basemap (ESRI). Colored lines correspond to major genetic clades. B) Bayesian phylogenetic relationships of *Grylloblatta* inferred from the mitochondrial COII locus and C) 322 concatenated nuclear genes. Nodal support values for the phylogeny are shown as Bayesian posterior probability estimates and branch lengths are in units of average substitutions per site.

281x289mm (300 x 300 DPI)

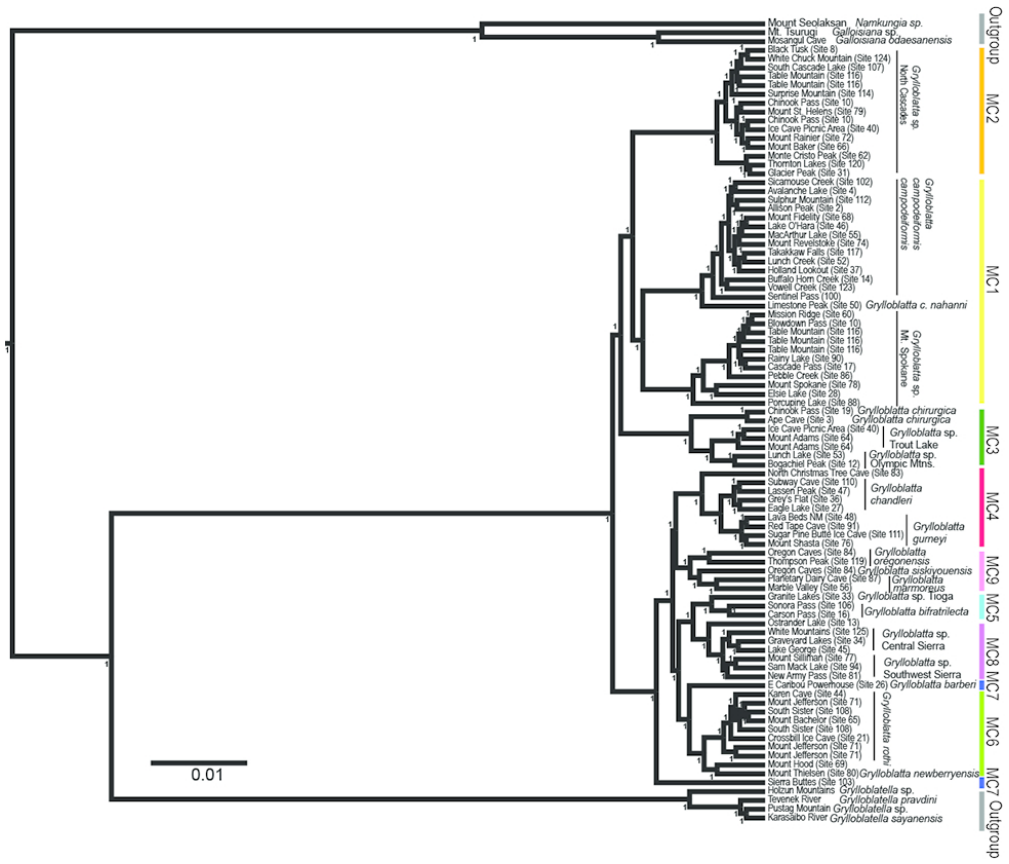

Fig. 4. Bayesian phylogenetic relationships of *Grylloblatta* inferred from the 322 concatenated nuclear genes, with terminal branches labeled by site and species. Nodal support values for the phylogeny are shown as Bayesian posterior probability estimates and branch lengths are in units of average substitutions per site.

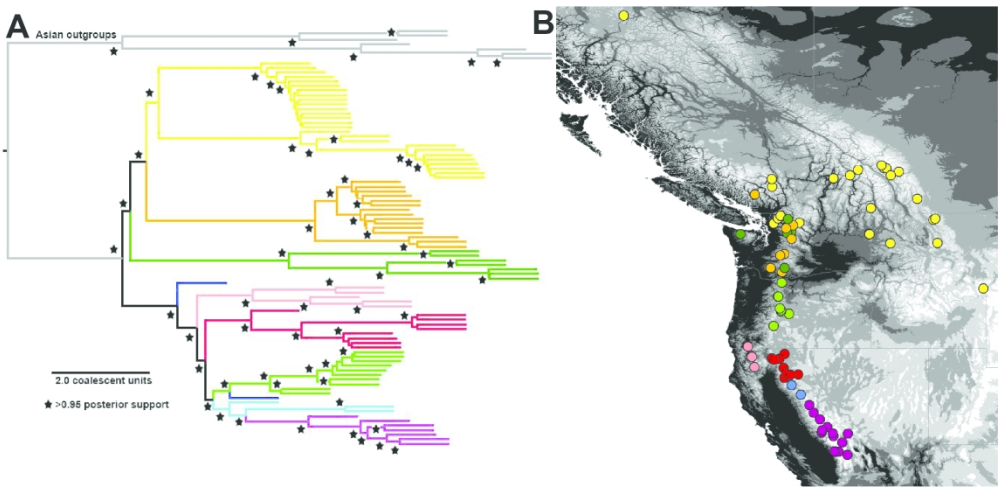

Fig. 5. A) Species tree relationships of *Grylloblatta* lineages inferred from 322 nuclear genes. Asterisks indicate local posterior probability >0.95 and branch lengths are represented in coalescent units. Major clades (species complexes) are uniquely colored. B) The distribution of major clades suggest high geographical cohesion and local endemism.

193x93mm (300 x 300 DPI)

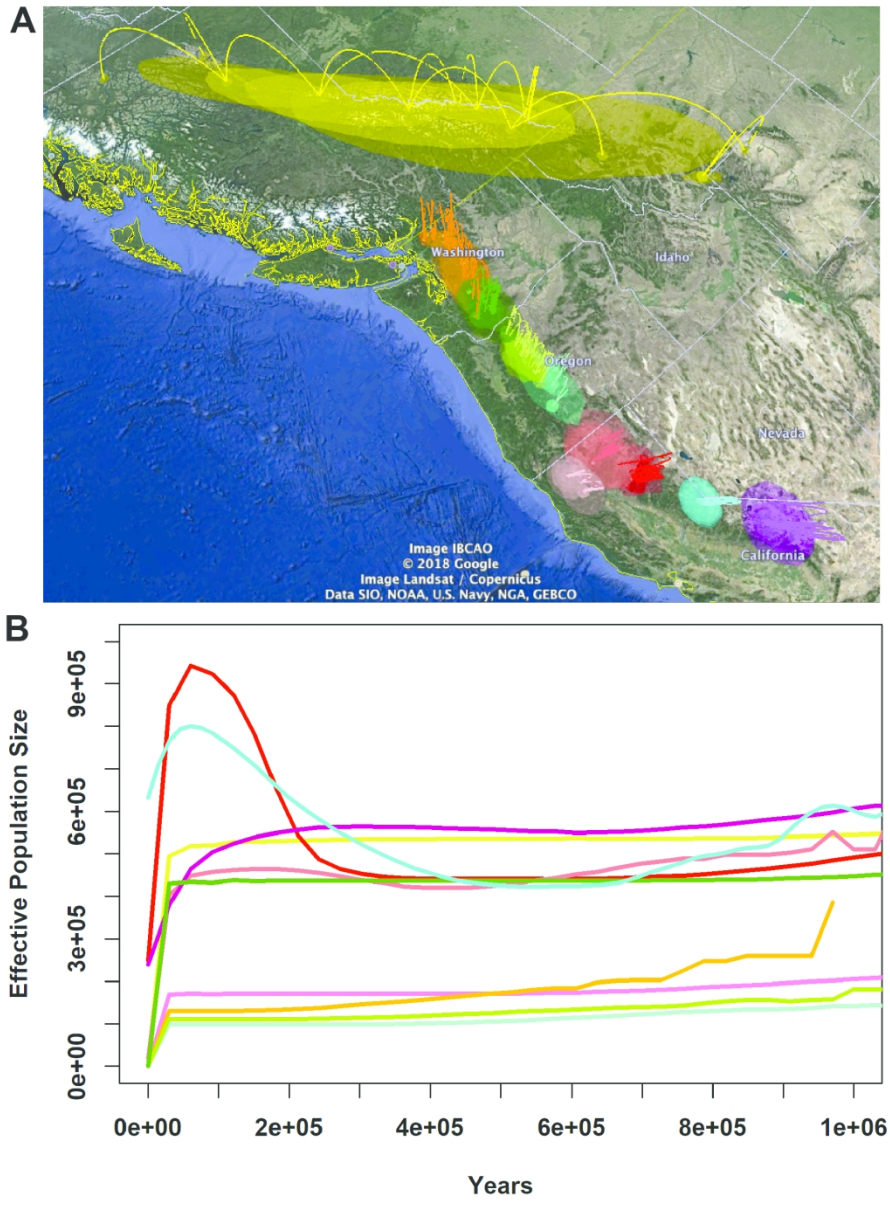

Fig. 6. For well-supported *Grylloblatta* lineages, COII data were used to estimate ancestral ranges and demographic trends. A) Ancestral range estimates are shown as colored polygons with distinct haplotypes as lines coalescing into the ancestral population. B) Bayesian skyline plots of effective population size change through time. Color codes for lineages are as follows: *G. campodeiformis* (yellow), *G.* sp. 'North Cascades' (orange), *G.* sp. 'Trout Lake' (dark green), *G. rothi* (light green), *G. marmoreus* (salmon), *G. chandleri* (red), *G. washoa* (turquoise), *G.* sp. 'Central Sierra' (violet), *G. newberryensis* (teal), and *G. gurneyi* (light red). Several other clades not shown (see Supplementary Figures S8-S20).

158x212mm (300 x 300 DPI)

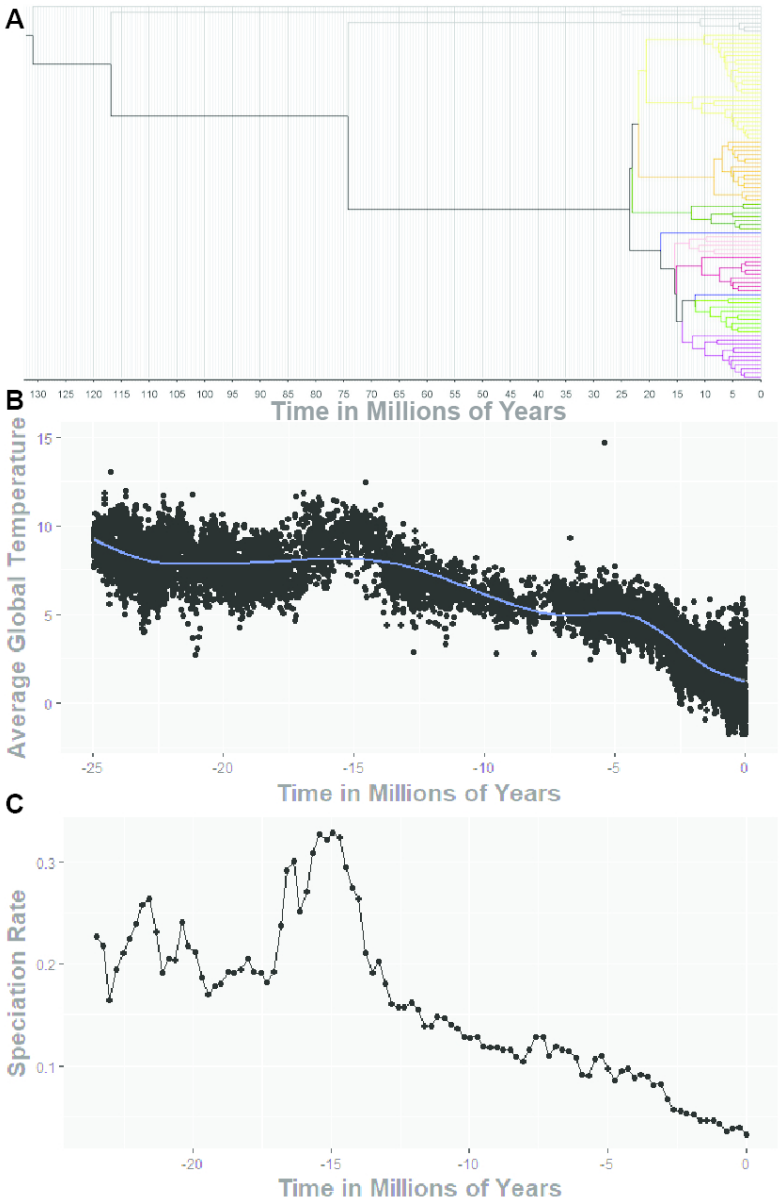

Fig. 7. Divergence time and diversification rate analysis. A) The ultrametric species tree showing estimates of the divergence time. The statistical association between B) global temperature change and C) diversification rate is shown relative to time in millions of years.

142x215mm (150 x 150 DPI)

Table S1. Collection sites of *Grylloblatta* that were sampled for genetic data in this study. Location numbers correspond to numbers shown in the phylogenetic trees. Accession numbers correspond to COII haplotypes found at each site. An asterisk (*) indicates a holotype or neotype locality, while bold type indicates sympatric lineages at a site.

Taxon	Major Clade Number	Site	Latitude	Longitude	Location Number	Genbank Accession #
Grylloblatta bifratrilecta	Northern Sierra Nevada	Above Winnemucca Lake	38.66666	-119.99001	1	MG824854
Grylloblatta campodeiformis	campodeiformis	Allison Peak	49.73802	-114.6302	2	MG824747, MG824748, MG824750, MG824817
Grylloblatta chirurgica, Grylloblatta sp. Trout Lake	chirurgica	**Ape Cave**	46.10794	-122.21149	3	MG824788, MG824828, MG824829, MG824855
Grylloblatta sp. Mt. Spokane	campodeiformis	Barrier Lake	51.25464	-119.86617	5	MG824845
Grylloblatta marmoreus	oregonensis	Big Foot Cave, Marble Valley*	41.5628	-123.1899	6	JN612963
Grylloblatta gurneyi	gurneyi + chandleri	Big Ice Cave	41.71831	-121.50561	7	MG824766
Grylloblatta sp. Mt. Spokane	campodeiformis	Black Tusk	49.9751	-123.0431	8	MG824845
Grylloblatta sp. Mt. Spokane	campodeiformis	Blackwall Peak on Paintbrush Trail	49.10033	-120.75788	9	MG824834
Grylloblatta sp. Mt. Spokane	campodeiformis	Blowdown Pass, Stein Valley	50.3626	-122.15	10	MG824834
Grylloblatta sp. North Cascades, Grylloblatta sp. Mt. Spokane	North Cascades, campodeiformis	**Blue Lake**	48.50509	-120.67184	11	MG824752, MG824834
Grylloblatta sp. Olympc Mtns.	chirurgica	Bogachiel Peak	47.90546	-123.78304	12	MG824820
Grylloblatta sp. Ostrander Lake	Southern Sierra Nevada	Bridalveil trail to Ostrander Lake	37.63543	-119.58054	13	FJ918583
Grylloblatta campodeiformis	campodeiformis	Buffalo Horn Creek	45.10403	-111.20861	14	DQ457367, MG824786, MG824816
Grylloblatta sp. Mt. Spokane	campodeiformis	Bull River Valley	50.17958	-115.22627	15	MG824845
Grylloblatta bifratrilecta	Northern Sierra Nevada	Carson Pass	38.6671	-119.98962	16	FJ918581, FJ918617, FJ918618
Grylloblatta sp. Mt. Spokane	campodeiformis	Cascade Pass	48.47445	-121.50604	17	MG824834
Grylloblatta sp. Trout Lake	chirurgica	Cheese Cave	45.98538	-121.54961	18	DQ457359
Grylloblatta chirurgica, Grylloblatta sp. North Cascades	chirurgica, North Cascades	**Chinook Pass**	46.87271	-121.51567	19	DQ457345, DQ457346, MG824829, MG824833, MG824858
Grylloblatta gurneyi	gurneyi + chandleri	Cox Cave	41.715	-121.498	20	DQ457349
Grylloblatta newberryensis	rothi	Crossbill Ice Cave	43.779	-121.261	21	MG824825
Grylloblatta gurneyi	gurneyi + chandleri	Crystal Ice Cave	41.71831	-121.50561	22	MG824765
Grylloblatta rothi	rothi	Cultus Mountain*	43.819	-121.8703	23	KF880959
Grylloblatta sp. Trout Lake	chirurgica	Dead Horse Cave	46.0672	-121.637	24	DQ457363
Grylloblatta sp. Trout Lake	chirurgica	Dry Creek Cave	45.99493	-121.67464	25	DQ457361
Grylloblatta barberi	barberi	E Caribou Powerhouse*	40.09134	-121.13687	26	FJ918624, MG824763

