Reliable, verifiable, and efficient monitoring of biodiversity via metabarcoding: Supporting Information

Yinqiu Ji1,*, Louise Ashton2,*, Scott M. Pedley3, David P. Edwards4,5,12,*,4, Yong Tang6, Akihiro Nakamura2,7, Roger Kitching2, Paul M. Dolman3, Paul Woodcock8, Felicity A. Edwards8, Trond H. Larsen9, Wayne W. Hsu10, Suzan Benedick11, Keith C. Hamer8, David S. Wilcove4,12, Catharine Bruce13, Xiaoyang Wang1, Taal Levi14,15, Martin Lott16, Brent C. Emerson17, Douglas W. Yu1,13,**

1 State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650223 China; jiyinqiu@hotmail.com, wangxiaoy0126@hotmail.com
2 Environmental Futures Centre and Griffith School of Environment, Griffith University, Nathan, Queensland 4111 Australia; l.ashton@griffith.edu.au, r.kitching@griffith.edu.au
3 School of Environmental Sciences, University of East Anglia, Norwich Research Park, Norwich, Norfolk NR47TJ UK; smpedley@gmail.com, p.dolman@uea.ac.uk
4 Woodrow Wilson School of Public and International Affairs, Princeton University, Princeton, New Jersey 08544 USA; dave.edwards@jcu.edu.au
5 Centre for Tropical Environmental and Sustainability Science (TESS), School of Marine and Tropical Biology, James Cook University, Cairns, Queensland 4878 Australia
6 Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Meng La, Yunnan 666303 China; tangy@xtbg.ac.cn
7 Queensland Museum, South Brisbane, Queensland 4101 Australia; a.nakamura@griffith.edu.au
8 Institute of Integrative and Comparative Biology and Faculty of Biological Sciences, University of Leeds, Leeds, West Yorkshire LS29JT UK; bgy3pw@leeds.ac.uk, bs082a@leeds.ac.uk, k.c.hamer@leeds.ac.uk
9 The Betty and Gordon Moore Center for Ecosystem Science and Economics, Conservation International, Arlington, Virginia 22202 USA; trondhilsen@gmail.com
10 Department of Ecology, Evolution, and Environmental Biology, Columbia University, New York City, New York 10027 USA; waynehsu@birdlover.com
11 School of Sustainable Agriculture, Universiti Malaysia Sabah, Sandakan, Sabah 9000 Malaysia; suzanben@ums.edu.my
12 Department of Ecology and Evolutionary Biology, Princeton University, Princeton, New Jersey 08544 USA; dwilcove@princeton.edu
13 School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich, Norfolk NR47TJ UK; catharine.powell@uea.ac.uk
14 Cary Institute of Ecosystem Studies, Millbrook, New York 12545 USA; thetaaltree@gmail.com
15 Department of Biology, University of Florida, Gainesville, Florida 32611 USA
16 School of Computing Sciences, University of East Anglia, Norwich Research Park, Norwich, Norfolk NR47TJ UK; martin@lott.co.uk
17 Island Ecology and Evolution Research Group, IPNA-CSIC, Tenerife, Canary Islands 38206 Spain; bemerson@ipna.csic.es

* These authors contributed equally.
** Correspondence. E-mail: dougwyu@gmail.com

*These authors contributed equally.
** Correspondence. E-mail: dougwyu@gmail.com

* These authors contributed equally.
** Correspondence. E-mail: dougwyu@gmail.com
Supplementary Materials deposited in the Dryad Repository: http://dx.doi.org/10.5061/dryad.t3v71

1. Example R script for analysis: Example_R_script.R
2. Example MBC arthropod dataset: MBCarthropods.txt
3. Example MBC environmental variables dataset: arthropods.env.txt
4. Monte Carlo R script for the RSW2 analysis in Figure 2: RSW2_Test.R
5. RSW2 dataset for Figure 2: RSW2SoftwareResults.txt

Supporting Information 1. Detailed sampling methods for the biodiversity datasets

