MOLECULAR ECOLOGY

Molecular characterisation of trophic ecology within an island radiation of insect herbivores (Curculionidae: Entiminae: Cratopus).

Journal:	<i>Molecular Ecology</i>
Manuscript ID:	MEC-13-0574
Manuscript Type:	Original Article
Date Submitted by the Author:	28-May-2013
Complete List of Authors:	Kitson, James; University of East Anglia, School of Biological Sciences Warren, Ben; UMR C53 PVBMT, Florens, F. B. Vincent; University of Mauritius, Faculty of Science Baider, Claudia; The Mauritius Herbarium, Strasberg, Dominique; UMR PVBMT, Université de La Réunion, Emerson, Brent; IPNA-CSIC, Ecology and Evolution; University of East Anglia, School of Biological Sciences
Keywords:	Adaptation, Insects, Natural Selection and Contemporary Evolution, Phylogeography, Population Genetics - Empirical, Species Interactions

1 **Molecular characterisation of trophic ecology within an island radiation of insect**
2 **herbivores (Curculionidae: Entiminae: *Cratopus*).**

3

4 James J. N. Kitson⁽¹⁾, Ben H. Warren⁽²⁾, F. B. Vincent Florens⁽³⁾, Claudia Baider⁽⁴⁾, Dominique
5 Strasberg⁽⁵⁾ and Brent C. Emerson^(1,6)

6

7 ¹ School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich, NR4
8 7TJ, UK.

9 ² UMR PVBMT, Université de La Réunion – CIRAD, Ligne Paradis, 97410 Saint Pierre, La
10 Réunion, France.

11 ³ Faculty of Science, University of Mauritius, Réduit, Mauritius.

12 ⁴ The Mauritius Herbarium, Mauritius Sugar Industry Research Institute, Réduit, Mauritius.

13 ⁵ UMR PVBMT, Université de La Réunion - CIRAD, 97715 Saint-Denis, La Réunion, France.

14 ⁶ Present address: Island Ecology and Evolution Research Group, IPNA-CSIC , C/Astroffsico
15 Francisco Sánchez 3, La Laguna, Tenerife, Canary Islands, 38206, Spain

16

17 Corresponding author: Brent Emerson. E-mail: bemerson@ipna.csic.es

18

19 Key words: insect, plant, trophic ecology, chloroplast, weevil

20

21 Running title: Molecular analysis of insect herbivore diets

22

23

24

Abstract

The phytophagous beetle family Curculionidae is the most species-rich insect family known, with much of this diversity having been attributed to both co-evolution with food plants and host-shifts at key points within the early evolutionary history of the group. Less well understood is the extent to which patterns of host use vary within or among related species, largely because of the technical difficulties associated with quantifying this. Here we develop a recently characterised molecular approach to quantify diet within and between two closely related species of weevil occurring primarily within dry forests on the island of Mauritius. Our aim is to quantify dietary variation across populations and assess adaptive and non-adaptive explanations for this, and to characterise the nature of a trophic shift within an ecologically distinct population within one of the species. We find that our study species are polyphagous, consuming a much wider range of plants than would be suggested by the literature. Our data suggest that local diet variation is largely explained by food availability, and locally specialist populations consume food plants that are not phylogenetically novel, but do appear to represent a novel preference. Our results demonstrate the power of molecular methods to unambiguously quantify dietary variation across populations of insect herbivores, providing a valuable approach to understanding trophic interactions within and among local plant and insect herbivore communities.

45 **Introduction**

46

47 Weevils are the most successful of all insect herbivore lineages, with approximately 62,000
 48 described species and an estimated 220,000 species in total (Oberprieler *et al.* 2007).
 49 Interactions between weevils and angiosperms have been hypothesised to be a significant
 50 driver of the great diversity we see today (Farrell 1998; Oberprieler *et al.* 2007) under a
 51 scenario where much of the family-level diversity evolved on gymnosperms and monocots
 52 before host shifting to, and radiating across, angiosperms (McKenna *et al.* 2009). However, at
 53 more recent species-level evolutionary timescales, patterns of host use and changes in host use
 54 among and within species through time are less understood. Progress in this area has largely
 55 been made through the investigation of species that inhabit tissue of a host plant at some stage
 56 of their life cycle (e.g. Barat *et al.* 2008; Hernandez-Vera *et al.* 2010; Jordal *et al.* 2004;
 57 Podlussany *et al.* 2001), where a reliable host/herbivore relationship can be inferred. Trophic
 58 relationships are much more complicated to describe for weevil species where adults feed on
 59 plant foliage and larvae do not develop inside host tissue. This is frequently the case for groups
 60 such as the entiminine weevils (Curculionidae: Entiminae) where trophic ecology is particularly
 61 complicated to quantify for three reasons. First, entiminine weevils generally feed at night so
 62 direct feeding observations are rare (Morris 1997). Second, multiple entiminine species may
 63 feed upon the same plant species (Williams 2000; Williams & Cox 2003) so directly attributing
 64 feeding damage to one species over another is difficult. Finally, weevils move through the
 65 environment (for example while searching for mates or looking for food plants) so there is no
 66 specific reason to suspect that the particular plant upon which a weevil is collected is indicative
 67 of what it feeds upon.

68

69 Host choice tests under laboratory conditions provide one way to quantify the trophic ecology
 70 of insect herbivores (e.g. Barone 1998; Novotny *et al.* 2002; Novotny *et al.* 2006), but such
 71 experiments may be costly, labour intensive or difficult to achieve in remote field conditions. In
 72 addition to these disadvantages, there is the obvious complication that feeding associations
 73 inferred under laboratory conditions may not be an accurate reflection of feeding associations
 74 under natural conditions. Studies such as that of Otte (1976) have attempted to circumvent
 75 these shortcomings by microscopically examining digestive system contents, but this approach
 76 is also extremely labour intensive. Recent advances in molecular ecology have provided a new
 77 set of tools for the direct quantification of diet that have been applied to a wide range of systems
 78 (e.g. Deagle *et al.* 2009; Hereward & Walter 2012; Staudacher *et al.* 2011) including trophic
 79 interactions between plants species and coleopteran herbivores (e.g. García-Robledo *et al.* 2013;
 80 Jurado-Rivera *et al.* 2009; Pinzón-Navarro *et al.* 2010).

81

82 Jurado-Rivera *et al.* (2009) first demonstrated that field caught chrysomelid beetles may contain
 83 sufficient plant tissue in their digestive system to permit the amplification of chloroplast DNA
 84 for the quantification of plant species consumed across a range of diverse chrysomelid species.
 85 Here we extend the approach of Jurado-Rivera *et al.* (2009) for an analysis of trophic ecology
 86 within a radiation of entiminine weevils across the Mascarene Islands, part of the southwest
 87 Indian Ocean area biodiversity hotspot (Myers *et al.* 2000). The genus *Cratopus* is represented

by approximately 100 species distributed across the islands of the Indian Ocean, with the great majority of species being restricted to the Mascarene Islands (La Réunion, Mauritius and Rodrigues), where approximately 80 species are found. The genus is particularly unusual for an oceanic island radiation of beetles, in that all but a few species are flighted. All other entiminine radiations studied on oceanic islands to date involve flightless lineages. These include *Syzygops* on the Mascarene Islands (Williams 2000), *Laparocerus* within Macaronesia (Machado 2006, 2007a, b, 2008; Machado *et al.* 2008), *Rhyncogonus* within Polynesia (Gillespie *et al.* 2008), *Ectemnorhinus* of the Prince Edward Islands (Grobler *et al.* 2006) and *Galapaganus* from the Galapagos Islands (Sequeira *et al.* 2008). While geographic isolation within islands has apparently acted to facilitate speciation within island invertebrate radiations (e.g., Holland & Hadfield 2002), this would appear to hold less explanatory power for a flighted group such as *Cratopus*.