Grylloblatta chandleri	gurneyi + chandleri	Eagle Lake*	40.64593	-120.79482	27	MG824781, MG824831
Grylloblatta sp. Mt. Spokane	campodeiformis	Elsie Lake	47.42846	-116.02541	28	MG824848, MG824849
Grylloblatta campodeiformis athapaska	campodeiformis	Gathto Creek	57.92708	-124.34231	30	MG824749
Grylloblatta sp. North Cascades	North Cascades, campodeiformis	Glacier Peak	48.04945	-121.15511	31	MG824822
Grylloblatta chirurgica	chirurgica	Goat Peak Trail	46.93901	-121.2652	32	DQ457362
Grylloblatta sp. Tioga Crest	Northern Sierra Nevada	Granite Lakes	37.92809	-119.28494	33	FJ918580
Grylloblatta sp. Central Sierra Nevada, Grylloblatta sp. Graveyard Lakes	Southern Sierra Nevada	**Graveyard Lakes**	37.44576	-118.97408	34	FJ918591, FJ918592, FJ918593, FJ918599
Grylloblatta sp. Tioga Crest	Northern Sierra Nevada	Greenstone Lake	37.97636	-119.29442	35	FJ918580, FJ918625
Grylloblatta chandleri	gurneyi + chandleri	Gray's Flat	40.628	-121.18932	36	FJ918575
Grylloblatta campodeiformis	campodeiformis	Holland Lookout	47.46465	-113.58217	37	MG824787, MG824847
Grylloblatta newberryensis	rothi	Ice Cave N Newberry Crater*	43.94	-121.27	39	KF880958
Grylloblatta sp. Trout Lake	chirurgica	Ice Caves Picnic Area	45.96123	-121.63244	40	DQ457343, DQ457344
Grylloblatta sp. Trout Lake	chirurgica	Ice Caves Picnic Area Trail	45.96012	-121.61714	41	MG824815
Grylloblatta sp. Ostrander Lake	Southern Sierra Nevada	Indian Cave	37.74816	-119.54898	42	GU013769
Grylloblatta newberryensis	rothi	Karen Cave	43.8923	-121.6973	44	MG824824
Grylloblatta sp. Central Sierra Nevada	Southern Sierra Nevada	Lake George	37.59764	-119.01425	45	FJ918619, MG824761
Grylloblatta campodeiformis	campodeiformis	Lake O'Hara	51.4017	-116.3437	46	MG824817
Grylloblatta chandleri	gurneyi + chandleri	Lassen Peak	40.47456	-121.50604	47	MG824774, MG824775, MG824776, MG824777, MG824778, MG824803, MG824804, MG824805, MG824830
Grylloblatta gurneyi	gurneyi + chandleri	Lava Beds NM	41.71831	-121.50561	48	MG824767
Grylloblatta campodeiformis nahanni	campodeiformis	Limestone Peak	59.25688	-129.83569	50	MG824750
Grylloblatta sp. Trout Lake	chirurgica	Little Red River Cave	46.134	-122.218	51	DQ457360
Grylloblatta sp. Olympc Mtns.	chirurgica	Lunch Lake	47.91166	-123.78175	53	MG824809, MG824810, MG824820
Grylloblatta sp. Tioga Crest	Northern Sierra Nevada	Lyell Canyon	37.76657	-119.25471	54	FJ918605
Grylloblatta marmoreus	oregonensis	Marble Valley	41.56687	-123.20684	56	MG824768, MG824769, MG824785, MG824826
Grylloblatta chintimini	rothi	Marys Peak*	44.50434	-123.55123	57	KF880960, DQ457368
Grylloblatta rothis	rothi	McKenzie Pass	44.26	-121.809	58	DQ457341
Grylloblatta gurneyi	gurneyi + chandleri	Merrill Ice Cave*	41.728	-121.547	59	DQ457347, DQ457348
Grylloblatta sp. Mt. Spokane	campodeiformis	Mission Ridge above Seton Portage	50.76355	-122.1698	60	MG824834

Grylloblatta sp. Southwest Sierra Nevada	Southern Sierra Nevada	Upper Monarch Lake	36.44792	-118.55756	61	FJ918604
Grylloblatta sp. North Cascades	North Cascades, campodeiformis	Monte Cristo	47.9766	-121.35822	62	MG824751, MG824753, MG824822, MG824836, MG824846
Grylloblatta sp. Trout Lake, Grylloblatta sp. North Cascades	chirurgica, North Cascades	**Mount Adams**	46.16379	-121.4902	64	MG824757, MG824812, MG824821, MG824827, MG824853
Grylloblatta rothi	rothi	Mount Bachelor	43.98064	-121.67994	65	MG824811,MG824842
Grylloblatta sp. Mt. Spokane	campodeiformis	Mount Baker	48.72919	-121.83435	66	MG824808, MG824834
Grylloblatta sp. Central Sierra Nevada	Southern Sierra Nevada	Mount Clarence-King	36.82479	-118.43382	67	FJ918590
Grylloblatta campodeiformis	N/A	Mount Fidelity	51.2411	-117.7	68	N/A
Grylloblatta rothi	rothi	Mount Hood	45.38869	-121.66738	69	MG824806, MG824807, MG824841
Grylloblatta sp. Southwest Sierra Nevada	Southern Sierra Nevada	Mount Hooper at Harvey Lake	37.30436	-118.91287	70	FJ918595
Grylloblatta rothi	rothi	Mount Jefferson	44.67262	-121.82511	71	MG824789, MG824837, MG824790, MG824791, MG824792, MG824793, MG824835, MG824842
Grylloblatta sp. Trout Lake, Grylloblatta sp. North Cascades	chirurgica, North Cascades	**Mount Rainier**	46.80372	-121.72292	72	MG824821, MG824836, MG824852
Grylloblatta campodeiformis	N/A	Mount Revelstoke	50.92708	-124.34231	73	N/A
Grylloblatta gurneyi	gurneyi + chandleri	Mount Shasta	41.47291	-122.16631	76	MG824796, MG824813
Grylloblatta sp. Mt. Spokane	campodeiformis	Mount Spokane	47.92438	-117.11488	78	MG824783, MG824784, MG824823, MG824843, MG824844, MG824845, MG824859
Grylloblatta sp. North Cascades	North Cascades, campodeiformis	Mount St. Helens	46.17097	-122.22319	79	MG824838, MG824846
Grylloblatta newberryensis	rothi	Mount Thielsen	43.1554	-122.06536	80	MG824795, MG824797, MG824798, MG824839, MG824851, MG824856
Grylloblatta sp. Southwest Sierra Nevada	Southern Sierra Nevada	New Army Pass	36.48459	-118.24693	81	MG824773
Grylloblatta sp. Trout Lake	chirurgica	New Cave	45.965	-121.62	82	DQ457365
Grylloblatta chandleri	gurneyi + chandleri	North Christmas Tree Cave	40.986	-121.564	83	MG824759
Grylloblatta oregonensis, Grylloblatta siskiyouensis	oregonensis	**Oregon Caves NM***	42.098	-123.406	84	DQ457366, JN612961, JN612962

Grylloblatta sp. Southwest Sierra Nevada	Southern Sierra Nevada	Pear Lake	36.59746	-118.66546	85	FJ918595, FJ918597, FJ918603
Grylloblatta sp. North Cascades	North Cascades, campodeiformis	Pebble Creek	46.80914	-121.72609	86	MG824836
Grylloblatta marmoreus	oregonensis	Planetary Dairy Cave, Marble Valley	41.56	-123.197	87	JN612964
Grylloblatta sp. North Cascades, Grylloblatta sp. Mt. Spokane	North Cascades, campodeiformis	**Rainy Lake**	48.4982	-120.73825	90	MG824822, MG824834
Grylloblatta gurneyi	gurneyi + chandleri	Red Tape Cave	41.45677	-121.76723	91	MG824779, MG824780
Grylloblatta sp. Lillburn Cave	Southern Sierra Nevada	Redwood Creek Trail	36.6975	-118.9128	92	MG824814
Grylloblatta sp. Lillburn Cave	Southern Sierra Nevada	Redwood Creek Trail at Lillburn Cave	36.66981	-118.90897	93	FJ918576, FJ918577
Grylloblatta sp. Central Sierra Nevada	Southern Sierra Nevada	Sam Mack Lake	37.11646	-118.51212	94	FJ918598
Grylloblatta sculleni	chirurgica	Santiam Junction	44.43952	-121.94283	95	KF880961
Grylloblatta sculleni	chirurgica	Sawyer's Ice Cave	44.426	-121.981	96	DQ457368
Grylloblatta washoa	Northern Sierra Nevada	SE Barker Pass	39.06661	-120.22999	98	FJ918620, FJ918621, FJ918622, FJ918623, MG824762
Grylloblatta sp. Central Sierra Nevada	Southern Sierra Nevada	Unnamed cirque lake SW of Selden Pass	37.28555	-118.88175	99	FJ918600, FJ918601
Grylloblatta sp. Mt. Spokane	campodeiformis	Sicamouse Creek	50.81111	-118.9587	102	MG824845
Grylloblatta sp. Sierra Buttes	barberi	Sierra Buttes	39.60174	-120.65625	103	MG824764
Grylloblatta sp. Southwest Sierra Nevada	Southern Sierra Nevada	Silliman Lake	36.63881	-118.69416	104	FJ918595, FJ918596
Grylloblatta washoa	Northern Sierra Nevada	Smith Lake	38.85681	-120.18805	105	FJ918578, FJ918579
Grylloblatta bifratrilecta	Northern Sierra Nevada	Sonora Pass*	38.31556	-119.66673	106	FJ918584, FJ918585, FJ918586, FJ918587
Grylloblatta sp. Mt. Spokane	campodeiformis	South Cascade Lake	48.36577	-121.06914	107	MG824834
Grylloblatta rothi	rothi	South Sister	44.06797	-121.754	108	MG824794, MG824840, MG824842
Grylloblatta sp. Central Sierra Nevada	Southern Sierra Nevada	Sphinx Lakes	36.71427	-118.51487	109	FJ918590, FJ918602
Grylloblatta chandleri	gurneyi + chandleri	Subway Cave	40.68493	-121.41901	110	MG824760
Grylloblatta gurneyi	gurneyi + chandleri	Sugar Pine Butte Ice Cave	41.40442	-122.02653	111	MG824782
Grylloblatta campodeiformis	campodeiformis	Sulphur Mountain*	51.14517	-115.5739	112	MG824847
Grylloblatta chirurgica	chirurgica	Surprise Cave	46.174	-122.216	113	DQ457364
Grylloblatta sp. North Cascades	North Cascades	Surprise Mountain	47.66508	-121.14173	114	MG824836
Grylloblatta washoa	Northern Sierra Nevada	Susie Lake	38.88574	-120.1345	115	FJ918607, FJ918608, FJ918609, FJ918610, FJ918611, FJ918612, FJ918613, FJ918614, FJ918615, FJ918616

Grylloblatta sp. North Cascades, Grylloblatta sp. Mt. Spokane	North Cascades, campodeiformis	**Table Mountain**	48.85504	-121.69711	116	MG824754, MG824755, MG824834, MG824836
Grylloblatta campodeiformis	campodeiformis	Takakkaw Falls	51.48581	-116.4766	117	MG824758
Grylloblatta oregonensis, Grylloblatta marmoreus	oregonensis	**Thompson Peak**	41.00307	-123.05009	119	MG824770, MG824771, MG824772, MG824799, MG824800, MG824801, MG824802, MG824818, MG824819, MG824832, MG824850, MG824860
Grylloblatta sp. North Cascade	North Cascades	Thornton Lakes	48.69096	-121.33401	120	MG824756, MG824836
Grylloblatt rothi	rothi	Tilly Jane Creek	45.39242	-121.66016	121	MG824841
Grylloblatta sp. Tioga Crest	Northern Sierra Nevada	Vogelsang Lake	37.78619	-119.34689	122	FJ918582
Grylloblatta sp. Mt. Spokane	campodeiformis	Vowell Creek Valley	50.952	-115.99064	123	MG824845
Grylloblatta sp. Central Sierra Nevada	Southern Sierra Nevada	White Mountains	37.58823	-118.23377	125	FJ918588, FJ918589
Grylloblatta chandleri	gurneyi + chandleri	Wilson Ice Cave	40.33592	-121.42169	126	DQ457339, DQ457340
Grylloblattina djakonovi djakonovi	Outgroups	Petrov Island, Russia				DQ457355, DQ457356
Grylloblattina djakonovi kurentzovi	Outgroups	Beriozovii Stream, Russia				DQ457357
Grylloblattina djakonovi kurentzovi	Outgroups	Mt. Krinitshnaya, Russia				DQ457358
Galloisiana yuasai	Outgroups	Tokugo Pass, Japan*				DQ457350
Galloisiana sp. Kakuma	Outgroups	Kakuma Valley, Japan				DQ457351, DQ457352
Galloisiana nipponensis	Outgroups	Lake Chuzenji, Japan*				DQ457353, DQ457354
Galloisiana sinensis	Outgroups	China, Changbaishan*	42.03954	128.05562		MG824857
Tyrannophasma gladiator	Outgroups	Namibia, South Africa				DQ457369

Table S2. List of the 322 single copy orthologs used for nuclear gene phylogeny, with length in base pairs and functional annotation.

HAMSTR Ortholog #	Max Sequence Length	Annotation
411851	1118	vacuolar protein-sorting protein
411852	606	mitochondrial import inner membrane translocase, subunit TIM23
411853	1379	vacuolar assembly/sorting protein VPS9
411858	498	Ufm1-conjugating enzyme
411870	528	NADH dehydrogenase
411875	1072	replication factor C, subunit RFC3
411877	2131	dystroglycan
411881	1371	ribosome biogenesis protein - Nop58p/Nop5p
411883	735	pur-alpha
411889	984	malate dehydrogenase
411892	653	60S ribosomal protein L10A
411896	1068	splicing factor u2af large subunit
411899	934	conserved hypothetical protein
411901	834	fatty acyl-CoA elongase
411905	1004	failed axon connections
411907	995	secreted protein
411909	795	WW domain-binding protein
411915	1875	SEL-1
411916	2413	rad25/xp-B DNA repair helicase
411917	429	40S ribosomal protein S23
411935	2767	heat shock protein 70 (HSP70)-interacting protein
411938	444	actin depolymerizing factor
411945	562	U4/U6 small nuclear ribonucleoprotein Prp31
411953	2151	translation elongation factor EF-1 alpha/Tu
411955	836	carbon-nitrogen hydrolase
411956	1010	O-sialoglycoprotein endopeptidase
411958	2872	glycine dehydrogenase
411965	447	calmodulin
411968	404	clathrin adaptor complex, small subunit
411972	1378	26S proteasome regulatory complex, subunit RPN5/PSMD12
411973	750	ATP synthase B chain
411981	609	RNA-binding protein musashi
411985	428	NADH:ubiquinone oxidoreductase, NDUFB9/B22 subunit
411991	798	conserved hypothetical protein
411992	847	SPRY domain SOCS box-containing protein
411997	727	ubiquitin fusion-degradation protein
411998	1014	folate carrier protein
412000	1870	neuroendocrine convertase
412020	1247	translation initiation factor 4F, helicase subunit
412021	996	coproporphyrinogen III oxidase
412031	1453	6-phosphogluconate dehydrogenase
412038	330	ubiquinol cytochrome C reductase, subunit QCR7
412044	431	prefoldin
412045	384	box H/ACA snoRNP
412047	379	conserved hypothetical protein
412051	1096	seven transmembrane receptor
412054	517	MACRO domain-containing protein
412058	978	succinyl-CoA synthetase, alpha subunit
412064	643	N(6)-adenine-specific DNA methyltransferase
412067	1975	armadillo repeat protein
412071	437	asparaginase

412075	697	membrane protein
412076	1066	sorbitol dehydrogenase
412085	460	NEFA-interacting nuclear protein NIP30
412088	583	RWD domain-containing protein 4A
412090	1059	conserved hypothetical protein
412091	1761	kinesin
412093	1057	protein arginine N-methyltransferase PRMT1
412095	1176	heat shock protein 70 (HSP70)-interacting protein
412096	990	adipophilin
412101	687	cell division protein FtsJ
412109	531	lysophosphatidic acid acyltransferase endophilin
412110	2770	4SNc-Tudor domain protein
412116	877	conserved hypothetical protein
412123	501	protein involved in membrane traffic
412136	636	conserved hypothetical protein
412140	756	ubiquitin-conjugating enzyme
412141	810	mannosyl-oligosaccharide glucosidase
412143	498	alpha catenin
412146	2652	eukaryotic translation initiation factor 3 subunit C
412148	848	esterase D
412165	2579	elongation factor
412166	1862	intermediate peptidase
412169	922	geranylgeranyl diphosphate synthase
412170	1218	FK506 binding protein (FKBP)
412172	619	conserved hypothetical protein
412174	1212	ribosomal protein L4
412185	864	17 beta-hydroxysteroid dehydrogenase
412200	1326	succinyl-CoA synthetase, beta subunit
412207	647	mitotic spindle assembly checkpoint protein MAD2A
412208	1206	S-adenosylmethionine synthetase
412216	780	caspase, apoptotic cysteine protease
412217	552	striatin
412220	411	coiled-coil protein
412223	727	mRNA-decapping enzyme
412224	929	rfx transcription factor
412226	651	conserved hypothetical protein
412235	880	protein kinase
412240	432	ubiquitin-conjugating enzyme
412248	1779	conserved hypothetical protein
412257	576	conserved hypothetical protein
412258	991	arsenite-translocating ATPase
412268	471	suppressor of cytokine signaling-2
412271	537	ADP-ribosylation factor, ARF
412279	563	ribosomal protein L43
412282	1157	C subunit of V-ATPase
412299	549	RNA-binding protein musashi
412304	714	U2 snRNP splicing factor, small subunit
412307	882	nitric oxide synthase interacting protein
412317	456	tetratricopeptide repeat protein
412319	885	transcription factor S-II
412323	630	RAS-related protein
412324	1026	transferase
412325	1746	conserved hypothetical protein
412329	923	RING finger protein
412332	508	signal sequence receptor delta