1.1 Ailaoshan.

Motivation. – The rainforest of south-western China is a biodiversity hotspot, with high levels of endemism; it is also an area that is subject to extensive human impact. This area is predicted to be impacted by climate change, resulting in increased average temperature and precipitation and more frequent extreme weather events (Ding et al. 2007; IPCC 2007). In order to understand the impacts of future climate change on the biodiversity of this region, we first need to establish baseline information on the current distributions of taxa. Altitudinal gradients are an ideal study system with steep shifts in environmental variables in a small geographical area. Using mountain systems as surrogates for changes in temperature and precipitation, we are able to examine the current distribution patterns of taxa, and make predictions about future shifts, based on each species’ current climate envelope. Little attention, in general, has been paid to invertebrates in conservation assessments of the impacts of climate change, yet terrestrial arthropods are, on the one hand, key drivers of ecological processes (Wilson 1987) and, on the other, excellent predictors of environmental change (Basset et al. 1998). Understanding how arthropod assemblages respond to climate (and the associated vegetation assemblages) is key in understanding likely future changes to diversity and distribution of arthropods. We established a set of three permanent altitudinal transects in Yunnan province, China (tropical, sub-tropical and temperate forest). Here we present the results of the first transect, in subtropical rainforest at Ailaoshan.

Study site – The Ailao Mountains Reserve is a protected forest covering 504 km², at around 24.5° N latitude, approximately 200 km south-west of Yunnan capital of Kunming. The Ailao Mountains occur at a major climatic border between the south west and south east monsoon systems of China (Young & Wang 1989). The sub-tropical climate of the Qian Jia Zai area (24.28° N, 101.26° S) has an average temperature of 11 °C, with average annual rainfall around 1900 mm with a dry season between December and April. This area encompasses a large tract of evergreen broad-leaved forests primarily dominated by Lithocarpus and
Castanopsis at mid elevation (ca. 2200-2600 m a.s.l.) with dense understory of bamboo, and Rhododendron dwarf forest towards higher elevations.

**Sampling protocol.** - Sampling of the moth assemblage occurred from 31 July to 21 August 2011. Moths were sampled using modified Pennsylvania light traps (Yang et al. 2004; Kitching et al. 2005). These light traps employ a vertical actinic tube, run from dusk to dawn, using 12 V lead acid batteries. The Pennsylvania light trap design has been modified, with a larger hole in the trap opening for large moths, lightweight large tin lids to keep out rain, and automatic timers. A Dichlorvos-impregnated strip, wrapped in paper towelling, is placed in the trap to kill moths *in situ*, along with cardboard material to reduce damage to moth specimens by beetles.

At each of the four altitudes (2000, 2200, 2400, and 2600 m a.s.l.), moths were sampled in five blocks. Two traps were run simultaneously in each block for three nights, with one trap approximately 2 m above the ground, and one raised into the canopy. Canopy lines were placed as high as possible in the canopy. The average height of the canopy drops with increasing altitude, and therefore the average height of the canopy traps was higher at the lower altitude sites. The total number of collections was 40 (4 altitudes X 5 blocks X 2 strata), but one sample was lost and so the final metabarcoding dataset consists of 39 samples. For some of the 39 samples, moth numbers were great enough that half the sample could be set aside in 100% EtOH for subsequent metabarcoding and the other half used for morphospecies assignment. For the smaller-volume samples, moths were extracted from the collection, identified to morphospecies, and then placed in 100% EtOH, along with the non-indicator-taxa specimens, for subsequent metabarcoding. For each pinned moth (see next paragraph), two legs were removed and placed in the corresponding metabarcoding sample.

All moths with a wing length greater than 1 cm were sorted to morphospecies. In the case of moths belonging to the superfamily Pyraloidea, moths with a wing length less than 1 cm were also processed. During the field session, at least 5 representatives from each morphospecies were pinned and dried. Moths were sorted into approximate morphospecies groups by volunteers, and then sorted and identified to family by Roger Kitching and Louise Ashton. Moths that could be confidently identified as belonging to an existing morphospecies were then recorded and discarded. Any moths where the identification was ambiguous were pinned for later identification.
1.2 Thetford

Motivation. - In Europe, much conservation concern is for species now dependent on semi-natural habitats. Large areas of lowland heathland, developed on unproductive soils, were maintained by livestock grazing and episodic cultivation over centuries. Heathland assemblages are of high biodiversity value, recognised in inclusion of heathland biotopes in Annexes of the EC Habitats Directive (EC 1992), such that member states have an obligation to designate and ensure favourable conservation status of examples of the habitat (as Special Areas of Conservation). Lowland heathland has been considerably reduced by 60-94% across Western Europe with remnants often remaining as isolated fragments. Reconnecting fragmented heathland is important to conserve its biodiversity in the longer term. Many stenotopic heath species require physical disturbance that exposes substrate, creates ruderal resources and sparse early-successional structures. There is potential to use disturbance management to enhance ecological connectivity for heathland biodiversity, taking advantage of existing trackway networks, but there is a need for more robust evidence across multiple taxa (Pedley et al. 2013).