The available information on the trophic ecology of *Cratopus*, although limited, indicates that they feed on a broad range of plant species, with Mauritian species ranging from generalists to specialists (Williams & Cox 2003). This suggests the potential for *Cratopus* species to have diversified with respect to plant feeding relationships, but given the complications described above for quantifying trophic ecology within Entiminine species, details of specific plant feeding relationships are lacking. In this study we focus attention on *Cratopus murinus*, and its close relative *C. ovalis*, both endemic to Mauritius. Both species are most commonly found in the isolated dry or semi-dry forests and scrub that cover the mountains of the Mauritian central plateau (Williams & Cox 2003), with *C. murinus* also occurring on several small coralline islets off the east coast of Mauritius. Within these habitats, a range of plant species exhibit foliage damage typical of Entiminine feeding and both species can be collected from a range of plant species, including species not exhibiting feeding damage to foliage. In contrast, the coralline islet populations of *C. murinus* are found only on one plant species, *Scaevola taccada*, a species that does not naturally occur within the dry forest habitat of the main island. *Cratopus murinus* populations from the islets also have atrophied remnants of wings, suggesting these populations to be derived from mainland source populations.

We have sampled *C. murinus* and *C. ovalis* from across Mauritius and coralline islets to apply a molecular approach to quantify (1) geographic structure of genetic variation within species and (2) trophic ecology. We adopt a multi-gene approach to quantify population structure and phylogeographic history within both species, with a specific focus on identifying the evolutionary and temporal origin of the coralline islet populations of *C. murinus*. We then assess the efficiency of the cpDNA trnL UAA gene (Jurado-Rivera *et al.* 2009) with an alternative set of primers to those used previously (Jurado-Rivera *et al.* 2009; Pinzón-Navarro *et al.* 2010) to address the following aims: (1) determine the plant species consumed by *C. ovalis* and *C. murinus*; (2) quantify trophic variation among dry forest populations of *C. murinus* and test between explanations of adaptive shift, or neutral response to resource availability; (3) determine whether *S. taccada* represents the acquisition of a phylogenetically novel food plant (i.e. is *S. taccada* closely or distantly related to one or more plant species consumed within mainland populations); (4) distinguish between a specialist-to-specialist transition, or a

generalist-to-specialist transition, for the trophic shift to *S. taccada* with the colonisation of coralline islets.

Materials and Methods

Beetle sampling

Beetles were collected by foliage beating on Mauritius in April 2009. Samples were placed directly in 99% ethanol solution with a single individual per tube to prevent cross-contamination of plant DNA present. Several individuals were kept alive until the same evening for photographing, before being placed individually in 99% ethanol. Sampling sites were recorded on a handheld GPS unit (Garmin GPS 60, Garmin Ltd). A total of 201 beetles were collected across 9 sites (136 *C. murinus* and 65 *C. ovalis*). See Table S1 for descriptions and GPS co-ordinates of sampling sites.

Plant sampling

Plant sampling was carried out by CB and FBVF between 5th September 2010 and 1st October 2010. For every GPS co-ordinate where beetles were sampled in Corps de Garde and Trois Mamelles (the two sites with sufficient samples of *C. murinus* for more detailed analysis), a 5m-radius circular plot was marked and a complete survey of vascular plants was undertaken. A total of twenty-five circular plots were surveyed for vascular plants, 13 in Corps de Garde and 12 in Trois Mamelles. To construct a DNA sequence reference database for the identification of plant species consumed by beetles, leaf samples were taken from single individuals of each surveyed plant species and preserved in silica gel for eventual chloroplast DNA amplification and sequencing, using the same primers used for the beetle dietary analysis (see below). To control against sample contamination within the diet analyses, sampled plant leaf material was maintained in Mauritius until the completion of trnL sequencing from beetle specimens, after which it was transported to Norwich for DNA extraction and sequencing.

DNA extraction, PCR amplification and sequencing

DNA was extracted from the head and prothorax of beetles using the DNeasy 96 well Blood and Tissue Extraction kit (QIAGEN, West Sussex, UK) with the digestion buffer volumes amended for large specimens as recommended by the manufacturer. Plant leaf samples were ground by hand and extracted using the DNeasy Plant kit (QIAGEN, West Sussex, UK) as per the manufacturer's instructions. Plant DNA extraction was performed in a separate laboratory to beetle DNA extractions. To undertake an analysis of population structure within beetle species we sequenced the mitochondrial gene Cytochrome Oxidase II (COII), and two nuclear genes: arginine kinase (ArgK) and internal transcribed spacer 2 (ITS2). Beetle feeding records were

obtained directly from beetle DNA extractions by amplifying the plant plastid gene transfer RNA Leucine UAA (trnL). Primers for each gene are listed in Table S2. Primers COIICraF and COIICraR were designed by aligning Curcujiform COII sequences from GenBank and selecting conserved regions as close to the COII gene as possible. COII amplification conditions were; 0.5 mM of each primer, 5 mM MgCl₂ and a thermal profile of: 95°C for 60s, 58°C for 60s and 72°C for 90s, 40 cycles. For ArgK, the primers ArgKforB2 and ArgKrevB1 (attributed to Danforth *et al.* (2005) in McKenna *et al.* (2009)) were used with a thermal profile of: 95°C for 60s, 50°C for 60s and 72°C for 120s, 37 cycles. ITS2 was amplified using the primers CAS28sB1d (Ji *et al.* 2003) and M13REV-CAS5 (CAS5p8sFt in Ji *et al.* (2003) but modified to have a M13REV tail to improve PCR yield (Cho *et al.* 1995; Regier & Shi 2005)) with a thermal profile of: 95°C for 40s, 50°C for 60s and 72°C for 40s, 35 cycles. All PCRs had an initial denaturing step of 94°C for 5 minutes and a final extension step of 72°C for 2 minutes.

For the trnL amplifications we explored a range of primer combinations and amplification conditions with primers described in Taberlet *et al.* (2007) and tailed modifications of these. Previous studies (Jurado-Rivera *et al.* 2009; Pinzón-Navarro *et al.* 2010) have used primers A49325 and B49863 (Taberlet *et al.* 2007), but with limited success. Although Jurado-Rivera (2009) did not report amplification success rates, Pinzón-Navarro *et al.* (2010) reported only a 36% amplification success rate for beetles sampled from the field. Our preliminary investigations achieved improved amplification success with an alternative pair of primers, A49425 (Taberlet *et al.* 2007) and M13(-21)B49863, a modification of B49863 by the incorporation of an M13(-21) tail. A touchdown PCR protocol (Don *et al.* 1991) was used with a thermal profile of: 94°C for 60s, 60°C for 60s (touched down by 1°C per cycle for 17 cycles) and 72°C for 60s, followed by 29 cycles with an annealing temperature of 42°C, giving a range of annealing temperatures between 60°C and 42°C and 46 cycles in total. PCR products comprised of more than one band were separated on a 2% agarose gel and individual bands were sampled from the gel with a pipette tip and resuspended in 30ul of AE buffer (QIAGEN, West Sussex, UK). Resuspensions were used as template for subsequent PCR reactions using the same primers and conditions as above.