412337	864	heparan sulfate 2-O-sulfotransferase
412343	822	prohibitin
412347	1092	fructose 1,6-bisphosphate aldolase
412354	720	protein pob
412355	1268	3-phosphoglycerate kinase
412356	1854	65-kDa macrophage protein
412359	1402	citrate synthase
412365	541	splicing factor SPF30
412383	1467	protein disulfide isomerase
412389	2952	lysosomal alpha-mannosidase
412392	1462	adenylosuccinate lyase
412400	1998	vacuolar protein-sorting protein
412401	252	histone H4
412407	978	conserved hypothetical protein
412411	948	Na+/K+ ATPase, beta subunit
412412	588	conserved hypothetical protein
412413	1080	branched chain alpha-keto acid dehydrogenase
412418	546	PUA domain-containing RNA-binding protein
412420	2465	nuclear pore complex protein Nup93
412422	3430	fh1/fh2 domains-containing protein
412425	1137	beta-ureidopropionase
412428	1530	vesicle coat complex COPI, beta' subunit
412432	1665	T-cell immunomodulatory protein
412438	748	20S proteasome, regulatory subunit alpha type PSMA7/PRE6
412442	1261	protein phosphatase 2C
412446	706	ribosomal protein L16
412449	390	conserved hypothetical protein
412451	1008	conserved hypothetical protein
412463	577	anti-silencing protein
412474	510	mediator of RNA polymerase II transcription subunit
412475	1402	translation elongation factor
412476	504	conserved hypothetical protein
412478	645	glutathione S-transferase
412500	732	reductase
412516	1055	60S ribosomal protein L3
412541	1359	ATP-dependent RNA helicase pitchoune
412551	1979	chromosome 9 open reading frame
412552	996	glyceraldehyde 3-phosphate dehydrogenase
412552	996	glyceraldehyde 3-phosphate dehydrogenase
412565	880	centaurin alpha
412566	1428	protein transport protein SEC61 alpha subunit
412572	1188	conserved hypothetical protein
412585	2103	translation initiation factor 3, subunit B
412592	409	Core histone H2A/H2B/H3/H4
412595	620	cargo transport protein EMP24
412598	951	pseudouridylate synthase
412609	3984	lipophorin receptor
412612	300	cysteine-rich PDZ-binding protein
412613	3313	ubiquitin carboxyl-terminal hydrolase
412617	1008	arginase
412621	3007	E1-E2 ATPase
412625	487	microspherule protein
412626	1433	nucleolar complex protein
412631	495	extended synaptotagmin-2
412654	4037	U2 small nuclear ribonucleoprotein

412658	564	calponin
412665	383	p53 and DNA damage-regulated protein
412670	756	conserved hypothetical protein
412671	1260	conserved hypothetical protein
412676	865	secreted inorganic pyrophosphatase
412677	1002	lipoic acid synthase
412690	1650	F0F1-type ATP synthase, alpha subunit
412694	453	ubiquitin protein ligase
412699	1302	presenilin
412701	378	transcription initiation factor TFII-D component
412709	395	40S ribosomal protein S30
412711	1734	negative elongation factor C/D
412712	527	29 kDa ribonucleoprotein B
412731	1173	transcription factor NFAT, subunit NF45
412733	630	translation initiation factor 4F, cap-binding subunit
412738	962	hypothetical protein
412742	1053	dihydroorotate dehydrogenase
412747	312	60S ribosomal protein L44
412748	612	ribosomal protein L13A
412756	336	transcription initiation factor IIA gamma chain
412757	998	Shk1 kinase-binding protein (Skb1)
412765	1504	cpg binding protein
412766	930	P53 induced protein
412768	776	CDK inhibitor P21 binding protein
412769	918	cytochrome c1
412774	1514	inositol-1-phosphate synthetase
412777	868	translation initiation factor 3, subunit G
412780	1611	ATP-dependent RNA helicase pitchoune
412790	720	glutathione S-transferase
412794	1521	UTP-glucose-1-phosphate uridylyltransferase
412796	1376	secreted protein
412799	848	ribosomal DEAD box protein
412819	3008	sarm1
412820	1953	phosphoenolpyruvate carboxykinase
412821	751	40S ribosomal protein S6
412826	1383	serine/threonine protein kinase
412830	673	coatomer gamma subunit
412832	660	calcyclin-binding protein CacyBP
412834	468	N-terminal Asn amidase
412839	586	methyl-CpG binding transcription regulator
412840	471	NADH-ubiquinone oxidoreductase, subunit NDUFB10/PDSW
412846	1628	heat shock protein
412847	951	60S acidic ribosomal protein P0
412851	831	dihydrolipoamide acetyltransferase
412852	639	V-SNARE
412860	543	U3 small nucleolar ribonucleoprotein (snoRNP) component
412862	774	ribosomal protein S28
412864	1350	alpha tubulin
412865	880	carnitine-acylcarnitine carrier protein
412870	324	FKBP-type peptidyl-prolyl cis-trans isomerase
412871	625	synaptogyrin
412872	999	adenosine A3 receptor
412873	375	MP1 adaptor interacting protein P14
412874	1662	guanylate-binding protein
412875	741	adenylate kinase

412878	586	alkyl hydroperoxide reductase, thiol specific antioxidant
412880	650	Plexin domain-containing protein
412884	442	40S ribosomal protein S15
412895	1070	heat shock protein
412897	2034	long chain fatty acid CoA ligase
412902	927	conserved hypothetical protein
412911	1062	glycerol-3-phosphate dehydrogenase
412923	686	conserved hypothetical protein
412926	1039	WD-repeat protein
412934	432	ribosomal protein L27
412937	552	60S ribosomal protein L23
412943	984	poly(p)/ATP NAD kinase
412944	702	RAB-23
412945	843	hypothetical protein
412951	1533	coatomer delta subunit
412959	667	N-acetyltransferase
412963	702	ribose-5-phosphate isomerase
412971	339	SCF ubiquitin ligase, Rbx1 component
412972	1571	glucose 6-phosphate dehydrogenase
412983	1130	adhesion regulating molecule
412984	2199	dipeptidyl peptidase IV
412994	453	ubiquitin protein ligase
412999	899	steroid reductase
413001	2238	F-spondin
413004	340	parafibromin
413006	546	conserved hypothetical protein
413010	3514	chromodomain helicase DNA binding protein
413015	1171	serine/threonine protein phosphatase 2A regulatory subunit B" subunit gamma
413019	1113	signal peptide peptidase
413020	618	DNAJ protein
413021	657	phosphoserine phosphatase
413025	1080	leucine rich domain-containing protein
413032	1395	GPI mannosyltransferase
413034	764	short-chain alcohol dehydrogenase
413053	1340	arsenite-resistance protein
413059	783	actin-bundling protein
413060	1278	sodium/calcium exchanger
413072	570	Ca2+ sensor
413073	789	G2/mitotic-specific cyclin A
413074	645	manganese superoxide dismutase
413080	1542	conserved hypothetical protein
413083	336	tubulin-specific chaperone A
413085	310	heat shock protein
413088	393	ribosomal protein S17
413095	1905	apoptosis-promoting RNA-binding protein TIA-1/TIAR
413101	2629	26S proteasome regulatory subunit rpn1
413108	646	vesicle coat protein clathrin, light chain
413126	612	ribosomal protein L15
413130	681	perixosomal biogenesis factor
413141	321	30S ribosomal protein S12
413147	759	nuclear pore complex protein Nup205
413153	864	thioredoxin H2 protein
413155	2096	conserved hypothetical protein
413168	942	asparaginase
413178	1044	vacuolar H+ ATPase

413180	1944	signal recognition particle protein
413187	2158	RNA helicase
413192	438	conserved hypothetical protein
413202	786	reductase
413207	594	guanylate kinase
413210	1491	vacuolar H+-ATPase V1 sector, subunit B
413216	417	deoxyuridine 5'-triphosphate nucleotidohydrolase
413226	1053	conserved hypothetical protein
413230	1090	glyoxylate/hydroxypyruvate reductase
413236	495	ADP ribosylation factor 79F
413242	2499	ATP-dependent protease PIM1/LON
413247	1734	translation elongation factor EF-1 alpha/Tu
413248	1692	conserved hypothetical protein
413250	1513	conserved hypothetical protein
413257	1526	carbohydrate kinase
413262	745	aminopeptidase
413269	648	conserved hypothetical protein
413271	1035	membrane protein
413273	462	superoxide dismutase Cu-Zn
413275	2332	vesicle coat complex COPII, subunit SEC23
413284	404	conserved hypothetical protein
413285	1164	medium-chain acyl-CoA dehydrogenase
413286	1101	phosphoserine aminotransferase
413289	822	methyltransferase
413300	1448	conserved hypothetical protein
413304	654	thioredoxin domain-containing protein
413314	624	bhlh transcription factor
413316	1185	conserved hypothetical protein
413317	2029	vitellogenin receptor
413322	879	RING finger motif containing protein
413324	1224	U3 snoRNP-associated protein
413335	1137	alpha-methylacyl-CoA racemase
413336	563	actin
413340	754	oxidative stress-induced growth inhibitor
413347	453	ribosomal protein L26
413350	1115	1-acyl-glycerol-3-phosphate acyltransferase
413351	546	ADP-ribosylation factor
413359	656	conserved hypothetical protein
413365	210	G-protein gamma subunit
413368	1834	conserved hypothetical protein
413369	538	signal peptidase
413371	558	ADP-ribosylation factor
413372	1057	RNA-binding protein
413392	1376	thioredoxin reductase
413394	1110	conserved hypothetical protein
413397	1245	N2,N2-dimethylguanosine tRNA methyltransferase
413398	387	U3 small nucleolar ribonucleoprotein imp4
413402	350	molecular chaperone
413405	288	small nuclear ribonucleoprotein (snRNP) splicing factor
413407	774	hypothetical protein
413413	325	cytochrome C
413419	522	membrane protein

Table S3. Collection sites in the nuclear gene tree with latitude and longitude coordinates and corresponding location numbers used on the map. Genbank accessions are available in Bioproject PRJNA488740. An asterisk (*) indicates a holotype or neotype locality, while bold type indicates sympatric lineages at a site.

Taxon	Location	Latitude	Longitude	Location Number	Major Clade	Specimen IDs
Grylloblatta campodeiformis	Allison Peak	49.73802	-114.6302	2	MC1	SDS15-030A
Grylloblatta chirurgica	Ape Cave*	46.10794	-122.21149	3	MC3	SDS14-033H
Grylloblatta campodeiformis	Avalanche Lake	48.65538	-113.78838	4	MC1	SDS14-015
Grylloblatta sp. North Cascades	Black Tusk	49.9751	-123.0431	8	MC2	BlackTusk
Grylloblatta sp. Mt. Spokane	Blowdown Pass, Stein Valley	50.3626	-122.15	10	MC1	ENTO12_008744
Grylloblatta sp. Olympic Mountains	Bogachiel Peak	47.90546	-123.78304	12	MC3	SDS14-480
Grylloblatta sp. Ostrander Lake	Bridalveil trail to Ostrander Lake	37.63543	-119.58054	13	MC8	SDS05-580
Grylloblatta campodeiformis	Buffalo Horn Creek	45.10403	-111.20861	14	MC1	SDS14-011A
Grylloblatta bifratrilecta	Carson Pass	38.6671	-119.98962	16	MC5	SDS07-006E
Grylloblatta sp. Mt. Spokane	Cascade Pass	48.46938	-122.06721	17	MC1	SDS14-349C
Grylloblatta chirurgica, Grylloblatta sp. North Cascades	**Chinook Pass**	46.88307	-121.51062	19	MC2, MC3	SDS14-022A, SDS14-022E, SDS13-022H
Grylloblatta rothi	Crossbill Ice Cave	43.779	-121.261	21	MC6	SDS13-553
Grylloblatta barberi	E Caribou Powerhouse*	40.09134	-121.13687	26	MC7	SDS08-001
Grylloblatta chandleri	Eagle Lake*	40.64554	-120.7946	27	MC4	SDS13-530B
Grylloblatta sp. Mt. Spokane	Elsie Lake	47.42846	-116.02541	28	MC1	SDS13-551
Grylloblatta sp. North Cascades	Glacier Peak	48.04945	-121.15511	31	MC2	SDS14-454
Grylloblatta sp. Tioga	Granite Lakes	37.92809	-119.28494	33	MC5	SDS05-0157
Grylloblatta sp. Central Sierra Nevada	Graveyard Lakes	37.44576	-118.97408	34	MC8	SDS06-068C
Grylloblatta chandleri	Grey's Flat	40.628	-121.18932	36	MC4	GREYFLAT
Grylloblatta campodeiformis	Holland Lookout	47.46465	-113.58217	37	MC1	SDS14-013A
Grylloblatta sp. Trout Lake, Grylloblatta sp. North Cascades	**Ice Cave Picnic Area**	45.96123	-121.63244	40	MC2, MC3	SDS14-028A, SDS14-028D
Grylloblatta rothi	Karen Cave	43.8929	-121.699	44	MC6	SDS13-521A
Grylloblatta sp. Central Sierra Nevada	Lake George	37.59764	-119.01425	45	MC8	SDS07-014
Grylloblatta campodeiformis	Lake O'Hara	51.4017	-116.3437	46	MC1	ENTO13_007677A
Grylloblatta chandleri	Lassen Peak	40.46656	-121.51067	47	MC4	SDS14-227B
Grylloblatta gurneyi	Lava Beds NM*	41.71831	-121.50561	48	MC4	SDS13-009A
Grylloblatta campodeiformis nahanni	Limestone Peak*	59.25688	-129.83569	50	MC1	SDS14-540
Grylloblatta campodeiformis	Lunch Creek	48.7038	-113.70258	52	MC1	SDS15-010B
Grylloblatta sp. Olympic Mountains	Lunch Lake	47.91166	-123.78175	53	MC3	SDS14-476A
Grylloblatta campodeiformis	MacArthur Lake	51.3329	-116.3402	55	MC1	ENTO13-0077091

Grylloblatta marmoreus	Marble Valley	41.56687	-123.20684	56	MC9	SDS13-019D
Grylloblatta sp. Mt. Spokane	Mission Ridge above Seton Portage	50.76355	-122.1698	60	MC1	ENTO11-010819
Grylloblatta sp. North Cascades	Monte Cristo Peak	47.97966	-121.36305	62	MC2	SDS14-316A
Grylloblatta sp. Trout Lake	Mount Adams	46.16379	-121.4902	64	MC3	SDS15-206A, SDS15-212
Grylloblatta rothi	Mount Bachelor	43.98064	-121.67994	65	MC6	SDS15-109
Grylloblatta sp. North Cascades	Mount Baker	48.72919	-121.83435	66	MC2	SDS14-380A
Grylloblatta campodeiformis	Mount Fidelity	51.2411	-117.7	68	MC1	ENTO12-008092_3
Grylloblatta rothi	Mount Hood	45.38869	-121.66738	69	MC6	SDS13-444B
Grylloblatta rothi	Mount Jefferson	44.67262	-121.82511	71	MC6	SDS14-066D, SDS14-067, SDS14-080B
Grylloblatta sp. North Cascades	Mount Rainier	46.80372	-121.72292	72	MC2	SDS14-296I
Grylloblatta campodeiformis	Mount Revelstoke Ski Resort	50.96318	-118.10114	74	MC1	ENTO12-008551
Grylloblatta gurneyi	Mount Shasta	41.47291	-122.16631	76	MC4	SDS14-147A
Grylloblatta sp. Southwest Sierra Nevada	Mount Silliman, SE upper lake	36.63881	-118.69416	77	MC8	SDS06-092F
Grylloblatta sp. Mt. Spokane	Mount Spokane	47.92438	-117.11488	78	MC1	SDS13_536
Grylloblatta sp. North Cascades	Mount St. Helens	46.17097	-122.22319	79	MC2	SDS14-035C
Grylloblatta newberryensis	Mount Thielsen	43.1554	-122.06536	80	MC6	SDS14-133A
Grylloblatta sp. Central Sierra Nevada	New Army Pass	36.48459	-118.24693	81	MC8	SDS13-056
Grylloblatta sp. North Christmas Tree Cave	North Christmas Tree Cave	40.986	-121.564	83	MC4	GG01
Grylloblatta oregonensis, Grylloblatta siskiyouensis	**Oregon Caves***	42.09571	-123.40332	84	MC9	SDS10-002B, SDS10-002C
Grylloblatta sp. Mt. Spokane	Pebble Creek	46.80914	-121.72609	86	MC1	SDS13-380A
Grylloblatta marmoreus	Planetary Dairy Cave, Marble Valley *	41.56	-123.197	87	MC9	SDS11-002
Grylloblatta sp. Mt. Spokane	Porcupine Lake	49.27091	-117.02075	88	MC1	ENTO13-003218-8
Grylloblatta sp. Mt. Spokane	Rainy Lake	48.4982	-120.73825	90	MC1	SDS14-328E
Grylloblatta gurneyi	Red Tape Cave	41.45677	-121.76723	91	MC4	SDS13-515A
Grylloblatta sp. Central Sierra Nevada	Sam Mack Lake	37.11646	-118.51212	94	MC8	SDS06-109
Grylloblatta campodeiformis	Sentinel Pass	51.33236	-116.22242	100	MC1	SDS14-594
Grylloblatta campodeiformis	Sicamouse Creek	50.81111	-118.9587	102	MC1	DH_02
Grylloblatta sp. Sierra Buttes	Sierra Buttes	39.60174	-120.65625	103	MC7	SDS13-002G
Grylloblatta bifratrilecta	Sonora Pass*	38.31556	-119.66673	106	MC5	SDS06-014B
Grylloblatta sp. North Cascades	South Cascade Lake	48.36577	-121.06914	107	MC2	SDS13-407B
Grylloblatta rothi	South Sister	44.06797	-121.754	108	MC6	SDS14-088A, SDS14-096C
Grylloblatta chandleri	Subway Cave	40.68493	-121.41901	110	MC4	GG02
Grylloblatta gurneyi	Sugar Pine Butte Ice Cave	41.40442	-122.02653	111	MC4	SDS13-516
Grylloblatta campodeiformis	Sulphur Mountain*	51.14517	-115.5739	112	MC1	SDS14-617B
Grylloblatta sp. North Cascades	Surprise Mountain	47.66508	-121.14173	114	MC2	SDS14-297

Grylloblatta sp. Mt. Spokane, Grylloblatta sp. North Cascades	**Table Mountain**	48.85243	-121.68925	116	MC1, MC2	SDS13-390A, SDS13-395B, SDS14-047A, SDS14-047B, SDS14-399
Grylloblatta campodeiformis	Takakkaw Falls	51.48581	-116.4766	117	MC1	ENTO13-007732B
Grylloblatta oregonensis	Thompson Peak	41.01723	-123.05129	119	MC9	SDS13-026M
Grylloblatta sp. North Cascades	Thornton Lakes	48.69096	-121.33401	120	MC2	SDS14-335D
Grylloblatta campodeiformis	Vowell Creek Valley	50.952	-115.99064	123	MC1	VowellCreek
Grylloblatta sp. North Cascades	White Chuck Mountain	48.21517	-121.42823	124	MC2	SDS13-431A
Grylloblatta sp. Central Sierra Nevada	White Mountains	37.58823	-118.23377	125	MC8	SDS06-018D
Namkungia sp.	Mount Seolaksan NP, South Korea	38.215444	128.438	Outgroup	Outgroup	BK_001
Galloisiana odaesanensis	Mosangul Cave, South Korea	36.154	128.686	Outgroup	Outgroup	SDS10K-002
Galloisiana sp.	Mt. Tsurugi, Shikoku, Japan	33.86727	134.08855	Outgroup	Outgroup	SDS09JP-303
Grylloblattella sp.	Pustag Mountain, Russia	53.03493	88.00301	Outgroup	Outgroup	SDS13-178A
Grylloblattella sayanensis	Karasaibo River, Russia*	52.15164	90.18011	Outgroup	Outgroup	SDS13-216A
Grylloblattella pravdini	Tevenek River, Russia*	51.79202	87.32421	Outgroup	Outgroup	SDS13-103A
Grylloblattella sp.	Holzun Mountains, Russia	50.20821	84.55408	Outgroup	Outgroup	SDS12-059A

Table S4. Results for phylogenetic model selection in BEAST of the COII dataset using AICM.