We examined responses of carabid, spider, and ant assemblages to physical disturbance treatments in a trackway network within a large pine plantation in eastern England that was planted over areas of lowland heathland and marginal farmland (Pedley et al. 2013). Within the forest, 1290 km of trackways provide potential for connectivity, both among permanent and ephemeral open habitat elements within the forest landscape, and across the forest to re-connect external heathland remnants. We examined whether physical disturbance treatments enhanced the value of this ecological network. Response to treatments was examined in terms of assemblage composition, species richness and abundance of early-successional specialist and generalist forest species; invertebrate assemblages were also compared to reference heath sites.

Study site. - Thetford Forest was planted in the early 20th century and occupies 185 km² of Breckland in eastern England (0°40' E, 52°27' N). Breckland is characterised by semi-continental climate, sandy, nutrient-poor soil and a long history of grazing and episodic cultivation supporting a regional biota that includes coastal, continental and Mediterranean elements. Physically disturbed heathland and ruderal habitats support at least 542 priority species (rare, scarce, range-restricted or UK Biodiversity Action Plan species) (Dolman et al. 2012). The forest is dominated by conifer plantations, with 80% comprised of Corsican
(Pinus nigra) and Scots (P. sylvestris) pine, managed by clear-felling (typically at 60-80 years) and replanting of even-aged coupes (mean area 9.0 ha ± 8.6 SD). Coupes are subdivided by a network of forestry trackways that provide management access. Trackways comprised two elements: central wheelings with sparse vegetation and exposed substrate, flanked by vegetated verges that are cut annually to facilitate access but lack bare substrate. Trackways vary in width (mean 13.7 m ± 5.8 SD, range 5-50 m, sample size n=93), substrate (sand, gravel), vegetation and shading due to adjacent tree height. Approximately 50% of heathland-associated carabid species have been recorded from this trackway network, as well as many characteristic heathland spider species (Pedley et al. 2013); however, some of the region’s rarest and most exacting species appear absent.

Experimental protocol. - Six physical disturbance treatments that varied in intensity, plus a set of non-managed controls, each replicated nine times across a total of 63 plots (length 150 m, average width 4 m), were established within the trackway system in February 2009. Treatments included two cutting treatments: swiping (sward cut with tractor mounted blades, clippings left in-situ) and forage harvesting (sward cut and removed with silage harvester) and four soil disturbance treatments ranging from mild disruption by disc ploughing (tractor-pulled disc harrow, disrupting but not destroying vegetation with shallow soil disturbance, 10-20 cm deep), to moderate disturbance by forest ploughing (soil and litter inverted in plough lines producing bare mineral substrate in the furrow, width 30-40 cm, depth 40-50 cm, alternating with strips 40-50 cm of intact vegetation), heavy disturbance by agricultural ploughing (turf and top-soil inverted producing bare-substrate across the plot, with biomass retained and buried to 20-30 cm), and the most destructive treatment turf stripping (removal of vegetation, root mat, litter and organic soil exposing mineral subsoil at a depth of 15-30 cm). Photographs of treatments are below.

Plots were placed within trackways at least 9 m wide, within coupes aged 10-25 years that comprise closed-canopy stands lacking open-habitat carabid-beetle, spiders, or plant species. All plots were located a minimum of 100 m away from other treatments, open area or felled coupes to ensure samples were not capturing open-habitat species from adjacent areas.

Treatments were allocated randomly to suitable trackways, stratifying between acidic soils 1) lacking, or 2) dominated by, bracken Pteridium aquilinum and 3) calcareous soils. To reduce shading effects, plots were established in the widest verge of trackways oriented north-south, or the northern verge of trackways oriented east-west.
Thetford Forest Restoration treatments

- Swipe
- Forage harvest
- Disc plough
- Control
- Forestry plough
- Agricultural plough
- Turf stripping
Invertebrate sampling in control and treatment plots. - In both 2009 and 2010, ground-active invertebrates were sampled in each plot on three occasions: in May, June and late July / early August (Pedley et al. 2013). In each period, six pitfall traps (each 7.5 cm deep, 6.5 cm diameter, filled with 50 ml of 70% ethylene glycol) set 15 m apart in a single transect along the centre of each plot (beginning 37.5 m from each end) were opened for seven consecutive days. Traps in each transect were combined giving one composite sample per plot-year.