Sequencing reactions were performed with the Big Dye Terminator v3.1 Cycle Sequencing kit (Applied Biosystems, California, USA). For COII, the primer COIICraF was used, and an internal sequencing primer, ArgMurF1 was designed as an internal sequencing primer for ArgK. CAS28sB1d and M13REV were used for ITS2 sequencing and for trnL the primer H-rev (a reverse form of primer H (Taberlet *et al.* 2007)) was used. The thermal profile used for all sequencing primers was: 96°C for 10s, 50°C for 5s and 60°C for 240s, 25 cycles, with the exception of COIICraF, for which an annealing temperature of 58°C was used. Sequences were read on a 3730XL sequencer (Applied Biosystems). All sequences were checked and ambiguous bases called in Geneious Pro version 5.4 (Drummond *et al.* 2011). Sequences were aligned using MAFFT v6.814b (Katoh *et al.* 2002) with the following parameter values: scoring matrix 200PAM/k=2, Gap open penalty = 1.53, Offset value = 0.123. The alignments produced were then checked by eye. For the nuclear loci, alleles were resolved for heterozygotes involving nucleotide polymorphism either by direct comparison to homozygous sequences or with PHASE 2.1 (Stephens & Scheet 2005; Stephens *et al.* 2001) as implemented in DNASP v5 (Librado &

Rozas 2009). Indel heterozygotes within ITS2 were resolved by sequencing individuals with both forward and reverse primers and then recoding sections with double peaks as ambiguities. The program Indelligent v1.2 (Dmitriev & Rakitov 2008a; Dmitriev & Rakitov 2008b) was then used to predict the indels present in the mixed trace. The SNPs and indels in the predicted haplotypes for each individual were then compared to known alleles for verification.

Phylogenetic, network and population genetic analyses

For the mtDNA COII sequence data phylogenetic trees were constructed with MrBayes 3.2 (Huelsenbeck & Ronquist 2001) with *C. melanocephalus* and *C. vulgaris* as outgroups. Four analyses were run, each for 10 million generations using 4 MCMC (Markov chain Monte Carlo) chains, discarding 25% of the samples as burnin, with the HKY+G model of sequence evolution as determined by jModeltest (Posada 2008) using BIC (Bayesian Information Criterion) and AIC (Akaike Information Criterion). All parameters permitted under this model were estimated. The output was assessed for stationarity and convergence in Tracer v1.5.0 (Rambaut & Drummond 2007) with only estimated sample sizes (ESS) of greater than 200 being accepted. The consensus tree was visualised in FigTree v1.3.1 (Rambaut 2011). To visualise relationships among the less divergent nuclear gene sequences, haplotype networks were constructed with TCS 1.21 (Clement *et al.* 2000) for non-recombinant sections within each gene, as in Jordal *et al.* (2006). For the quantification of non-recombinant sections both genes were analysed with the recombination detection programs, Dual Brothers (Minin *et al.* 2005) with the default MCMC settings and the program RDP3 (Martin *et al.* 2010) using the default settings. Average and maximum pairwise distances within gene partitions were calculated in MEGA 5 (Tamura *et al.* 2011). Pairwise F_{st} statistics were also generated among sampling sites for nuclear sequence data using Arlequin 3.5 (Excoffier & Lischer 2010).

Divergence time estimation

Two approaches were taken to estimate the timing of divergence between the flightless coral islet populations of *C. murinus* and populations of the main island using a coleopteran COII mutation rate of 0.0154 substitutions/site/myr reported in Cicconardi *et al.* (2010). As a first approach we undertook a BEAST analysis with only the COII partition, employing a constant size coalescent model for all nodes corresponding to intra-population level variation, and a Yule model specified for the remainder of the tree, following the approach of Ho *et al.* (2008). Three replicate analyses were performed using the best fit model characterised by jModelTest and an uncorrelated relaxed clock (Drummond *et al.* 2006). Analyses were run for 100 million generations, and convergence and stationarity were assessed with Tracer v1.5.0. For second approach *BEAST was used with the complete COII and ITS2 (gaps included) data sets and the largest non-recombinant section of ArgK. An uncorrelated relaxed clock was used for all three partitions and substitution models were determined using jModelTest. Three separate analyses, each of 500 million generations were run and convergence and stationarity were assessed with Tracer v1.5.0.

Species assignment of *trnL* sequences

Two approaches were taken to assign *trnL* sequences obtained from beetle specimens to plant species. Sequences were analysed with the program SAP (Munch *et al.* 2008) which seeks to assign taxonomy by MCMC sampling 10,000 unrooted phylogenetic trees constructed with a query sequence and its GenBank homologues. As only 2 sites were sampled for our plant reference library, SAP permits taxonomic inference for food plant species that may not have been represented within these sites. We used a minimum threshold matching probability of 100% for species, genus and family assignment. Each beetle derived *trnL* sequence was then compared to a local plant species reference database constructed from taxonomically identified reference material (see Plant sampling). We sequenced a single individual for each plant species and adopted an acceptance threshold of 95% sequence matching to a beetle derived *trnL* sequence to consider a plant species to be a potential match

To compare the breadth of the observed diets of both *C. murinus* and *C. ovalis* we used the current APG III classification of angiosperms (The Angiosperm Phylogeny Group 2009). For our purpose we define dietary breadth as the sum of the branch lengths among food plant species within a pruned APG III tree representing potential food plant species for *C. murinus* and *C. ovalis*. The phylogenetic placement of *Scaevola taccada*, with regard to the food plants consumed by mainland populations of *C. murinus*, is used as a measure of the novelty of *S. taccada* as a food plant.

Statistical analyses

To assess whether or not changes in diet have evolved among populations of *C. murinus*, we must first test for significant dietary differences among populations, and then evaluate whether any such differences are functional, or a response to differing resource availability among collecting sites. Analyses were performed on the two sites sampled for the greatest number of *C. murinus* where vegetation surveys were subsequently undertaken (see table S4), providing data on both the availability and frequency of each plant species. Chi-square tests were used to evaluate the significance of differences between the frequencies for which each plant species is consumed across both sites, and food plant species availability across sites. The number of sampling plots containing a given plant species was used as a proxy for the availability of each plant species within a site. As a second measure of dietary overlap, Schoener's index (Schoener 1968) was calculated across the two sites. A critical value of 0.6 was taken to mean diets were not significantly different, as suggested by Wallace (1981) and Wallace and Ramsey (1983). For each of the sampled sites Spearman's rank correlation of feeding frequency for each food plant species vs. the proportion of plots in which each food plant species occurred was used to test for a relationship between feeding frequency and resource availability. All analyses were performed in R 2.12.1 (R Development Core Team 2010).

Results

MtDNA and nuclear gene sequencing

Of the 201 beetles sampled, 182 (123 *C. murinus* and 59 *C. ovalis*) were successfully amplified and sequenced for the mtDNA COII gene (GenBank accession numbers JN982748 - JN982931), yielding sequences of 607bp. One hundred and forty sites were variable across both species, with 107 sites variable within *C. murinus* and 53 within *C. ovalis*. There were 69 unique mitotypes (41 in *C. murinus* and 28 in *C. ovalis*) and the average pairwise p-distance was 7.23% across both species with a maximum of 10.38%. Average pairwise p-distance for *C. murinus* sequences was 5.21% with a maximum of 10.38% and the average pairwise distance for *C. ovalis* was 2.07% with a maximum of 3.62%. For ArgK 554bp of sequence was obtained for 117 *C. murinus* (234 haplotypes) and 57 *C. ovalis* (114 haplotypes) with 34 unique haplotypes found within *C. murinus* and 21 within *C. ovalis* (GenBank accession numbers JN987963 - JN988312). Across both species there are 41 variable sites with an average pairwise p-distance of 1.32% and a maximum of 2.53%. The average pairwise distance for *C. murinus* sequences was 1.01% with a maximum of 2.53% and the average pairwise distance for *C. ovalis* was 0.56% with a maximum of 1.08%. ITS2 sequences were obtained for 101 *C. murinus* (202 haplotypes) and 62 *C. ovalis* (124 haplotypes) (a dryad accession number for the alignment of all inferred and sequenced haplotypes will be provided upon acceptance). *Cratopus murinus* ITS2 sequences varied in length from 420bp to 431bp and *C. ovalis* sequences varied from 422bp to 423bp with 22 unique haplotypes within *C. murinus*, and 13 within *C. ovalis*. An alignment of both species was 464bp long with 115 variable sites, an average pairwise p-distance of 4.66% and a maximum pairwise distance of 9.58%. An alignment of *C. murinus* sequences was 448bp in length with 72 variable sites, an average pairwise p-distance of 3.11% and a maximum of 6.94%. An alignment of *C. ovalis* sequences was 424bp long with 13 variable sites, an average pairwise p-distance of 0.16% and a maximum pairwise p-distance of 0.95%. There are at least two *C. murinus* ITS2 alleles with indels and SNPs that could not be attributed to known alleles from homozygotes. These are shared between Corps de Garde and Snail Rock, but are not present in any other site. Samples containing these alleles have been excluded from further analyses. Sequencing coverage for all beetle loci is listed in Table S4.