	AICM	S.E.	Nonpartitioned	Partitioned	Yule	Birth Death
Nonpartitioned	23603.753	+/- 1.863	-	-147.277	-	-
Partitioned	23751.03	+/- 8.14	147.277	-	-	-
Yule	23725.643		-	-	-	229.029
Birth Death	23496.614	+/-0.258	-	-	-229.029	-

Table S5. Pairwise differences among trees using the Kendall Colijn metric. Below the diagonal, the pairwise distance is based on topological differences only. Above the diagonal, the pairwise distance among trees is based on branch lengths (species tree not included as branch lengths were measured in coalescent units).

	Full dataset	30% missing dataset	20% missing dataset	10% missing dataset	5% missing dataset	Species tree
Full dataset	0	0.0300151	0.0812129	0.2021176	0.8538241	N/A
30% missing dataset	14.45683	0	0.0520087	0.2299312	0.8819482	N/A
20% missing dataset	12.20656	9.69536	0	0.2808592	0.9330955	N/A
10% missing dataset	15.19868	18.60108	15.6205	0	0.6535073	N/A
5% missing dataset	36.22154	38.8201	37.53665	33.92639	0	N/A
Species tree	79.51729	78.56208	79.05694	82.31646	81.38182	0

Table S6. Estimated root age for well-supported *Grylloblatta* lineages examined in the ancestral range reconstruction and demographic analysis.

Species or Clade	Mean	Median	Lower 95% HPD	Upper 95% HPD	Calibrated Mean	Calibrated Median	Calibrated Lower 95% HPD	Calibrated Upper 95% HPD
G. campodeiformis	1.325	1.3099	0.8956	1.7943	1,325,000	1,309,900	895,600	1,794,300
G. chandleri	1.091	1.078	0.6803	1.5372	1,091,000	1,078,000	680,300	1,537,200
G. chirurgica	0.1359	0.1204	0.0142	0.295	135,900	120,400	14,200	295,000
G. gurneyi	0.5565	0.5436	0.3107	0.8193	556,500	543,600	310,700	819,300
G. sp. 'Southwest Sierra Nevada'	0.8132	0.8007	0.5038	1.1435	813,200	800,700	503,800	1,143,500
G. marmoreus	0.8807	0.8669	0.5296	1.2666	880,700	866,900	529,600	1,266,600
G. newberryensis	0.6151	0.5941	0.3054	0.9254	615,100	594,100	305,400	925,400
G. rothi	0.457	0.4423	0.2051	0.7326	457,000	442,300	205,100	732,600
G. sp. 'North Cascades'	0.3374	0.325	0.144	0.5389	337,400	325,000	144,000	538,900
G. sp. 'Central Sierra Nevada'	0.8692	0.8496	0.5355	1.2496	869,200	849,600	535,500	1,249,600
G. sp. 'Mt. Spokane'	0.2406	0.2281	0.0867	0.4164	240,600	228,100	86,700	416,400
G. sp. 'Trout Lake'	1.2154	1.2051	0.8121	1.6459	1,215,400	1,205,100	812,100	1,645,900
G. washoa	0.565	0.5505	0.316	0.837	565,000	550,500	316,000	837,000

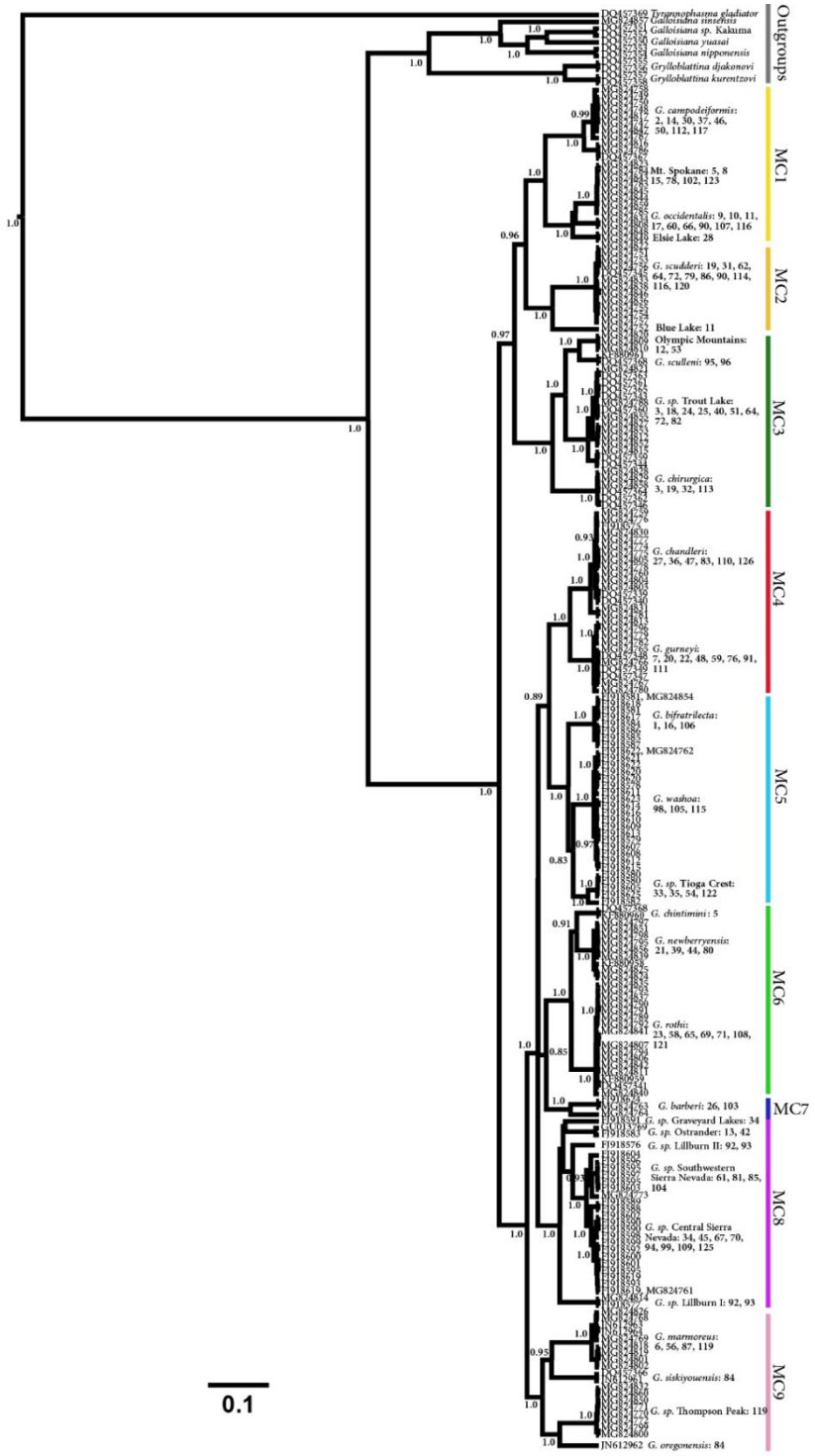

Figure S1. Bayesian phylogenetic relationships of *Grylloblatta* inferred from the COII mitochondrial gene. Numbers and names at the tips refer to sample sites and known taxon names, respectively. Well-supported genetic lineages are grouped into major clades (MC), represented as colored groups and referred to in the text. Nodal support values for the phylogeny are shown as Bayesian posterior probability estimates and branch lengths are in units of substitutions per site.

Species	Site Name	Site Number	Species	Site Name	Site Number	Species	Site Name	Site Number
Grylloblatta campodeiformis	Allison Peak	2	Grylloblatta campodeiformis nahanni	Limestone Peak*	50	Grylloblatta sp. Mt. Spokane	Rainy Lake	90
Grylloblatta chirurgica	Ape Cave*	3	Grylloblatta campodeiformis	Lunch Creek	52	Grylloblatta gurneyi	Red Tape Cave	91
Grylloblatta campodeiformis	Avalanche Lake	4	Grylloblatta sp. Olympic Mountains	Lunch Lake	53	Grylloblatta sp. Central Sierra Nevada	Sam Mack Lake	94
Grylloblatta sp. North Cascades	Black Tusk	8	Grylloblatta campodeiformis	MacArthur Lake	55	Grylloblatta campodeiformis	Sentinel Pass	100
Grylloblatta sp. Mt. Spokane	Blowdown Pass, Stein Valley	10	Grylloblatta marmoreus	Marble Valley	56	Grylloblatta campodeiformis	Sicamouse Creek	102
Grylloblatta sp. Olympic Mountains	Bogachiel Peak	12	Grylloblatta campodeiformis	Mission Ridge above Seton Portage	60	Grylloblatta sp. Sierra Buttes	Sierra Buttes	103
Grylloblatta sp. Ostrander Lake	Bridalveil trail to Ostrander Lake	13	Grylloblatta sp. North Cascades	Monte Cristo Peak	62	Grylloblatta bifratrilecta	Sonora Pass*	106
Grylloblatta campodeiformis	Buffalo Horn Creek	14	Grylloblatta sp. Trout Lake	Mount Adams	64	Grylloblatta sp. North Cascades	South Cascade Lake	107
Grylloblatta bifratrilecta	Carson Pass	16	Grylloblatta rothi	Mount Bachelor	65	Grylloblatta rothi	South Sister	108
Grylloblatta sp. Mt. Spokane	Cascade Pass	17	Grylloblatta sp. North Cascades	Mount Baker	66	Grylloblatta chandleri	Subway Cave	110
Grylloblatta chirurgica, Grylloblatta sp. North Cascades	**Chinook Pass**	19	Grylloblatta campodeiformis	Mount Fidelity	68	Grylloblatta gurneyi	Sugar Pine Butte Ice Cave	111
Grylloblatta rothi	Crossbill Ice Cave	21	Grylloblatta rothi	Mount Hood	69	Grylloblatta campodeiformis	Sulphur Mountain*	112
Grylloblatta barberi	E Caribou Powerhouse*	26	Grylloblatta rothi	Mount Jefferson	71	Grylloblatta sp. North Cascades	Surprise Mountain	114
Grylloblatta chandleri	Eagle Lake*	27	Grylloblatta sp. North Cascades	Mount Rainier	72	Grylloblatta sp. Mt. Spokane	Table Mountain	116
Grylloblatta sp. Mt. Spokane	Elsie Lake	28	Grylloblatta campodeiformis	Mount Revelstoke Ski Resort	74	Grylloblatta campodeiformis	Takakkaw Falls	117
Grylloblatta sp. North Cascades	Glacier Peak	31	Grylloblatta gurneyi	Mount Shasta	76	Grylloblatta oregonensis	Thompson Peak	119
Grylloblatta sp. Tioga	Granite Lakes	33	Grylloblatta sp. Southwest Sierra Nevada	Mount Silliman, SE upper lake	77	Grylloblatta sp. North Cascades	Thornton Lakes	120
Grylloblatta sp. Central Sierra Nevada	Graveyard Lakes	34	Grylloblatta sp. Mt. Spokane	Mount Spokane	78	Grylloblatta campodeiformis	Vowell Creek Valley	123
Grylloblatta chandleri	Grey's Flat	36	Grylloblatta sp. North Cascades	Mount St. Helens	79	Grylloblatta sp. North Cascades	White Chuck Mountain	124
Grylloblatta campodeiformis	Holland Lookout	37	Grylloblatta newberryensis	Mount Thielsen	80	Grylloblatta sp. Central Sierra Nevada	White Mountains	125
Grylloblatta sp. Trout Lake, Grylloblatta sp. North Cascades	**Ice Cave Picnic Area**	40	Grylloblatta sp. Central Sierra Nevada	New Army Pass	81	Namkungia sp.	Mount Seolaksan NP, South Korea	Outgroup
Grylloblatta rothi	Karen Cave	44	Grylloblatta sp. North Christmas Tree Cave	North Christmas Tree Cave	83	Galloisiana odaesanensis	Mosangul Cave, South Korea	Outgroup
Grylloblatta sp. Central Sierra Nevada	Lake George	45	Grylloblatta oregonensis, Grylloblatta siskiyouensis	**Oregon Caves***	84	Galloisiana sp.	Mt. Tsurugi, Shikoku, Japan	Outgroup
Grylloblatta campodeiformis	Lake O'Hara	46	Grylloblatta sp. Mt. Spokane	Pebble Creek	86	Grylloblattella sp.	Pustag Mountain, Russia	Outgroup
Grylloblatta chandleri	Lassen Peak	47	Grylloblatta marmoreus	Planetary Dairy Cave, Marble Valley *	87	Grylloblattella sayanensis	Karasaibo River, Russia*	Outgroup
Grylloblatta gurneyi	Lava Beds NM*	48	Grylloblatta sp. Mt. Spokane	Porcupine Lake	88	Grylloblattella pravdini	Tevenek River, Russia*	Outgroup
						Grylloblattella sp.	Holzun Mountains, Russia	Outgroup

Legend for Figures S2-S7.

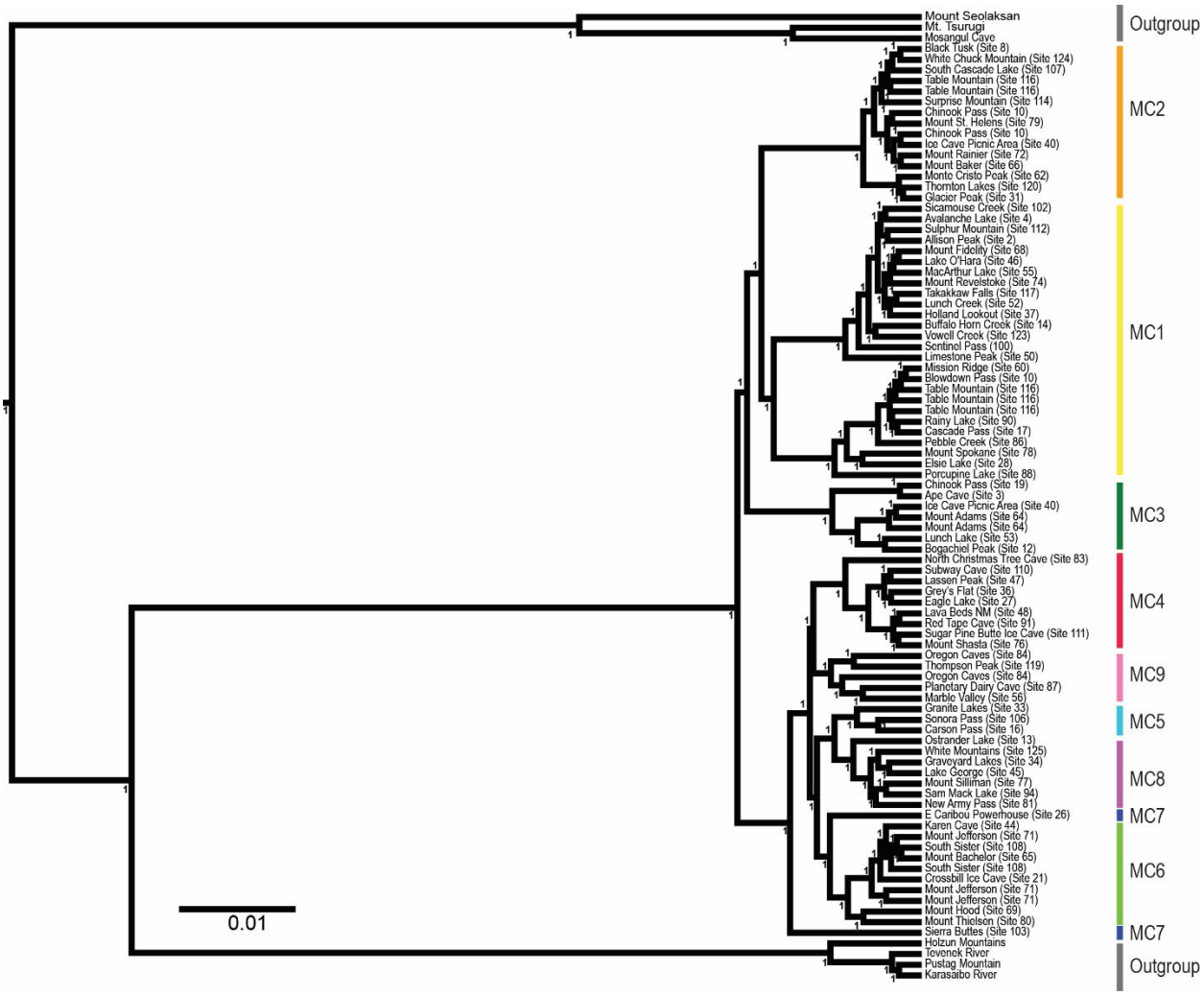

Figure S2. Bayesian phylogenetic relationships of *Grylloblatta* inferred from 322 concatenated nuclear genes (full dataset, with data exceeding 30% missing genes per individual in some cases). Numbers and names at the tips refer to sample sites and known taxon names, respectively. Well-supported genetic lineages are grouped into major clades (MC), represented as colored groups and referred to in the text. Nodal support values for the phylogeny are shown as Bayesian posterior probability estimates and branch lengths are in units of average substitutions per site. See detached legend for taxonomic assignment.