Invertebrate sampling in heath reference sites. - Ground-active invertebrates were also sampled on seven heath reference sites, located within 8 km of treatment plots, of which eight were designated under EU and or UK conservation legislation (Pedley et al. 2013). All were subject to conservation management, predominantly rabbit and sheep grazing, with some mechanical disturbance. Within each site, three transects were set (each of six pitfall traps of the same dimensions used in experimental plots) at least 50 m apart, open for seven consecutive days over three trapping periods (May, June and August) in 2009.

Indicator taxa were first sorted to carabids, ants, and spiders using paid undergraduates. The ants were identified by the Norfolk county ant recorder, the carabids were identified by the Norfolk beetle recorder and an amateur coleopterist, and Scott Pedley identified all spiders. Adult spiders were identified to species following Roberts (1987, 1996); juveniles and sub-adults were not identified due to the lack of developed reproductive structures. Identification of carabids followed Luff (2007) and ants followed Bolton and Collingwood (1975), Skinner and Allen (1996) and Blacker and Collingwood (2002).

Samples from 68 of the 70 sites in the 2010 trackway and 2009 heathland collections were available for metabarcoding and contained legs of all the designated indicator taxa (spiders, carabids, and ants) plus the unidentified bycatch, which included Insecta (Orthoptera, Diptera, non-carabid Coleoptera, Plecoptera, Lepidoptera, and Hemiptera), Isopoda, Collembola, and Myriapoda. After metabarcoding, one control (undisturbed trackway) site was deleted from both the STD and MBC datasets because it was discovered that it had been inadvertently mowed before sampling, leaving 67 sites for analysis.
1.3 Danum Valley

Motivation. - We need an improved understanding of how tropical production forests fit into conservation agendas both ecologically and financially. Understanding the trade-offs between biodiversity conservation and financial returns from logging is critical for developing effective conservation strategies. Efficiency is an important goal for conservation, given the increasing isolation of protected areas, increasing habitat loss, and limited conservation funding. We use biodiversity and financial datasets to provide an empirical analysis of the trade-offs between human modification of a landscape via logging and biodiversity protection. In doing so, it allows us to explore the role of costs in identifying efficient conservation agendas.

Study site. - Our study is based in and around the Yayasan Sabah (YS) logging concession in eastern Sabah, Malaysian Borneo. Within the YS concession is the Danum Valley Conservation Area and Palum Tambun Watershed Reserve, comprising a combined area of 45,200 ha of unlogged (primary) lowland dry Dipterocarp rainforest, which is dominated by valuable timber species of the Dipterocarpaceae (Fisher et al. 2011). Contiguous with this primary forest is the 238,000 ha Ulu Segama-Malua Forest Reserve (US-MFR; again part of the YS concession), which includes selectively logged forests that have undergone either one or two rotations of timber extraction. Sampled locations in once-logged forest were logged between 1987 and 1991 using a modified uniform system in which all commercial stems >0.6 m diameter were removed (yielding an average of 120 m$^3$ of timber per ha) (Fisher et al. 2011). Our twice-logged locations were logged using the same methods during the first rotation, and again between 2001-2007, employing the same logging techniques but with the minimum tree diameter reduced to >0.4 m (>0.25 m in some cases) and resulting in an additional 15–72 m$^3$ of timber extracted per ha (Edwards et al. 2011; Fisher et al. 2011). To the north, east and south of the US-MFR are oil palm plantations, where sampled sites had mature palms (20-30 years old) at a density of 100 trees per ha (Edwards et al. 2010).