Phylogenetic analyses and haplotype networks

The Bayesian analysis (Fig. 1) reveals 8 well supported mtDNA lineages. *Cratopus murinus* is divided into 6 main lineages: (1) all Corps de Garde samples, plus some individuals from Mondrain and Yemen; (2) all samples from Snail Rock; (3) all coralline islet samples; (4 & 5) all samples from Trois Mamelles, with a single sample from Yemen within lineage 4, and (6) all samples from Lion Mountain. *Cratopus ovalis* is divided into two distinct lineages: (7) samples from Snail Rock, and (8) samples from all sites. Recombination analyses performed on the ArgK alignment revealed there to be at least 3 recombination breakpoints within this locus, so the alignment was divided into three partitions of length 30bp, 200bp and 208bp for network analyses. The remainder of the locus was unusable due to limited information content within non-recombinant regions. No recombination was detected in the ITS2 alignment and gaps were included for network analysis. Haplotype networks were generated for each of the three

informative non-recombinant ArgK partitions (Fig. 2). The first ArgK partition recovers *C. ovalis* as distinct from *C. murinus*, with no haplotype sharing and monophyly for both species. The ancestral haplotype within *C. murinus* is defined by the network placement of *C. ovalis* haplotypes and reveals that haplotypes from Trois Mamelles and Lion Mountain are derived from haplotypes shared across Snail Rock, Corps de Garde, and the coralline islets. The second ArgK partition distinguishes *C. murinus* from *C. ovalis*, with one haplotype being shared across these two species. Haplotypes unique to the coralline islets are derived from a haplotype distributed across the islets and several mainland sites. The ancestral *C. murinus* haplotype is found in Trois Mamelles with the single haplotype from Lion Mountain being derived from this. As with the first partition, haplotypes from the coralline islets are shared with, or derived from, a haplotype found in Snail Rock and Corps de Garde. The third ArgK partition has few informative sites, so lacks resolution, but does resolve lineage 6 in a manner consistent with the COII tree. A single haplotype network was generated from the ITS2 haplotype data (Fig. 3). *Cratopus murinus* and *C. ovalis* are clearly distinct.

Population genetic analyses and cryptic species assessment

Pairwise F_{st} values were calculated between sufficiently sampled sites for both nuclear loci for *C. murinus* and *C. ovalis*. Most sampling sites are significantly and often highly differentiated for both loci for *C. murinus* (Table 1). In contrast, populations of *C. ovalis* are not significantly differentiated, with most comparisons revealing no statistically significant differentiation (Table 2). Divergent mtDNA lineages are found within the Trois Mamelles population of *C. murinus* (average p-distance = 7.21%) and to a lesser extent the Snail Rock population of *C. ovalis* (average p-distance = 2.73%), the occurrence of which may be explained by high levels of intraspecific variation, but could also be indicative of sympatric cryptic species (e.g., Cicconardi *et al.* 2010). Our null hypothesis is that these divergent mtDNA lineages are consistent with intraspecific genetic variation, and that weevils corresponding to different divergent mitochondrial lineages do not exhibit nuclear genetic differentiation. Within Trois Mamelles, no significant nuclear genetic differentiation was found when *C. murinus* were partitioned into two divergent mitochondrial groups (ArgK: $F_{st} = -0.02$, $p = 0.62$; ITS2: $F_{st} = 0.13$, $p = 0.09$). Similarly, within Snail Rock, *C. ovalis* partitioned into two divergent mtDNA groups were not significantly differentiated at either nuclear locus (ArgK: $F_{st} = -0.01$, $p = 0.75$; ITS2: $F_{st} = 0.00$, $p = 0.48$).

Divergence times

The BEAST analysis estimated an age of 0.89 mya (95% HPD [Highest Probability Density] interval: 0.59 mya – 1.23 mya, ESS = 1992.68) for the divergence between the Snail Rock and coralline islet populations of *C. murinus* and a time for the most recent common ancestor of the coralline islet populations of 0.68 mya (95% HPD interval: 0.42 mya – 0.95 mya, ESS = 3789.91). The *BEAST estimate for the divergence between the coralline islet and Snail Rock populations is slightly older at 1.20 mya (95% HPD interval: 0.22 mya – 2.39 mya, ESS = 2168.04).

Plant analyses

Across all sites, clearly readable trnL sequences were obtained from 164 of the 201 sampled beetles, comprising 108 *C. murinus* and 56 *C. ovalis*. Of the 164 individuals, 16 yielded two differently sized bands that were individually reamplified and sequenced. An additional 5 individuals produced mixed trnL sequence traces that may be attributed to more than one trnL PCR product of similar length. Thirty-two individuals either did not amplify at all or amplified weakly and produced sequences of poor quality. Across all individuals of *C. murinus* and *C. ovalis* a total of 18 different trnL sequences were obtained. Three sequences yielded taxonomic matches within our SAP analyses. The first was identified with 100% probability as belonging to the genus *Betula* and the second with 100% probability to the genus *Pinus*. The *Betula* sequence was found in 19 individuals (13 *C. murinus* from Corps de Garde and Trois Mamelles and 6 *C. ovalis* from Snail Rock and Corps de Garde). The *Pinus* sequence was identified from a single specimen on Ile aux Vacoas. Neither of the plant genera identified were found during floral surveys, and given that both genera are exclusively represented by large woody species, it is unlikely they could have been overlooked. As such both sequences were attributed to contamination and excluded from further analysis. The third sequence was identified with 100% probability as belonging to the family Convolvulaceae, and was found in two individuals from the coralline islet of Ile Marianne. We assign this sequence to the only representative of the Convolvulaceae on Ile Marianne, *Ipomoea*, of which there are two species present, *I. violaceae* and *I. pes-caprae*. Thirteen sequences were each exclusively assigned to one of 13 species within the plant species reference database with a minimum identity of 99.6% (Table 3). Two sequences were assigned to the genus *Eugenia* because they exhibited 98.4-100% identity to multiple species within the genus (Table 3).

Excluding *Betula* and *Pinus* sequences, we obtained trnL sequences from 156 out of 201 individuals (102 *C. murinus* and 54 *C. ovalis*), yielding a total of 158 feeding records, representing an amplification success of 79%. This represents a substantial improvement over the 36% amplification success rate reported in the only other comparable study to date (Pinzón-Navarro *et al.* 2010) and is consistent with our evaluation of different primer combinations. TrnL sequence length varied between 255bp and 409bp and, excluding *Betula* and *Pinus*, a total of 14 different plant species were unambiguously identified from beetle DNA extracts. Across all sites, 13 plant species were consumed by *C. murinus* (Table 3) and five by *C. ovalis* (Table 4), with four species in common. Species of *Eugenia* could not be distinguished by trnL sequences, precluding specific identification of *Eugenia* food plants.