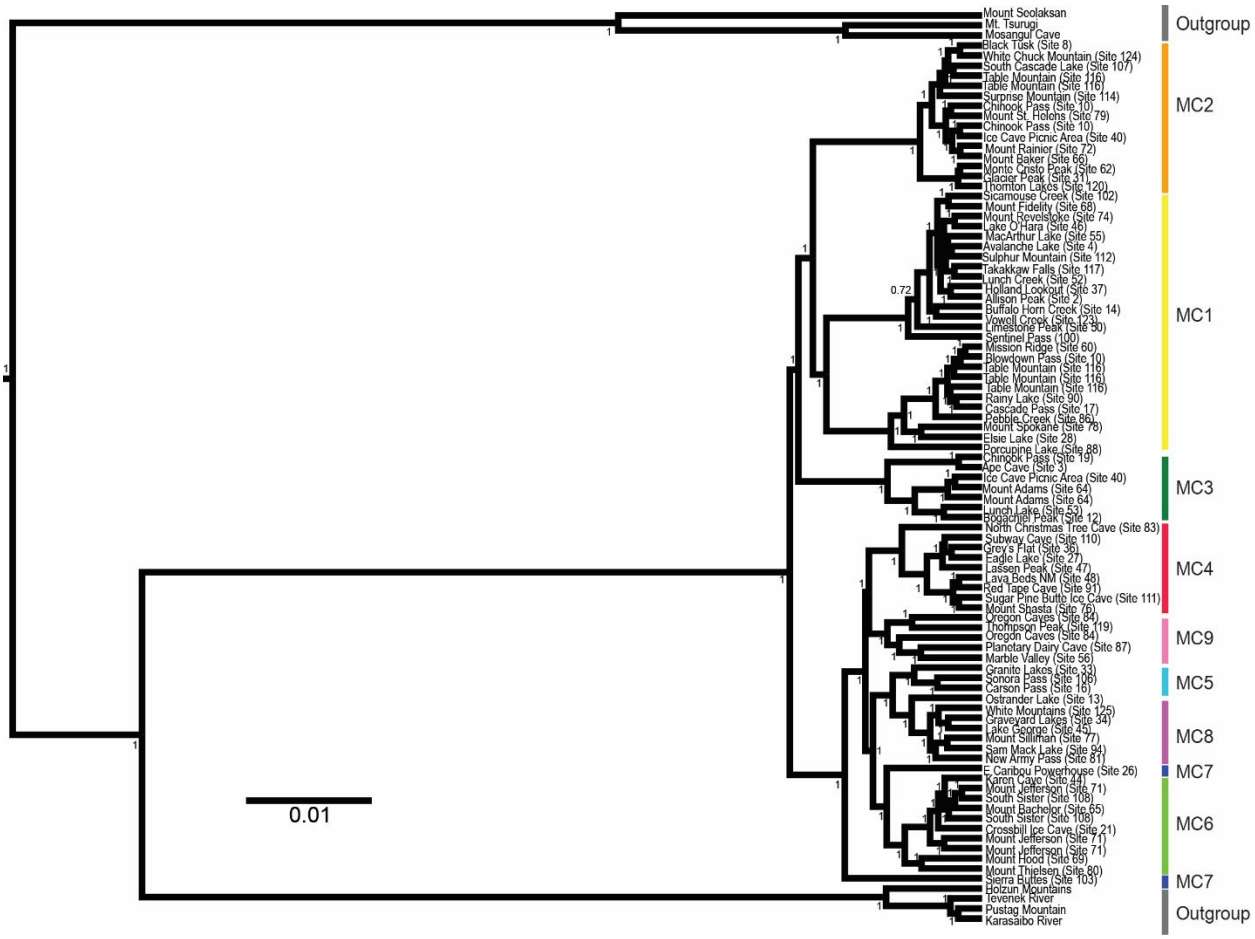

Figure S3. Bayesian phylogenetic relationships of *Grylloblatta* inferred from 266 concatenated nuclear genes (<30% missing genes per individual allowed). Numbers and names at the tips refer to sample sites and known taxon names, respectively. Well-supported genetic lineages are grouped into major clades (MC), represented as colored groups and referred to in the text. Nodal support values for the phylogeny are shown as Bayesian posterior probability estimates and branch lengths are in units of average substitutions per site. See detached legend for taxonomic assignment.

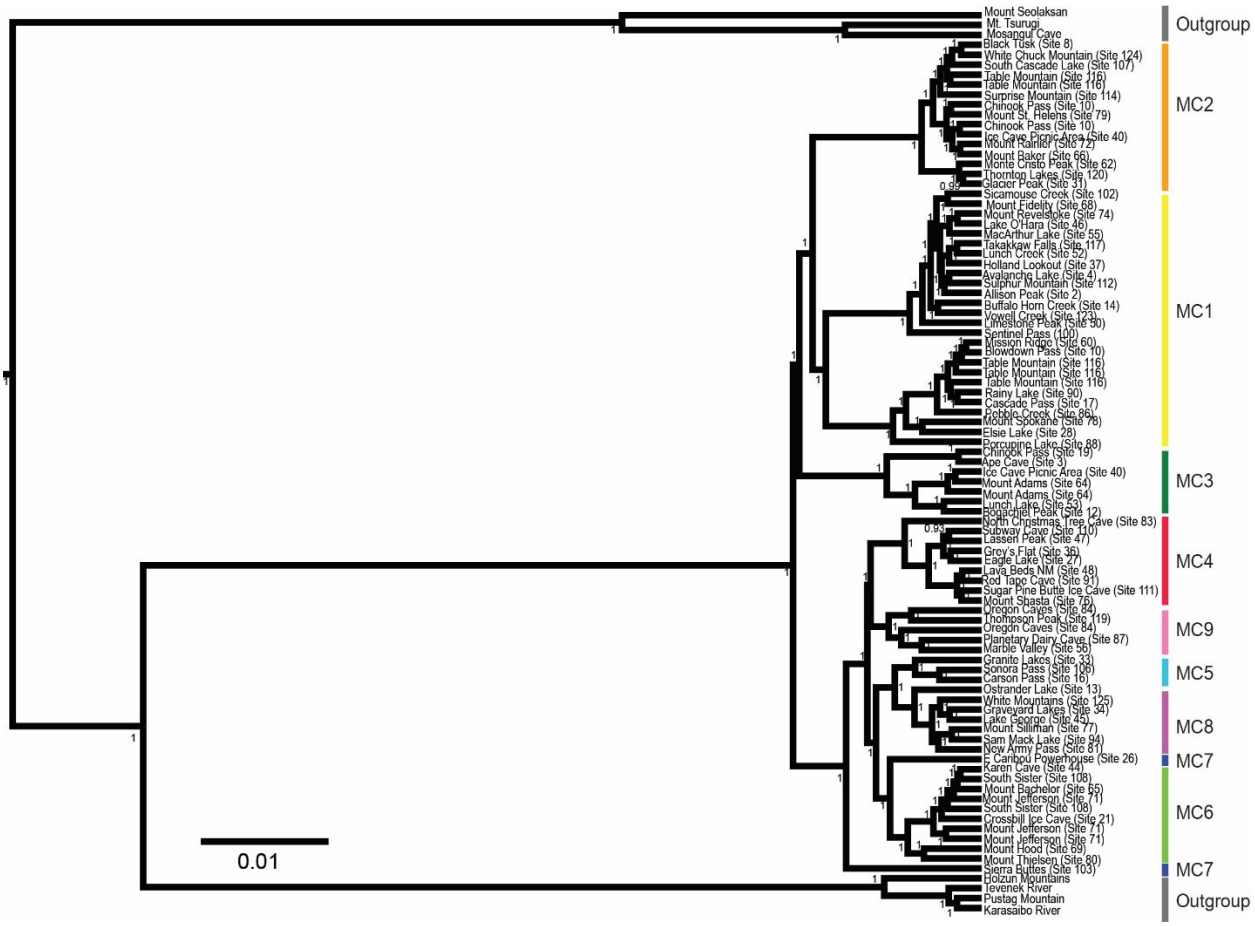

Figure S4. Bayesian phylogenetic relationships of *Grylloblatta* inferred from 191 concatenated nuclear genes (<20% missing genes per individual allowed). Numbers and names at the tips refer to sample sites and known taxon names, respectively. Well-supported genetic lineages are grouped into major clades (MC), represented as colored groups and referred to in the text. Nodal support values for the phylogeny are shown as Bayesian posterior probability estimates and branch lengths are in units of average substitutions per site. See detached legend for taxonomic assignment.

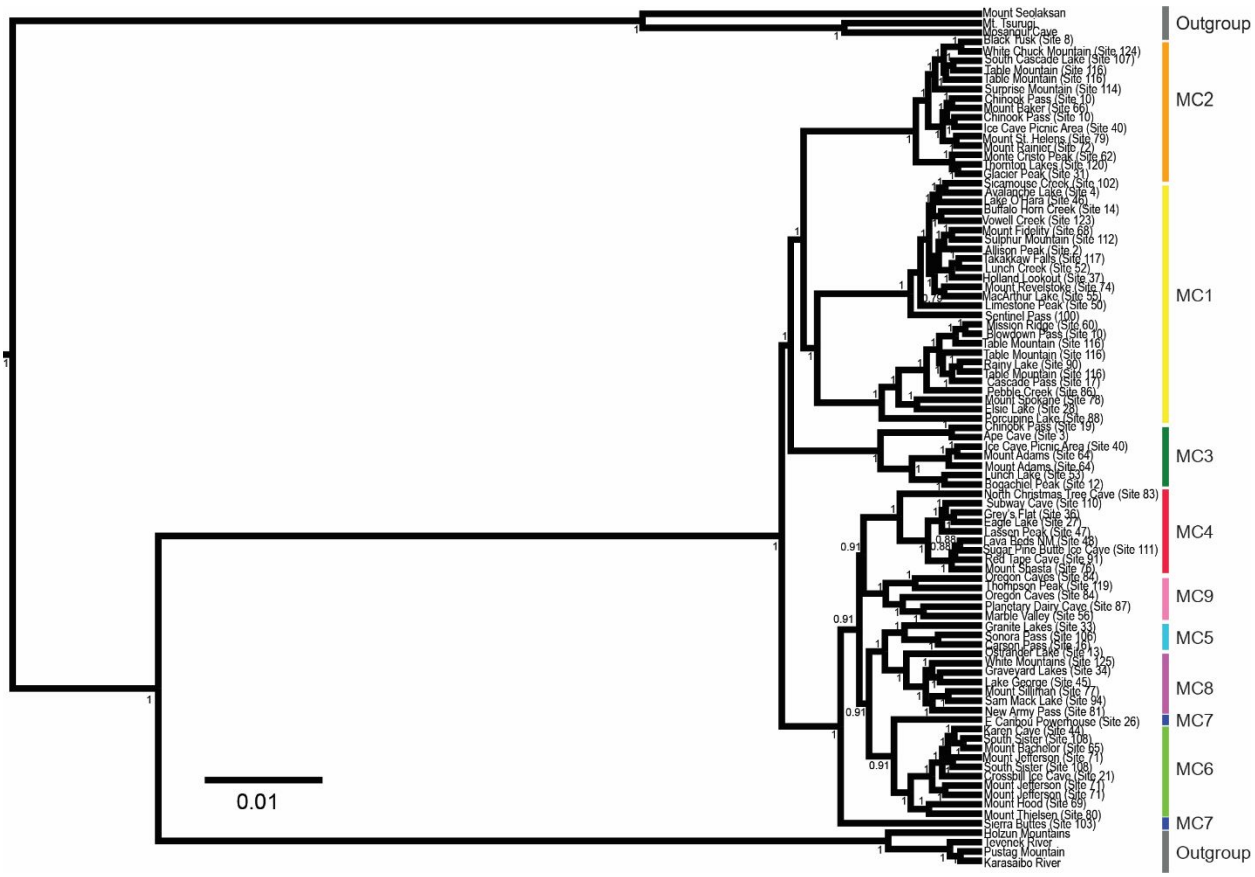

Figure S5. Bayesian phylogenetic relationships of *Grylloblatta* inferred from 77 concatenated nuclear genes (<10% missing genes per individual allowed). Numbers and names at the tips refer to sample sites and known taxon names, respectively. Well-supported genetic lineages are grouped into major clades (MC), represented as colored groups and referred to in the text. Nodal support values for the phylogeny are shown as Bayesian posterior probability estimates and branch lengths are in units of average substitutions per site. See detached legend for taxonomic assignment.

Figure S6. Bayesian phylogenetic relationships of *Grylloblatta* inferred from 26 concatenated nuclear genes (<less than 5% missing genes per individual allowed). Numbers and names at the tips refer to sample sites and known taxon names, respectively. Well-supported genetic lineages are grouped into major clades (MC), represented as colored groups and referred to in the text. Nodal support values for the phylogeny are shown as Bayesian posterior probability estimates and branch lengths are in units of average substitutions per site. See detached legend for taxonomic assignment.

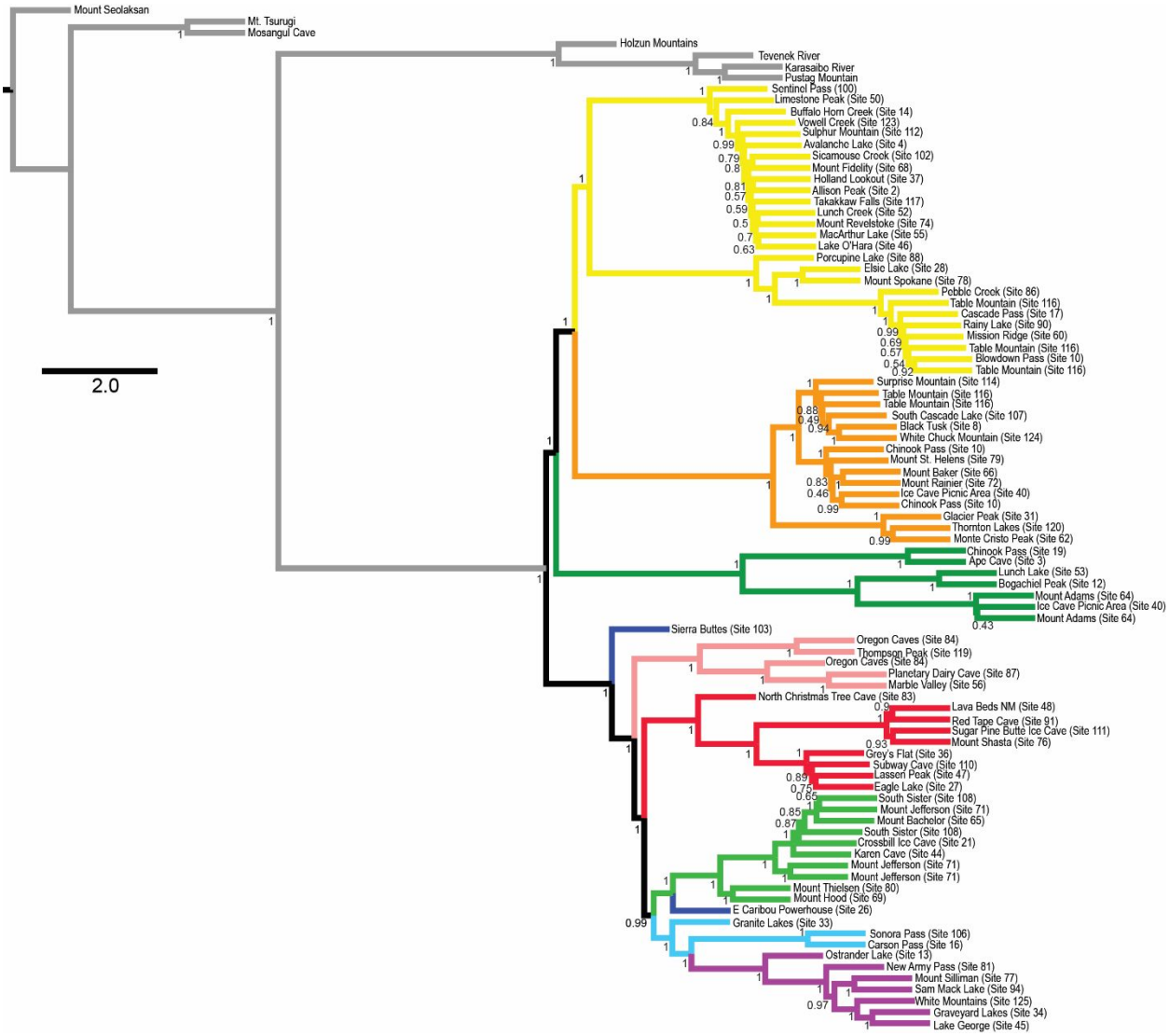

Figure S7. Multi-species coalescent tree inferred from ASTRAL. Numbers and names at the tips refer to sample sites and known taxon names, respectively. Well-supported genetic lineages are grouped into major clades, represented as colored groups and referred to in the text. Numbers at nodes indicate local posterior probability and branch lengths are represented in coalescent units. See detached legend for taxonomic assignment.

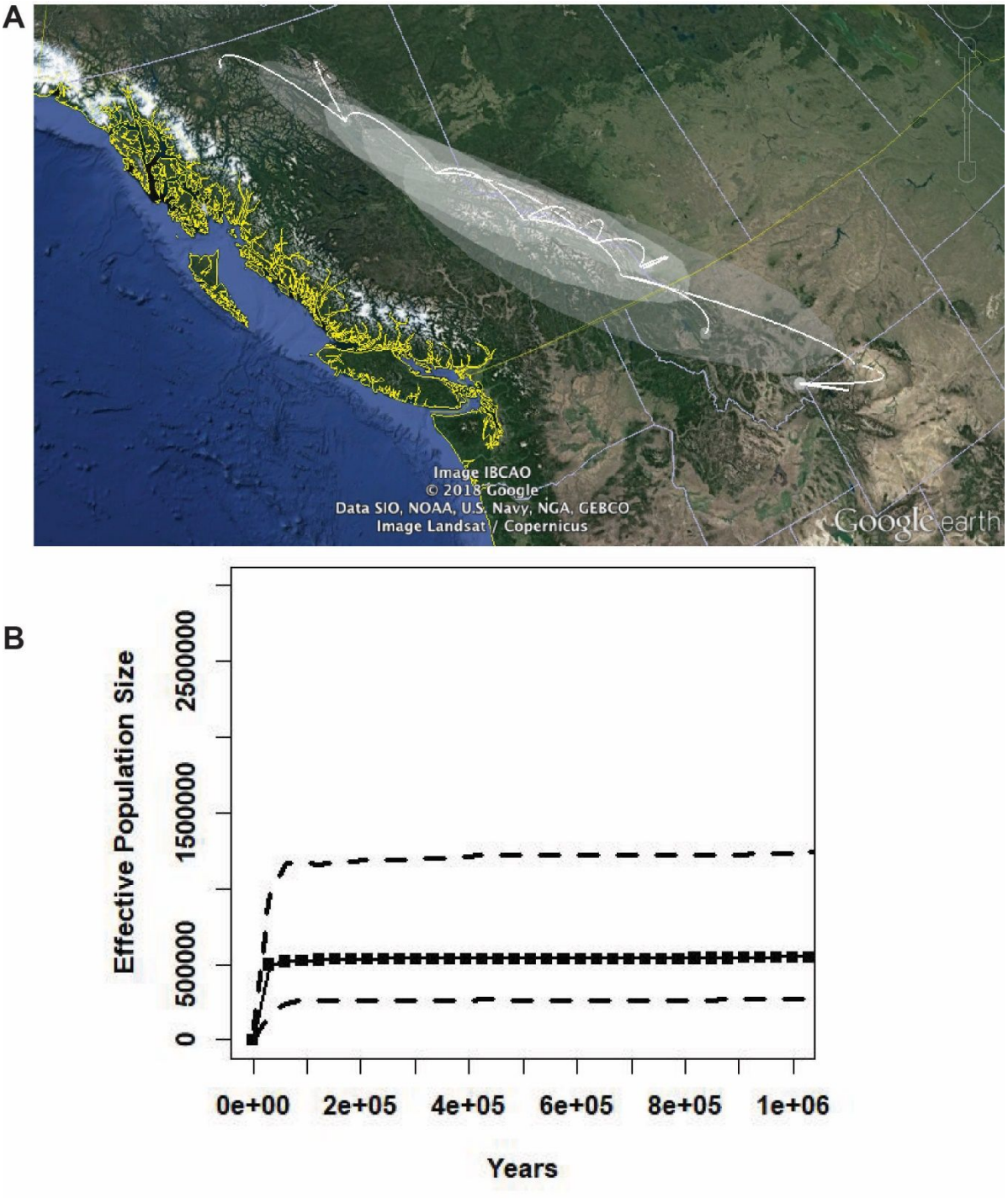

Figure S8. The estimated A) ancestral ranges and B) Bayesian skyline plots of effective population size change through time for the *Grylloblatta campodeiformis* clade, based on mitochondrial COII sequence data. Ancestral range estimates are shown as a white colored polygon with distinct haplotypes as lines coalescing into the ancestral population.

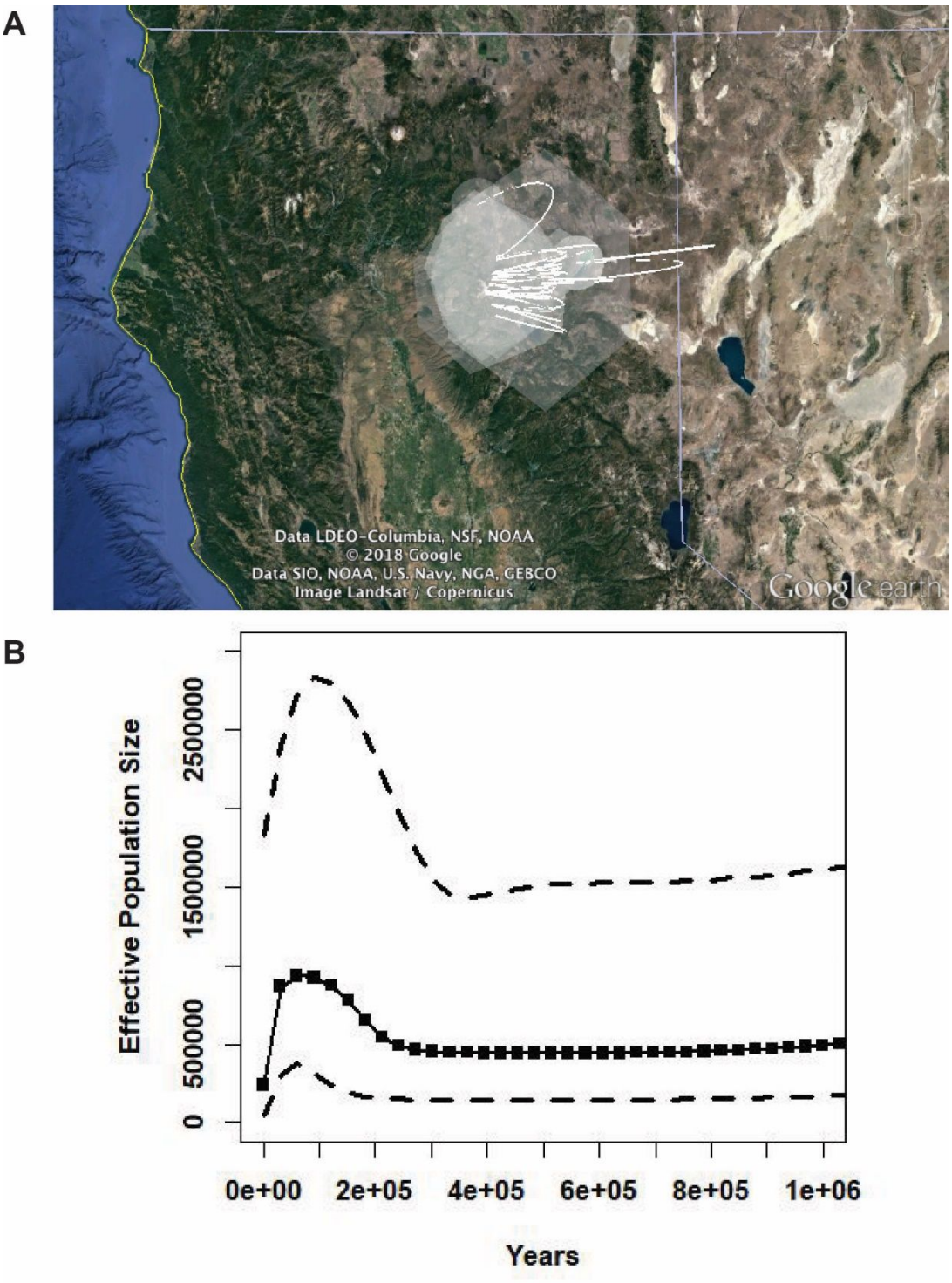

Figure S9. The estimated A) ancestral ranges and B) Bayesian skyline plots of effective population size change through time for the *Grylloblatta chandleri* clade, based on mitochondrial COII sequence data. Ancestral range estimates are shown as a white colored polygon with distinct haplotypes as lines coalescing into the ancestral population.

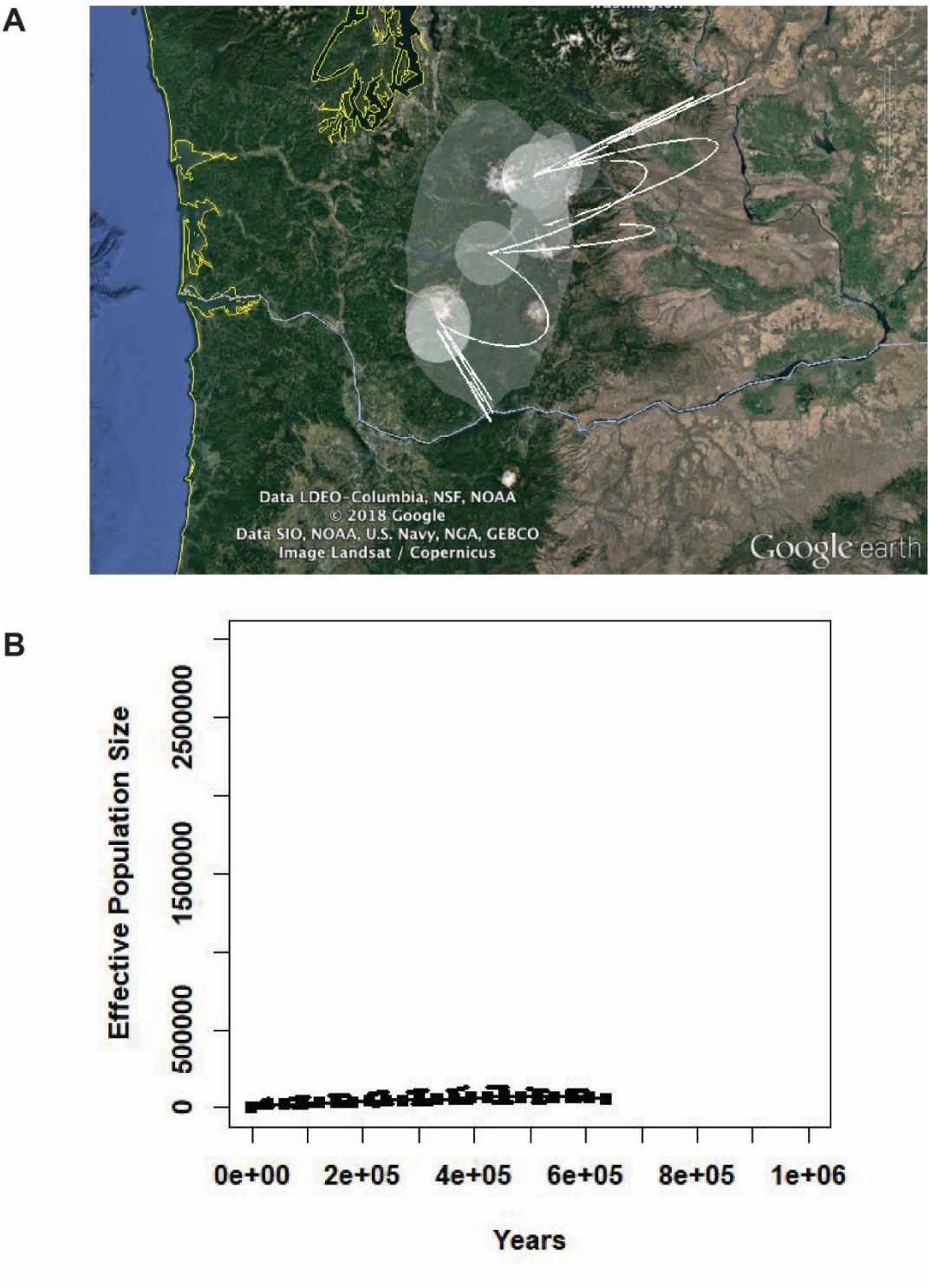

Figure S10. The estimated A) ancestral ranges and B) Bayesian skyline plots of effective population size change through time for the *Grylloblatta chirurgica* clade, based on mitochondrial COII sequence data. Ancestral range estimates are shown as a white colored polygon with distinct haplotypes as lines coalescing into the ancestral population.

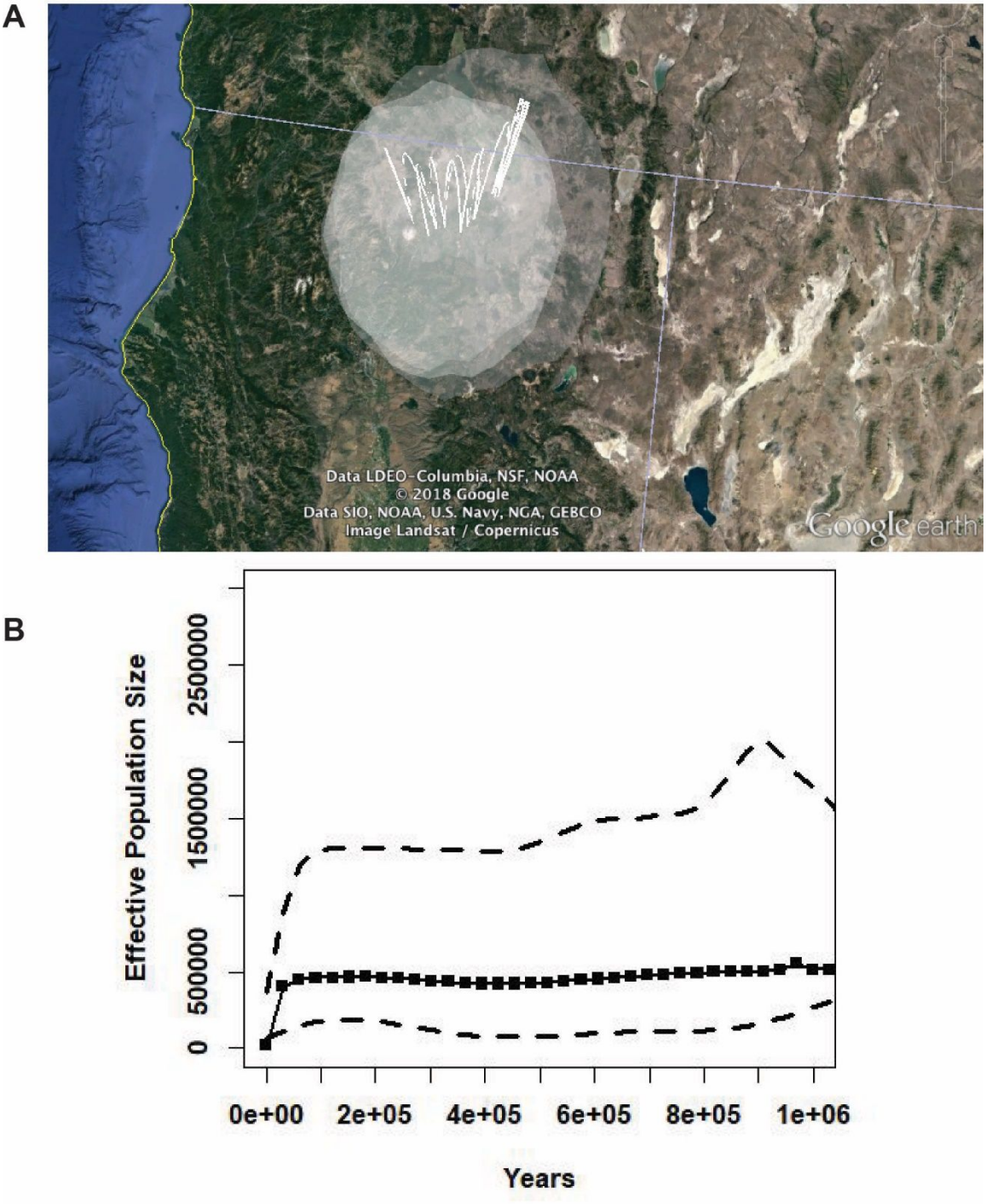

Figure S11. The estimated A) ancestral ranges and B) Bayesian skyline plots of effective population size change through time for the *Grylloblatta gurneyi* clade, based on mitochondrial COII sequence data. Ancestral range estimates are shown as a white colored polygon with distinct haplotypes as lines coalescing into the ancestral population.