A total of 64 plant species were identified across Corps de Garde and Trois Mamelles (Table S3). DNA sequences obtained from identified plant material can be found using GenBank accession numbers JF804883 - JF804940. Across both sites 10 species of plant were characterised from trnL sequences obtained from *C. murinus* samples (Table 5). Seven of the 10 plant species consumed were found in both Trois Mamelles and Corps de Garde, but only one of these plant species was consumed in both sites (Table 5). The frequency of feeding on each plant species was significantly different between the two sites ($\chi^2 = 31.47$, $df = 9$, $p < 0.001$), and a Schoener index of 0.09 also suggests that diet of *C. murinus* differs between the two sites. However, food plant availability was also significantly different across the two sites ($\chi^2 = 36.89$, $df = 9$, $p < 0.001$). A Spearman's rank correlation of feeding frequency vs. plant availability (proportion of

sites with food plant) was significant ($p = 0.50$, $p = 0.026$). A scatter plot of feeding frequency vs. proportion of sites (Fig. 4) shows that there is one outlier, *Erythroxylum sideroxyloides* in Trois Mamelles, which could be driving this relationship. All analyses were therefore repeated with *Erythroxylum sideroxyloides* omitted (Fig. 4). Both the feeding frequency and availability of each food plant species were still significantly different across the two sites ($\chi^2 = 18.31$, $df = 8$, $p = 0.019$; $\chi^2 = 31.90$, $df = 8$, $p < 0.001$, respectively), however, with the removal of *E. sideroxyloides* the Spearman's rank correlation was not significant ($\rho = 0.40$, $p = 0.103$).

Discussion

Geographic structuring of genetic variation within Cratopus murinus and Cratopus ovalis

Cratopus murinus populations are highly differentiated with mitotypes largely restricted to only one collecting site. This level of structure is reflected in the pairwise F_{ST} values between populations for the ArgK and ITS2 data sets, indicating that among our sampling sites *C. murinus* have limited ability or tendency for dispersal. Notable exceptions to this highly structured pattern can be seen in samples from Yemen, Mondrain, Ile de La Passe and Ile aux Vacoas. *Cratopus murinus* samples from Yemen and Mondrain possess mtDNA sequences that are phylogenetically closely related, or in several cases identical, to *C. murinus* samples from Corps de Garde and Trois Mamelles. Mondrain is a botanic garden, and as such plants introduced from various areas of Mauritius are present. It is therefore conceivable that *C. murinus* in Mondrain are the descendants of larvae or adults accidentally transported with dry forest plants sampled from Corps de Garde. However Yemen is considered one of the best examples of remnant dry forest on Mauritius, and thus it is possible that both Mondrain and Yemen represent natural populations recently founded from Corps de Garde and Trois Mamelles. Two mitotypes are shared across the islets of Ile aux Vacoas and Ile de la Passe, presumably a consequence of their close proximity (210m at their closest point) compared to the more distant Ile Marianne (2.4km from Ile aux Vacoas and 2.7km from Ile de la Passe). Fluctuating sea levels over the last 0.45 mya mean that all three islets were connected to each other and the mainland on at least several occasions (Warren *et al.* 2010). Due to their close proximity and the shallow water bordering the western shores of these islets, Ile aux Vacoas and Ile de la Passe would have had the most substantial connectivity, and as such they can be considered as one recently fragmented population.

Compared to *C. murinus*, *C. ovalis* populations exhibit greater homogeneity with regard to the structuring of genetic variation, suggesting perhaps a more recent origin for the distribution of this species, or higher levels of dispersal relative to *C. murinus*. Both explanations imply a dispersal difference between the two species, which would be intriguing for two ecologically and evolutionarily related species. It has been suggested that the ability to fly and dispersal itself are not necessarily linked, and McCulloch *et al.* (2009) have shown that some insect populations can retain wings while not routinely dispersing by flight. Associations between reduced flight and energetic savings, as well as life history benefits such as increased reproductive output, are well documented (Roff 1990). Zera and Denno (1997) suggest that

reduced flight musculature provides most of the energetic savings resulting from flight loss, meaning that individuals retain some flight capability with the benefits that it brings (e.g., escape from predators), while also achieving energetic savings. The flightlessness of coralline islet populations of *C. murinus* represents an extreme form of this variation, suggesting the potential for functional variation between mainland populations of *C. murinus* and *C. ovalis*, which could be evaluated with comparative data on flight musculature.

Trophic ecology of Cratopus murinus and Cratopus ovalis from dry forest sites

Dry forest populations of both *C. murinus* and *C. ovalis* are dietary generalists with diets encompassing one member of the Laurales and much of the higher order diversity within the core eudicots as defined in APG III (2009), but with *C. ovalis* having a narrower diet than *C. murinus*. However, most food plant species belong to orders that were also represented within each site by other species, none of which were identified as food plants. Diet within our two focal species is not restricted to a few closely related food plant species, as is suggested to be the norm for tropical coleopteran herbivores by Novotny *et al.* (2002) and Ødegaard *et al.* (2005). We suggest that within our study species diet is not constrained by phylogenetic conservatism, with a plausible explanation being that food plant species present a range of structural and chemical defences that fall within a spectrum of palatable values for *C. murinus* and *C. ovalis*. Further examination of the chemical and structural defences presented by the food plants consumed would be useful for determining the ecological constraints on the diet of our focal species.

Both the physiological efficiency hypothesis (Dethier 1954) and the enemy free space hypothesis (Jeffries & Lawton 1984) have been suggested to promote host specialisation in small insect herbivores (Tilmon 2008). In light of their broad diets, this would not seem to be the case for *C. murinus* and *C. ovalis*. However, selective advantages have been described for true polyphagy, where any individual can eat multiple food plants, as opposed to composite generalism proposed by Fox and Morrow (1981), where individuals are monophagous but the whole species or population is polyphagous. True polyphagy carries with it the advantage of being able to gain the full range of nutrients required, while reducing intake of any single plant secondary metabolite (Singer *et al.* 2002). Additionally, if enemy free space is variable in space or time (e.g. changing exposure or predator activity patterns during the day), herbivores that can feed on different plants in response to this would be at a selective advantage over those that cannot (Tilmon 2008). As we found several individuals with multiple food plant sequences in their digestive systems, it seems likely that dry forest *C. murinus* and *C. ovalis* are at least capable of true food mixing, and their evolutionary history may have been influenced by the aforementioned selective pressures.

Variation in Cratopus murinus diet across dry forest sites

By adopting a methodology that combines direct molecular quantification of diet at the level of the individual weevil with floral surveys of two densely sampled sites, we have been able to

examine how and why diet varies among populations of *C. murinus*. We have shown that while diet varies between the populations of Trois Mamelles and Corps de Garde so does the availability of each food plant, and a Spearman's rank correlation reveals that feeding frequency co-varies with availability, although the strength of the correlation is influenced by one particular food plant species. Our results suggest that dietary choices in dry forest populations of *C. murinus* are regulated by the availability of food plants, and that diet variation is a neutral response to the food plant community, and not indicative of any adaptive differences between populations.

Trophic ecology of coralline islet populations of Cratopus murinus

In contrast to the dry forest populations of *C. murinus*, populations from the coralline islets have a restricted diet that consists mostly of *Scaevola taccada*, with two records of a species of *Ipomoea*, despite other core eudicot genera being available. Both *S. taccada* and *Ipomoea* sp. represent novel food plant species, in the context of food plants consumed within dry forest populations of *C. murinus*. However, neither represents a phylogenetically novel food plant as they belong to the orders Asterales and Solanales respectively, both of which are represented by different species in the diet of *C. murinus*. However, a recent survey of another dry forest site, Chamarel, revealed artificially planted *S. taccada* with no feeding damage despite *C. murinus* being collected nearby. As such it appears that *S. taccada* does indeed represent an evolutionary change in host plant preference and/or tolerance in coralline islet populations. Although we can describe a change in the trophic ecology of coralline islet populations, the timing of this is less clear, with dating estimates indicating that coralline islet populations of *C. murinus* probably diverged from extant mainland dry forest populations somewhere between 0.68-1.2mya.

Conclusions

Quantifying trophic interactions between insect herbivore and plant species is important for the understanding of ecosystem functioning. However, it is complicated by the difficulty of obtaining plant-feeding records, particularly when an insect herbivore's life cycle does not involve direct association with plant tissue. We have demonstrated that population-level molecular analysis of insect digestive system contents using DNA sequences can overcome this complication. Our improved success rate in obtaining plant feeding records from DNA sequence data is at least in part due to the reduced fragment length amplified compared with previous studies. Our approach of sequencing plant material from only the upper digestive tract mostly yielded sequences for a single plants species from individual weevils, but in several cases sequences from two plant species were recovered, indicating that analysis of material from the entire digestive tract may increase the proportion of individuals yielding cpDNA sequence data from more than one plant species. While the quality of plant DNA extracted from the entire digestive tract may limit PCR success for larger amplicons, if shorter amplicons are capable of distinguishing among plant species within a study system then parallel sequencing platforms will be more advantageous over Sanger sequencing by avoiding the need for gel separation of different PCR products and/or cloning. When undertaking studies such as this, researchers should be aware that the possibility of contamination and the effects that it may have on data quality. Further studies combining DNA sequence-based measures of plant feeding, DNA

548 sequence libraries for plant species identification, population genetic data and local ecological
549 data, will help develop our understanding of trophic interactions between insect herbivore and
550 plant species.
551
552

For Review Only

553 Acknowledgments

554 We thank Juliane Casquet for field assistance, Sylvain Hugel for beetle samples from the Yemen
555 and Mondrain sites, Zayd Jhumka for beetle samples from Ile Marianne, the Mauritian Wildlife
556 Foundation for logistical assistance and the Mauritius National Parks & Conservation Service for
557 permits. We also thank José Jurardo-Rivera and Jesus Gomez-Zurita for advice regarding trnL
558 amplifications. Fieldwork and labwork were supported by funding from the Agence Nationale
559 de la Recherche, France (BIOTAS: The Southwest Indian Ocean biodiversity hotspot: a biota-
560 level study of diversification on land and sea), and a NERC funded PhD studentship to JJNK.
561 Molecular analysis training for JJNK was supported by the Company of Biologists and the John
562 and Pamela Salter Charitable Trust. The molecular analyses presented in this paper were
563 carried out on the High Performance Computing Cluster supported by the Research Computing
564 Service at the University of East Anglia.

565

566

References

- Barat M, Tarayre M, Atlan A (2008) Genetic divergence and ecological specialisation of seed weevils (*Exapion* spp.) on gorses (*Ulex* spp.). *Ecological Entomology* **33**, 328-336.
- Barone JA (1998) Host-specificity of folivorous insects in a moist tropical forest. *Journal of Animal Ecology* **67**, 400-409.
- Cho SW, Mitchell A, Regier JC, *et al.* (1995) A highly conserved nuclear gene for low-level phylogenetics - Elongation Factor-1-Alpha recovers morphologically-based tree for Heliothine moths. *Molecular Biology and Evolution* **12**, 650-656.
- Cicconardi F, Nardi F, Emerson BC, Frati F, Fanciulli PP (2010) Deep phylogeographic divisions and long-term persistence of forest invertebrates (Hexapoda: Collembola) in the North-Western Mediterranean basin. *Molecular Ecology* **19**, 386-400.
- Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9**, 1657-1659.
- Danforth BN, Lin CP, Fang J (2005) How do insect nuclear ribosomal genes compare to protein-coding genes in phylogenetic utility and nucleotide substitution patterns? *Systematic Entomology* **30**, 549-562.
- Deagle BE, Kirkwood R, Jarman SN (2009) Analysis of Australian fur seal diet by pyrosequencing prey DNA in faeces. *Molecular Ecology* **18**, 2022-2038.
- Dethier VG (1954) Evolution of feeding preferences in phytophagous insects. *Evolution* **8**, 33-54.
- Dmitriev DA, Rakitov RA (2008a) Decoding of Superimposed Traces Produced by Direct Sequencing of Heterozygous Indels. *Plos Computational Biology* **4**.
- Dmitriev DA, Rakitov RA (2008b) Indelligent v1.2, available from <http://imperialis.inhs.illinois.edu/dmitriev/indel.asp>.

- 590 Don RH, Cox PT, Wainwright BJ, Baker K, Mattick JS (1991) 'Touchdown' PCR to circumvent
591 spurious priming during gene amplification. *Nucleic Acids Research* **19**, 4008-4008.
- 592 Drummond A, Ashton B, Buxton S, *et al.* (2011) Geneious v5.4, Available from
593 <http://www.geneious.com/>.
- 594 Drummond AJ, Ho SYW, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with
595 confidence. *PLoS Biol* **4**, e88.
- 596 Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform
597 population genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10**,
598 564-567.
- 599 Farrell BD (1998) "Inordinate fondness" explained: Why are there so many beetles? *Science*
600 **281**, 555-559.
- 601 Fox LR, Morrow PA (1981) Specialization: species property or local phenomenon? *Science* **211**,
602 887-893.
- 603 García-Robledo C, Erickson DL, Staines CL, Erwin TL, Kress WJ (2013) Tropical plant-herbivore
604 networks: Reconstructing species interactions using DNA barcodes. *PLoS ONE* **8**, e52967.
- 605 Gillespie RG, Claridge EM, Goodacre SL (2008) Biogeography of the fauna of French Polynesia:
606 diversification within and between a series of hot spot archipelagos. *Philosophical*
607 *Transactions of the Royal Society B-Biological Sciences* **363**, 3335-3346.
- 608 Grobler GC, van Rensburg LJ, Bastos ADS, Chimimba CT, Chown SL (2006) Molecular and
609 morphometric assessment of the taxonomic status of *Ectemnorhinus* weevil species
610 (Coleoptera : Curculionidae, Entiminae) from the sub-Antarctic Prince Edward Islands.
611 *Journal of Zoological Systematics and Evolutionary Research* **44**, 200-211.
- 612 Hereward JP, Walter GH (2012) Molecular interrogation of the feeding behaviour of field
613 captured individual insects for interpretation of multiple host plant use. *PLoS ONE* **7**, e44435.

- 614 Hernandez-Vera G, Mitrovic M, Jovic J, *et al.* (2010) Host-associated genetic differentiation in a
615 seed parasitic weevil *Rhinusa antirrhini* (Coleoptera: Curculionidae) revealed by
616 mitochondrial and nuclear sequence data. *Molecular Ecology* **19**, 2286-2300.
- 617 Ho SYW, Larson G, Edwards CJ, *et al.* (2008) Correlating Bayesian date estimates with climatic
618 events and domestication using a bovine case study. *Biology Letters* **4**, 370-374.
- 619 Holland BS, Hadfield MG (2002) Islands within an island: phylogeography and conservation
620 genetics of the endangered Hawaiian tree snail *Achatinella mustelina*. *Molecular Ecology* **11**,
621 365-375.
- 622 Huelsenbeck JP, Ronquist F (2001) MrBayes: Bayesian inference of phylogenetic trees.
623 *Bioinformatics* **17**, 754-755.
- 624 Jeffries MJ, Lawton JH (1984) Enemy free space and the structure of ecological communities.
625 *Biological Journal of the Linnean Society* **23**, 269-286.
- 626 Ji YJ, Zhang DX, He LJ (2003) Evolutionary conservation and versatility of a new set of primers
627 for amplifying the ribosomal internal transcribed spacer regions in insects and other
628 invertebrates. *Molecular Ecology Notes* **3**, 581-585.
- 629 Jordal BH, Emerson BC, Hewitt GM (2006) Apparent 'sympatric' speciation in ecologically
630 similar herbivorous beetles facilitated by multiple colonizations of an island. *Molecular*
631 *Ecology* **15**, 2935-2947.
- 632 Jordal BH, Kirkendall LR, Harkstad K (2004) Phylogeny of a Macaronesian radiation: host-plant
633 use and possible cryptic speciation in *Liparthrum* bark beetles. *Molecular Phylogenetics and*
634 *Evolution* **31**, 554-571.
- 635 Jurado-Rivera JA, Vogler AP, Reid CAM, Petitpierre E, Gomez-Zurita J (2009) DNA barcoding
636 insect-host plant associations. *Proceedings of the Royal Society B: Biological Sciences* **276**,
637 639-648.