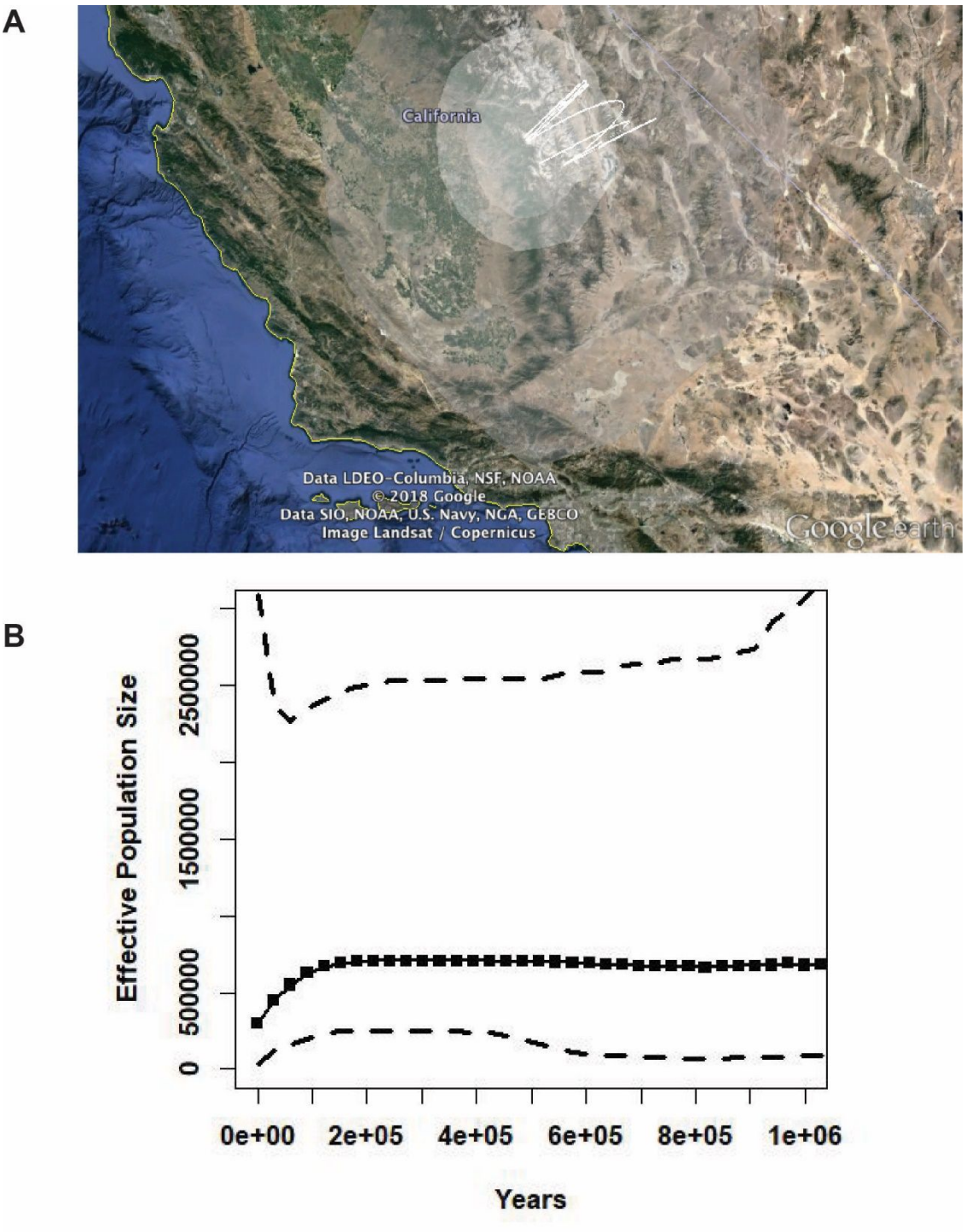

Figure S12. The estimated A) ancestral ranges and B) Bayesian skyline plots of effective population size change through time for the *Grylloblatta* sp. 'Southwest Sierra Nevada' clade, based on mitochondrial COII sequence data. Ancestral range estimates are shown as a white colored polygon with distinct haplotypes as lines coalescing into the ancestral population.

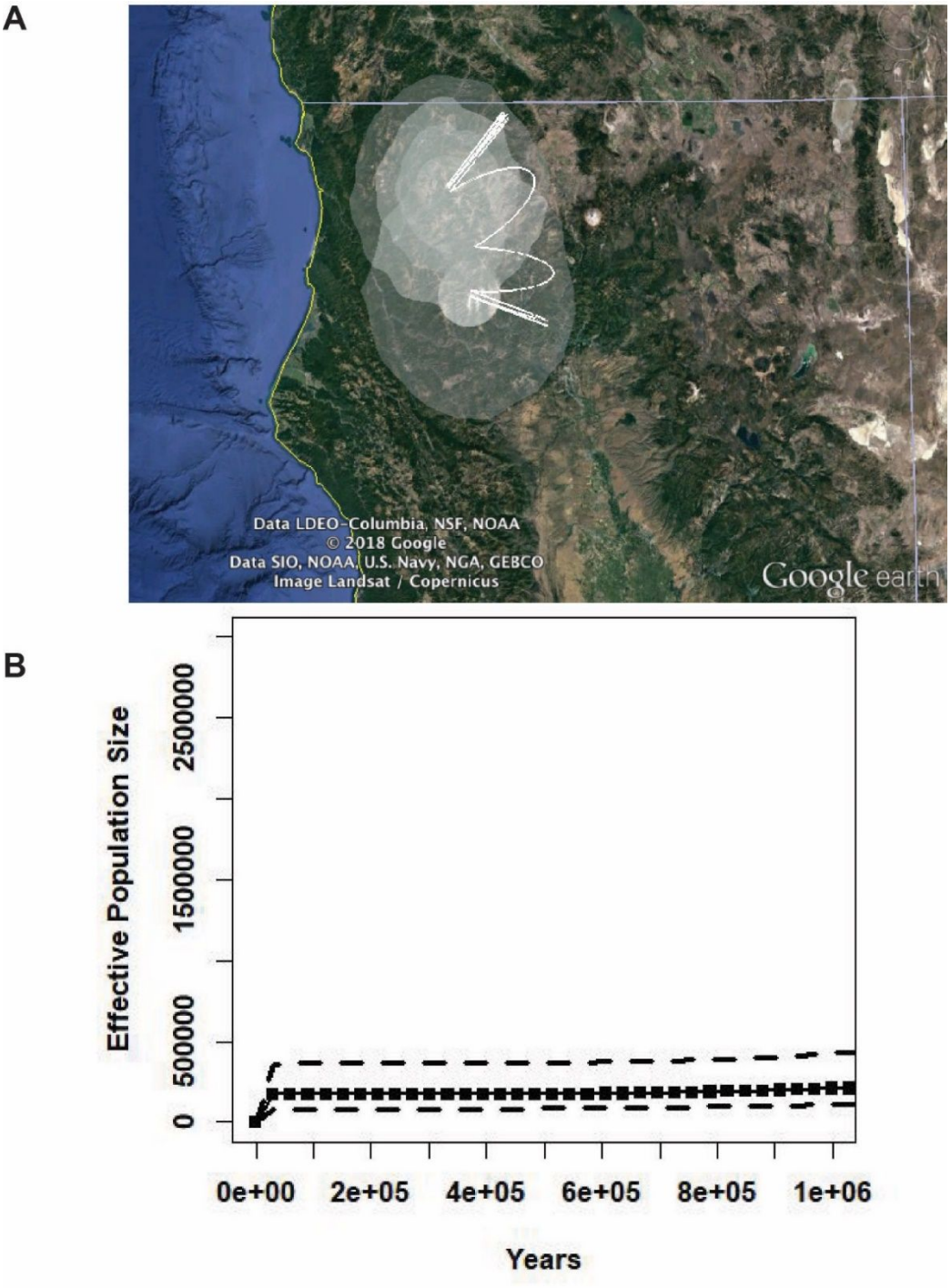

Figure S13. The estimated A) ancestral ranges and B) Bayesian skyline plots of effective population size change through time for the *Grylloblatta marmoreus* clade, based on mitochondrial COII sequence data. Ancestral range estimates are shown as a white colored polygon with distinct haplotypes as lines coalescing into the ancestral population.

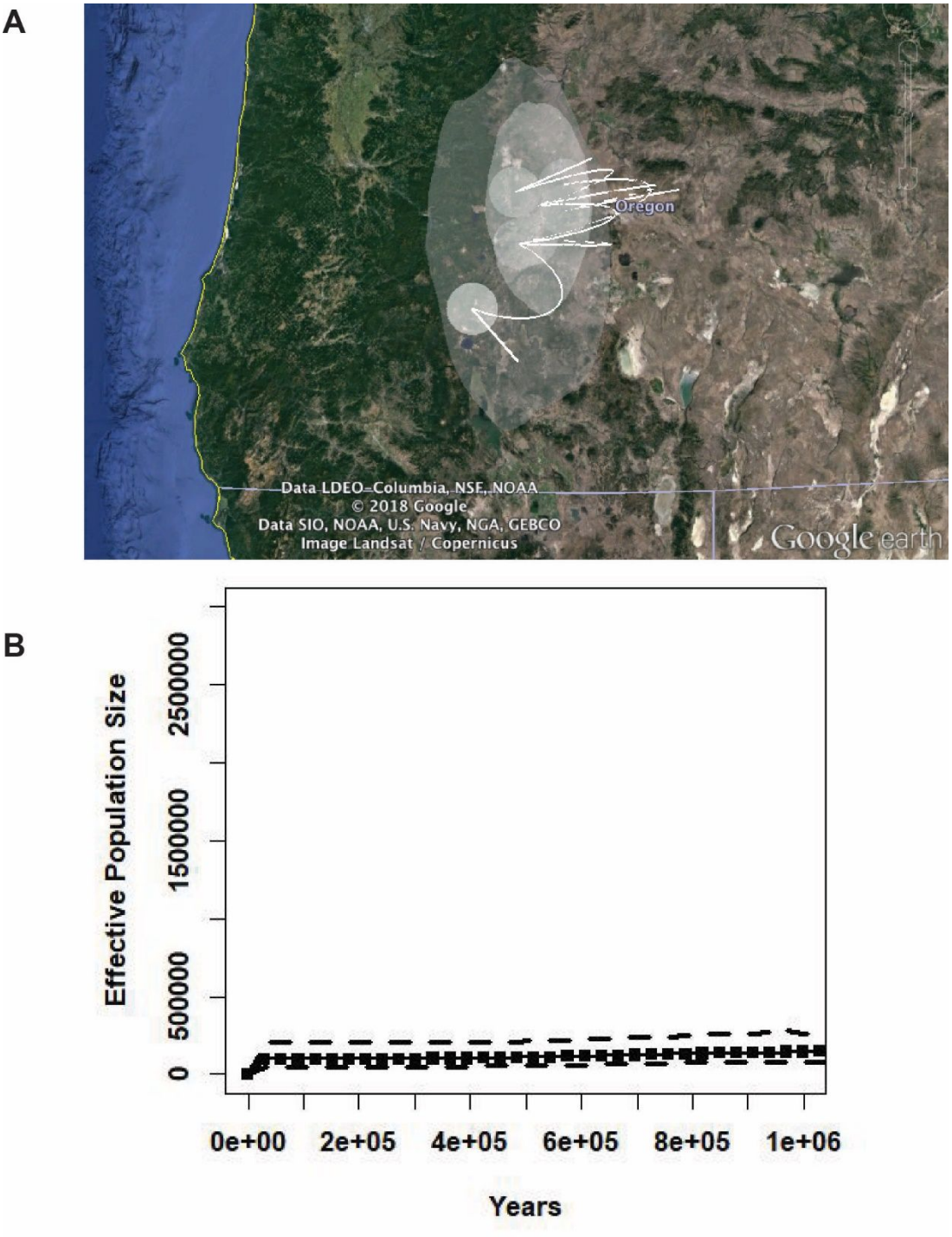

Figure S14. The estimated A) ancestral ranges and B) Bayesian skyline plots of effective population size change through time for the *Grylloblatta newberryensis* clade, based on mitochondrial COII sequence data. Ancestral range estimates are shown as a white colored polygon with distinct haplotypes as lines coalescing into the ancestral population.

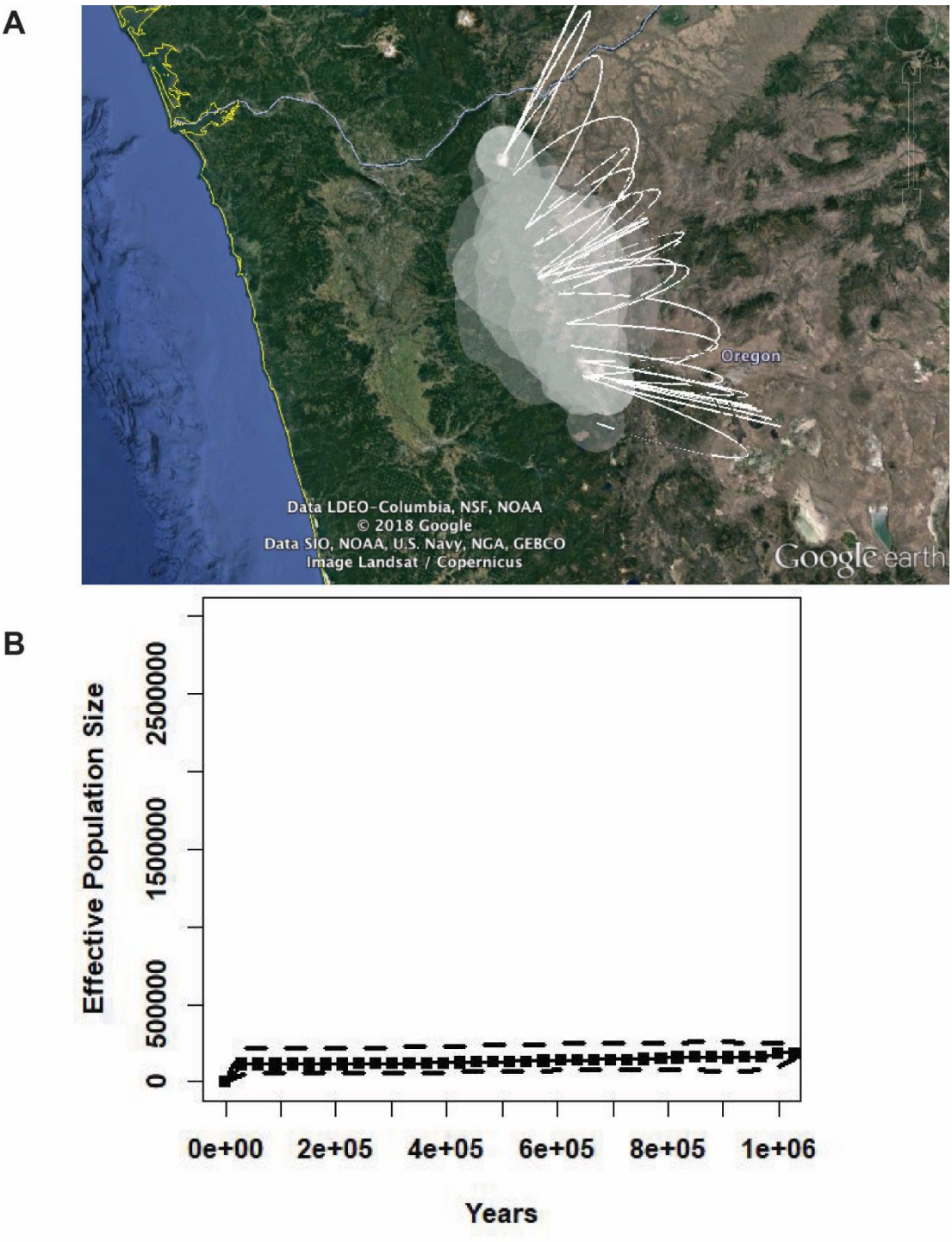

Figure S15. The estimated A) ancestral ranges and B) Bayesian skyline plots of effective population size change through time for the *Grylloblatta rothi* clade, based on mitochondrial COII sequence data. Ancestral range estimates are shown as a white colored polygon with distinct haplotypes as lines coalescing into the ancestral population.

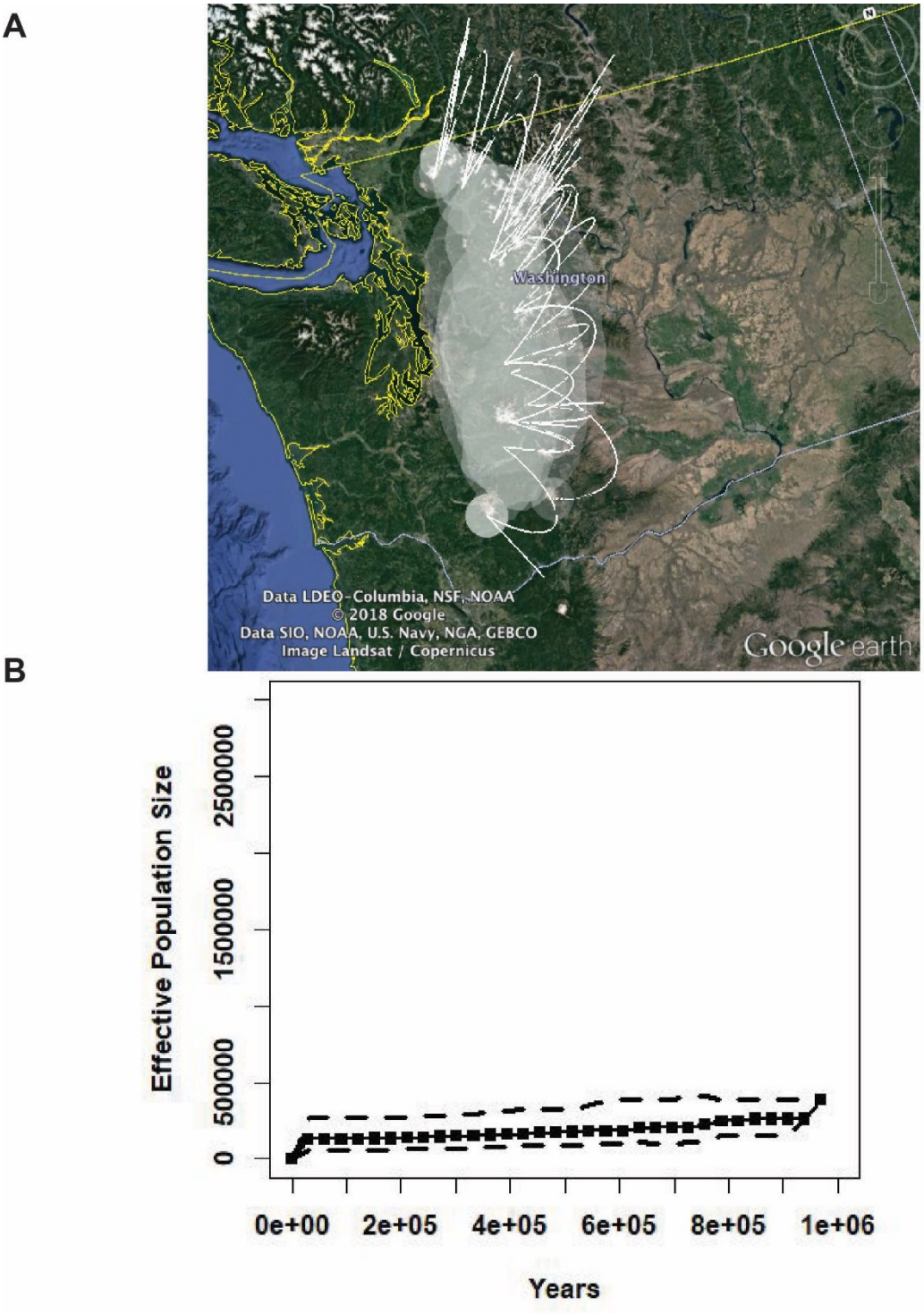

Figure S16. The estimated A) ancestral ranges and B) Bayesian skyline plots of effective population size change through time for the *Grylloblatta* sp. 'North Cascades' clade, based on mitochondrial COII sequence data. Ancestral range estimates are shown as a white colored polygon with distinct haplotypes as lines coalescing into the ancestral population.