- 638 Katoh K, Misawa K, Kuma K, Miyata T (2002) MAFFT: a novel method for rapid multiple
639 sequence alignment based on fast Fourier transform. *Nucleic Acids Research* **30**, 3059-3066.
- 640 Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA
641 polymorphism data. *Bioinformatics* **25**, 1451-1452.
- 642 Machado A (2006) The type material of the species of *Laparocerus* Schonherr, 1834 (Coleoptera,
643 Curculionidae, Entiminae). *Journal of Natural History* **40**, 2001-2055.
- 644 Machado A (2007a) New species of *Laparocerus* Schonherr, 1834 (Coleoptera, Curculionidae)
645 from the island of Tenerife (Canary Islands). *Zootaxa* **1457**, 35-48.
- 646 Machado A (2007b) New species of *Laparocerus* Schonherr, 1834 from La Gomera, Canary
647 Islands (Coleoptera, Curculionidae, Entiminae). *Zootaxa* **1643**, 1-38.
- 648 Machado A (2008) Three new endogean species of *Laparocerus* Schonherr, 1834 from the
649 Canary Islands (Coleoptera, Curculionidae). *Journal of Natural History* **42**, 1277-1288.
- 650 Machado A, Lopez M, Almeida T, Hernandez M (2008) Mitochondrial DNA phylogenetic analysis
651 of the genus *Laparocerus* (Coleoptera, Curculionidae, Entiminae). 1 The madeiran clade.
652 *Zoologica Scripta* **37**, 415-427.
- 653 Martin DP, Lemey P, Lott M, *et al.* (2010) RDP3: a flexible and fast computer program for
654 analyzing recombination. *Bioinformatics* **26**, 2462-2463.
- 655 McCulloch GA, Wallis GP, Waters JM (2009) Do insects lose flight before they lose their wings?
656 Population genetic structure in subalpine stoneflies. *Molecular Ecology* **18**, 4073-4087.
- 657 McKenna DD, Sequeira AS, Marvaldi AE, Farrell BD (2009) Temporal lags and overlap in the
658 diversification of weevils and flowering plants. *Proceedings of the National Academy of*
659 *Sciences of the United States of America* **106**, 7083-7088.

- 660 Minin VN, Dorman KS, Fang F, Suchard MA (2005) Dual multiple change-point model leads to
661 more accurate recombination detection. *Bioinformatics* **21**, 3034-3042.
- 662 Morris MG (1997) *Broad-Nosed Weevils*, Handbooks for the identification of British Insects: Vol
663 5 Part 17a, Royal Entomological Society.
- 664 Munch K, Boomsma W, Huelsenbeck JP, Willerslev E, Nielsen R (2008) Statistical assignment of
665 DNA sequences using Bayesian phylogenetics. *Systematic Biology* **57**, 750-757.
- 666 Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J (2000) Biodiversity hotspots
667 for conservation priorities. *Nature* **403**, 853-858.
- 668 Novotny V, Basset Y, Miller SE, *et al.* (2002) Low host specificity of herbivorous insects in a
669 tropical forest. *Nature* **416**, 841-844.
- 670 Novotny V, Drozd P, Miller SE, *et al.* (2006) Why are there so many species of herbivorous
671 insects in tropical rainforests? *Science* **313**, 1115-1118.
- 672 Oberprieler RG, Marvaldi AE, Anderson RS (2007) Weevils, weevils, weevils everywhere.
673 *Zootaxa*, 491-520.
- 674 Ødegaard F, Diserud OH, Østbye K (2005) The importance of plant relatedness for host
675 utilization among phytophagous insects. *Ecology Letters* **8**, 612-617.
- 676 Otte D (1976) On feeding patterns in desert grasshoppers and the evolution of specialized diets.
677 *Proceedings of the Academy of Natural Sciences of Philadelphia* **128**, 89-126.
- 678 Pinzón-Navarro S, Jurado-Rivera JA, Gómez-Zurita J, Lyal CHC, Vogler AP (2010) DNA profiling
679 of host-herbivore interactions in tropical forests. *Ecological Entomology* **35**, 18-32.
- 680 Podlussany A, Jermy T, Szentesi A (2001) On the leguminous host plants of seed predator
681 weevils (Coleoptera : Apionidae, Curculionidae) in Hungary. *Acta Zoologica Academiae*
682 *Scientiarum Hungaricae* **47**, 285-299.

- 683 Posada D (2008) jModelTest: Phylogenetic model averaging. *Molecular Biology and Evolution*
684 **25**, 1253-1256.
- 685 R Development Core Team (2010) R: A Language and Environment for Statistical Computing,
686 Vienna, Austria.
- 687 Rambaut A (2011) FigTree v1.3.1. In: <http://tree.bio.ed.ac.uk/software/figtree/>
- 688 Rambaut A, Drummond A (2007) Tracer v1.5.0. In: <http://beast.bio.ed.ac.uk/Tracer>
- 689 Regier JC, Shi D (2005) Increased yield of PCR product from degenerate primers with
690 nondegenerate, nonhomologous 5' tails. *Biotechniques* **38**, 34.
- 691 Roff DA (1990) The evolution of flightlessness in insects. *Ecological Monographs* **60**, 389-421.
- 692 Schoener TW (1968) The anolis lizards of Bimini: Resource partitioning in a complex fauna.
693 *Ecology* **49**, 704-726.
- 694 Sequeira AS, Lanteri AA, Albelo R, Bhattacharya S, Sijapati M (2008) Colonization history,
695 ecological shifts and diversification in the evolution of endemic Galapagos weevils. *Molecular*
696 *Ecology* **17**, 1089-1107.
- 697 Singer MS, Bernays EA, Carrière Y (2002) The interplay between nutrient balancing and toxin
698 dilution in foraging by a generalist insect herbivore. *Animal Behaviour* **64**, 629-643.
- 699 Staudacher K, Wallinger C, Schallhart N, Traugott M (2011) Detecting ingested plant DNA in soil-
700 living insect larvae. *Soil Biology and Biochemistry* **43**, 346-350.
- 701 Stephens M, Scheet P (2005) Accounting for decay of linkage disequilibrium in haplotype
702 inference and missing-data imputation. *American Journal of Human Genetics* **76**, 449-462.
- 703 Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction
704 from population data. *American Journal of Human Genetics* **68**, 978-989.

- 705 Taberlet P, Coissac E, Pompanon F, *et al.* (2007) Power and limitations of the chloroplast trnL
706 (UAA) intron for plant DNA barcoding. *Nucleic Acids Research* **35**, Article No.: e14.
- 707 Tamura K, Peterson D, Peterson N, *et al.* (2011) MEGA5: Molecular Evolutionary Genetics
708 Analysis using maximum likelihood, evolutionary distance, and maximum parsimony
709 methods. *Molecular Biology and Evolution* **28**, 2731–2739.
- 710 The Angiosperm Phylogeny Group (2009) An update of the Angiosperm Phylogeny Group
711 classification for the orders and families of flowering plants: APG III. *Botanical Journal of the*
712 *Linnean Society* **161**, 105–121.
- 713 Tilmon KJ (2008) *Specialization, Speciation, and Radiation. The Evolutionary Biology of*
714 *Herbivorous Insects*, University of California Press, Los Angeles.
- 715 Wallace RK (1981) An Assessment of Diet-Overlap Indexes. *Transactions of the American*
716 *Fisheries Society* **110**, 72–76.
- 717 Wallace RK, Ramsey JS (1983) Reliability in Measuring Diet Overlap. *Canadian Journal of*
718 *Fisheries and Aquatic Sciences* **40**, 347–351.
- 719 Warren BH, Strasberg D, Bruggemann JH, Robert PP-J, Christophe T (2010) Why does the biota
720 of the Madagascar region have such a strong Asiatic flavour? *Cladistics* **26**, 526–538.
- 721 Williams JR (2000) A revision of the Mascarene weevil genus *Syzygops* Schonherr (Coleoptera :
722 Curculionidae : Entiminae). *Invertebrate Taxonomy* **14**, 411–432.
- 723 Williams JR, Cox ML (2003) A Contribution to the study of Mascarene Weevils of the genus
724 *Cratopus* Schonherr (Coleoptera: Curculionidae: Entiminae: Cratopini): The species of
725 Mauritius and Rodrigues. *The Mauritius Institute Bulletin* **12**.
- 726 Zera AJ, Denno RF (1997) Physiology and ecology of dispersal polymorphism in insects. *Annual*
727 *Review of Entomology* **42**, 207–230.

728

729 **Data Accessibility**

730 - Genbank accessions for DNA sequences: COII (JN982748 - JN982931), ArgK
731 (JN987963 - JN988312) and reference TrnL sequences (JF804883 - JF804940).
732 - An alignment of phased ITS2 haplotypes for each individual included in
733 the analyses is available on Dryad (doi:10.5061/dryad.ss427).

734

735 **Author Contributions**

736 JJNK undertook most beetle sampling, performed laboratory work and data
737 analyses. BFVF and CB performed all plant sampling and identification. BW
738 and DS contributed to fieldwork. JJNK and BCE wrote the manuscript; all
739 authors discussed the results and commented on the manuscript. BCE
740 conceived and supervised the project.

741

Figure 1. MrBayes consensus tree of mtDNA COII DNA sequence data sampled from *Cratopus murinus* and *C. ovalis*, with *C. melanocephalus* and *C. vulgaris* as outgroups. The analysis was performed using 4 MCMC chains for 10 million generations. Nodal values are posterior probabilities for each node. Insert is a map of Mauritius with sampling sites colour coded with reference to the phylogeny.

Figure 2. Haplotype networks of putatively non-recombined sections of ArgK sampled from *Cratopus murinus* and *C. ovalis*. Circle area is proportional to the frequency of a haplotype. Colours correspond to the sample sites in Figure 1, and differently coloured segments within circles represent haplotype sharing across sites, but not haplotype frequency.

Figure 3. Haplotype network of ITS2 sequence data sampled from *Cratopus murinus* and *C. ovalis*. Each indel event is represented by a single step, irrespective of the size of an indel. Longer connections have been labelled with the number of SNPs and indel events responsible for the number of steps inferred. Circle area is proportional to the frequency of a haplotype. Colours correspond to the sample sites in Figure 1, and differently coloured segments within circles represent haplotype sharing across sites, but not haplotype frequency.

Figure 4. (A) Scatter plot of the number of feeding records vs. food plant availability for *Cratopus murinus* sampled from Corps de Garde and Trois Mamelles. The solid black line is a line of best fit though the data. The red stars highlight the points for *Erythroxylum sideroxyloides* in both sites for comparison. (B) As for (A) but with *Erythroxylum sideroxyloides* removed.

Table 1. Pairwise F_{st} values between sampling sites for *C. murinus*. Non-significant results are in bold (significance level = 0.05). Values above the diagonal are calculated using ITS2 data, those below the diagonal are calculated using ArgK.

	Corps de Garde	Coralline islets	Snail Rock	Lion Mountain	Trois Mamelles
Corps de Garde	-	0.57	0.12	0.94	0.95
Coralline islets	0.27	-	0.60	0.94	0.94
Snail Rock	0.15	0.13	-	0.98	0.97
Lion Mountain	0.80	0.83	0.85	-	0.93
Trois Mamelles	0.77	0.82	0.86	0.86	-

Table 2. Pairwise F_{st} values between sampling sites for *C. ovalis*. Non-significant results are in bold (significance level = 0.05). Values above the diagonal are calculated using ITS2 data, those below the diagonal are calculated using ArgK.

	Corps de Garde	Snail Rock	Trois Mamelles
Corps de Garde	-	-0.02	0.04
Snail Rock	0.05	-	0.03
Trois Mamelles	-0.03	0.02	-

Table 3. Plant species consumed by *C. murinus* by sampling locality. N = number of sequences retrieved. Match = % sequence similarity of sequences derived from *C. murinus* to a plant species reference database sequence.

Locality	Species consumed	N	Match
Corps de Garde	<i>Hilsenbergia petiolaris</i> (Malvales)	4	100
	<i>Pittosporum ferugineum</i> (Apiales)	6	100
	<i>Schinus terebinthifolius</i> (Sapindales)	4	100
	<i>Litsea glutinosa</i> (Laurales)	1	99.7
Coralline islets	<i>Scaevola taccada</i> (Asterales)	47	100
	<i>Ipomoea</i> sp. (Solanales)	2	-
Trois Mamelles	<i>Erythroxylum sideroxyloides</i> . (Malpighiales)	13	99.4-100
	<i>Eugenia</i> sp. (Myrtales)	2	98.4-100
	<i>Cassine orientalis</i> (Celestrales)	1	100
	<i>Flacourtia indica</i> (Malpighiales)	1	100
	<i>Margaritaria anomala</i> (Malpighiales)	2	99.7
	<i>Coffea myrtifolia</i> (Gentianales)	1	100
	<i>Hilsenbergia petiolaris</i> (Malvales)	2	100
Snail Rock	<i>Hilsenbergia petiolaris</i> (Malvales)	8	100
Lion Mountain	<i>Psiadia viscosa</i> (Asterales)	8	100

Table 4. Plant species consumed by *C. ovalis* by sampling locality. N = number of sequences retrieved. Match = % sequence similarity of sequences derived from *C. ovalis* to a plant species reference database sequence.

Locality	Species consumed	N	Match
Corps de Garde	<i>Flacourtia indica</i> (Malpighiales)	1	100
	<i>Schinus terebinthifolius</i> (Sapindales)	1	100
	<i>Hilsenbergia petiolaris</i> (Malvales)	2	100
Trois Mamelles	<i>Margarataria anomala</i> (Malpighiales)	1	99.7
	<i>Ageratum conysoides</i> (Asterales)	1	100
Snail Rock	<i>Hilsenbergia petiolaris</i> (Malvales)	48	99.7-100
	<i>Margarataria anomala</i> (Malpighiales)	2	99.7

Table 5. Feeding and abundance records for food plants of *C. murinus* in Corps de Garde and Trois Mamelles. Unbracketed numbers are the number of feeding records for a plant. Bracketed numbers are the proportion of samples plots in which a food plant was recorded.

Plant species	Corps de Garde	Trois Mamelles
<i>Litsea glutinosa</i>	1 (0.62)	0 (0.25)
<i>Hilsenbergia petiolaris</i>	4 (0.15)	2 (0.08)
<i>Pittosporum ferugineum</i>	6 (0.15)	0 (0.08)
<i>Erythroxylum sideroxyloides</i>	0 (0.15)	13 (1)
<i>Schinus terebinthifolius</i>	4 (0.77)	0 (0.08)
<i>Eugenia sp.</i>	0 (0.31)	2 (0.92)
<i>Cassine orientalis</i>	0 (0)	1 (0.33)
<i>Flacourtia indica</i>	0 (0.46)	1 (0.17)
<i>Margaritaria anomala</i>	0 (0)	2 (0.33)
<i>Coffea myrtifolia</i>	0 (0)	1 (0.25)