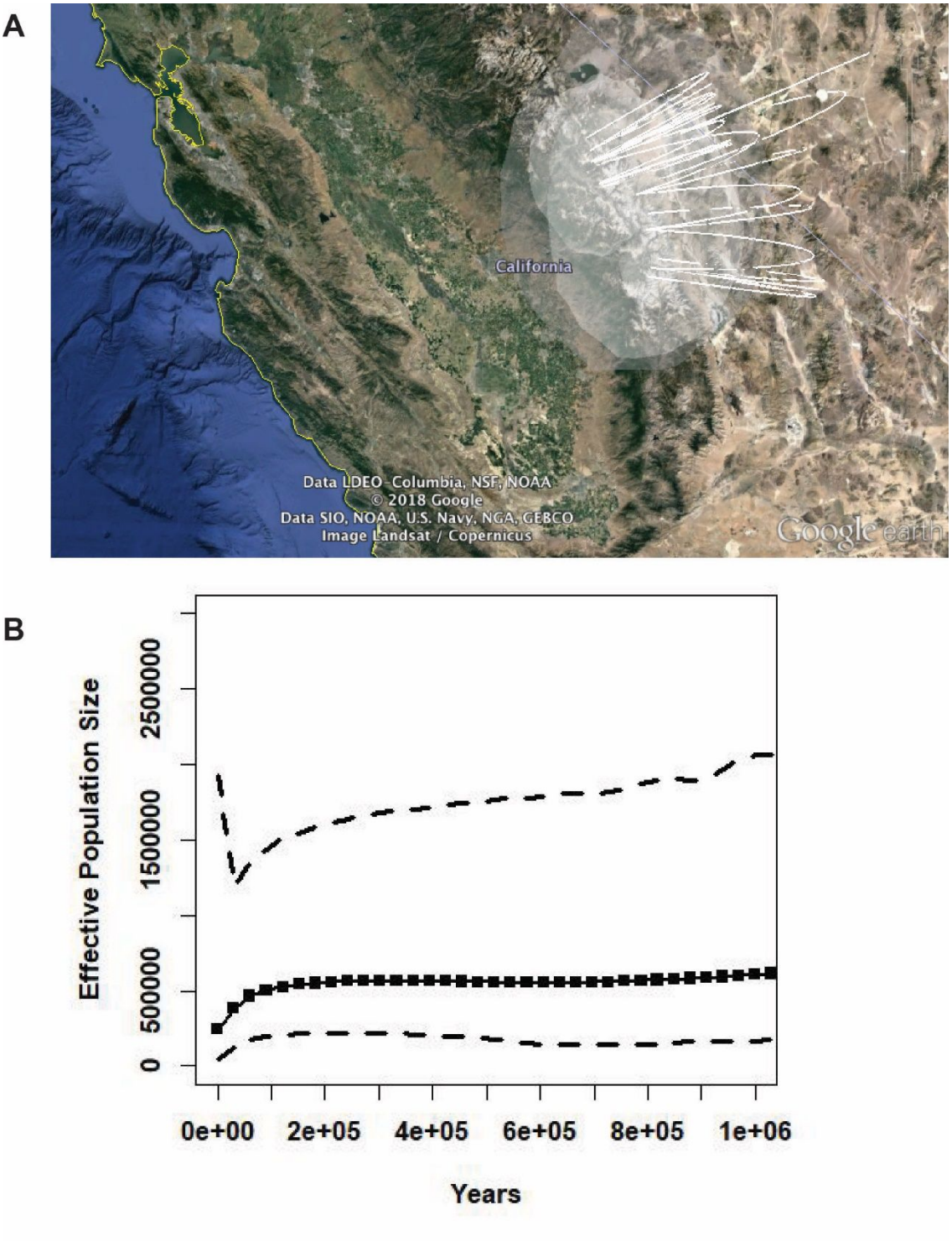

Figure S17. The estimated A) ancestral ranges and B) Bayesian skyline plots of effective population size change through time for the *Grylloblatta* sp. *'Central Sierra Nevada'* clade, based on mitochondrial COII sequence data. Ancestral range estimates are shown as a white colored polygon with distinct haplotypes as lines coalescing into the ancestral population.

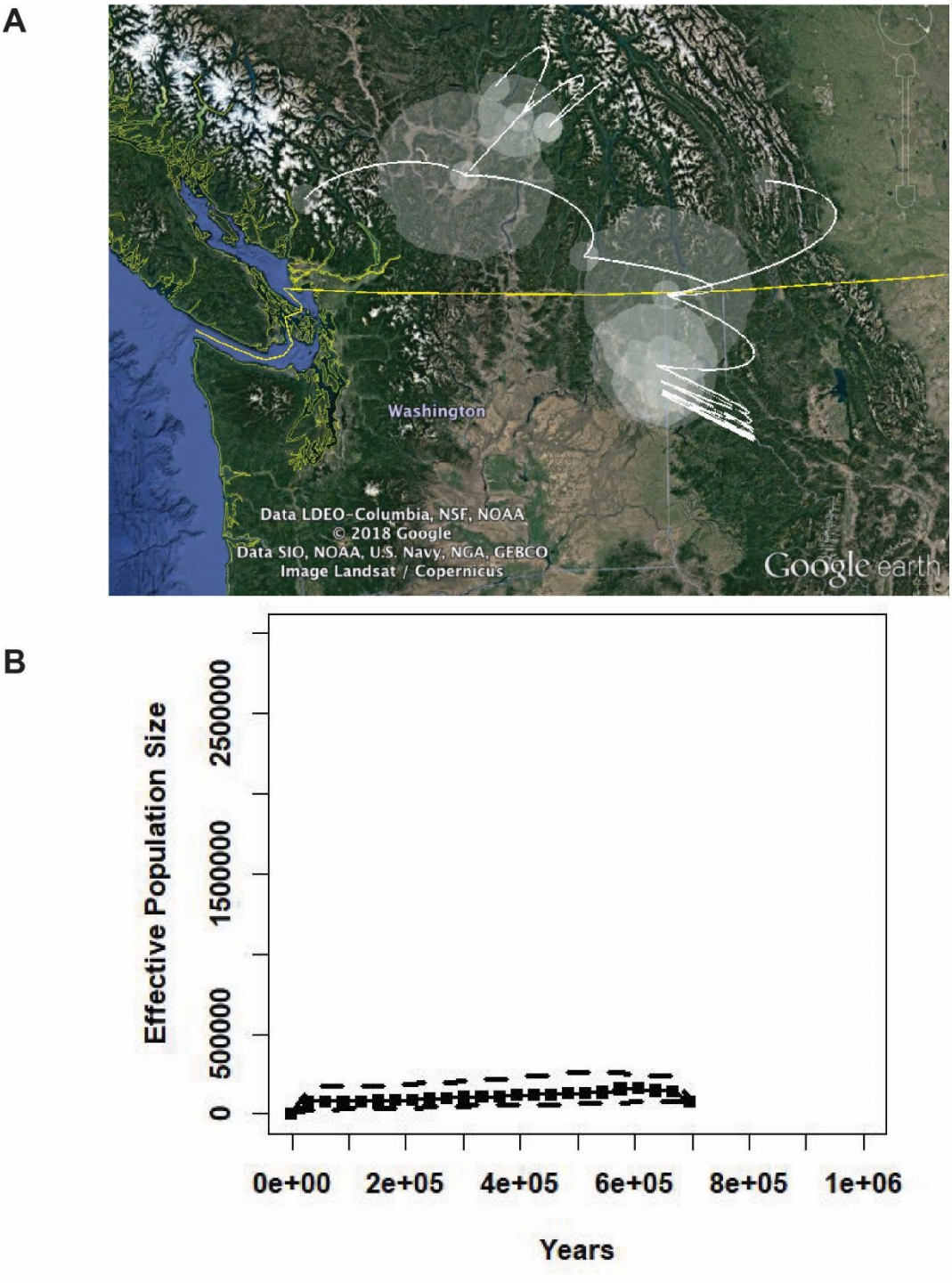

Figure S18. The estimated A) ancestral ranges and B) Bayesian skyline plots of effective population size change through time for the *Grylloblatta* sp. 'Mt. Spokane' clade, based on mitochondrial COII sequence data. Ancestral range estimates are shown as a white colored polygon with distinct haplotypes as lines coalescing into the ancestral population.

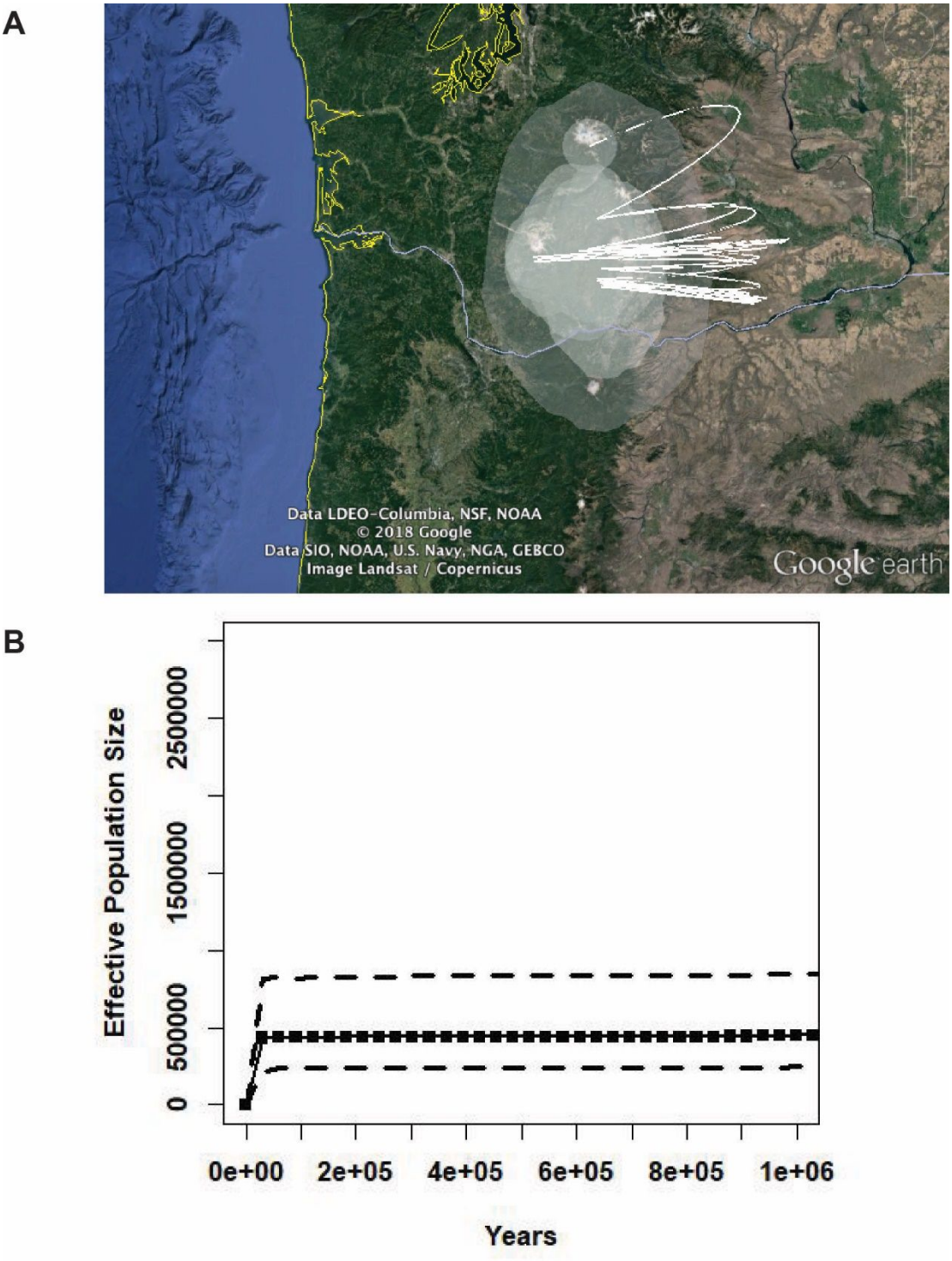

Figure S19. The estimated A) ancestral ranges and B) Bayesian skyline plots of effective population size change through time for the *Grylloblatta* sp. 'Trout Lake' clade, based on mitochondrial COII sequence data. Ancestral range estimates are shown as a white colored polygon with distinct haplotypes as lines coalescing into the ancestral population.

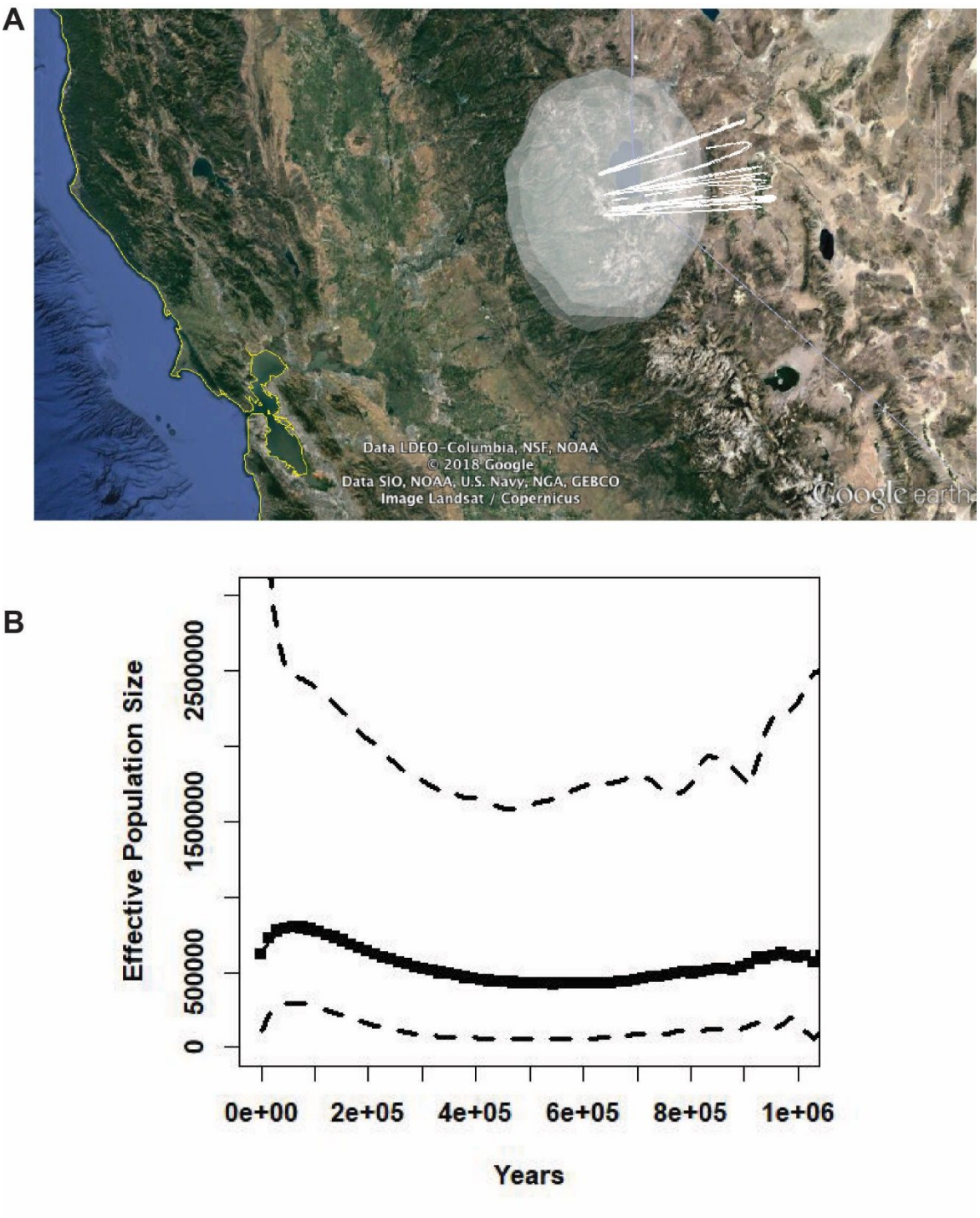

Figure S20. The estimated A) ancestral ranges and B) Bayesian skyline plots of effective population size change through time for the *Grylloblatta washoa* clade, based on mitochondrial COII sequence data. Ancestral range estimates are shown as a white colored polygon with distinct haplotypes as lines coalescing into the ancestral population.