Sampling protocols. - Fieldwork was conducted from July to October 2007, May to August 2008, May to October 2009, and February-April 2011. Fourteen widely spaced sites (between 1–43 km apart) were established within the unlogged, once-logged and twice-logged forests, and oil palm, comprising four sites >2 km apart within each forest type and two sites 3 km apart in oil palm. Each site had two line transects spaced 500-800 m apart (Edwards et al. 2011), and each transect is treated as a sample, for a total of 28 transects, 4 of which were
located in oil palm plantations, and 8 each in the forest sites. Our study taxa were sampled in subsets of these transects (STD: \(n_{\text{birds}} \times \text{mist nets} = 24\) [0 in oil palm]; \(n_{\text{dung beetles}} = 28\), \(n_{\text{leaf-litter ants}} = 26\) [6 in 1L]; MBC: \(n_{\text{malaise traps}} = 56\) traps at 2 traps per transect, subsequently pooled within-transect for \(n_{\text{malaise traps}} = 28\) samples). Birds were identified in the field; confusion species were identified from photographs with reference to field guides by DPE. Dung beetles (THL & FAE) and ants (PW) were identified with reference to collections made by taxonomic experts.

Birds (1,738 individuals; 89 species). We used mist nets to survey the lower-storey community in 2007–2009. On each transect, we erected 15 mist nets in a line; nets were opened from 06.00–12.00 on three consecutive days (Edwards et al. 2011) (6,480 mist net hours in total). To control for variation in canopy height, we restricted our analysis to species that are not defined as canopy specialists (n = 33 individuals of 14 species removed) (Edwards et al. 2011).

Dung beetles (25,319 individuals; 64 species). We used standardized pitfall traps baited with human dung (Larsen & Forsyth 2005) in 2009 and 2011. On each transect, five traps were spaced at 100-m intervals (140 traps in total) (Edwards et al. 2011); traps were collected every 24 h for four days and were re-baited after two days.

Leaf-litter ants (3,256 colonies; 256 species). We used mini-Winkler extractors in 2007–2009 and 2011 (Woodcock et al. 2011). On each transect, 1 m² of leaf litter and loose topsoil were collected from seven census points spaced at 25-m intervals (182 points in total)(Woodcock et al. 2011). Material was sieved to remove larger debris and hung inside the extractors for four days, after which minor workers were removed. Two transects were not collected for ants due to inclement weather.

Malaise traps. In 2011, we used Malaise traps to sample flying invertebrates for the metabarcoding dataset. On each transect, we erected two malaise traps spaced >150 m apart (56 malaise traps in total); traps were collected after four days. The two samples per transect were DNA-extracted and PCR-amplified separately, but the samples were pooled within transect (from n=56 to n=28) for metabarcoding and subsequent comparison with the bird, dung-beetle, and ant datasets.
Supporting Information 2. Ailaoshan MBC Arthropoda ordination

Non-metric multidimensional scaling (NMDS) ordination of the Ailaoshan MBC Arthropoda dataset. Points are census sites, and coloured ellipses are 95% confidence intervals of Altitude-Layer treatment combination species centroids (‘ordiellipses’ in Ref. Oksanen et al. 2012). Sites within the same altitude are connected by line segments, with ellipses are drawn for each treatment combination (light green is used only to differentiate the 2600 m ordiellipses).
Supporting Information 3. *mvabund* statistical tables

*mvabund* commands and statistical output tables for the STD and MBC datasets.
### Ailaoshan

**MBC dataset**

```r
MBCmvabund.glm1 <- manyglm(MBCmvabund ~ altitude*type, family="binomial")
MBCmvabund.anova1.pit <- anova(MBCmvabund.glm1, resamp="pit.trap", show.time=TRUE, nBoot=299)
MBCmvabund.anova1.pit$table

| Res.Df | Df.diff | Dev       | Pr(>Dev)    |
|--------|---------|-----------|-------------|
| (Intercept) | 38      | NA        | NA          |
| altitude | 37      | 1         | 4291.077    | 0.003344482 |
| type    | 36      | 1         | 2952.852    | 0.003344482 |
| altitude:type | 35    | 1      | 1514.186    | 0.481605351 |
```

MBCmvabund.glm2 <- manyglm(MBCmvabund ~ altitude+type, family="binomial")

```r
MBCmvabund.anova2.pit <- anova(MBCmvabund.glm2, p.uni="unadjusted", resamp="pit.trap", show.time=TRUE, nBoot=999)
MBCmvabund.anova2.pit$table

| Res.Df | Df.diff | Dev       | Pr(>Dev)    |
|--------|---------|-----------|-------------|
| (Intercept) | 38      | NA        | NA          |
| altitude | 37      | 1         | 4291.077    | 0.001001001 |
| type    | 36      | 1         | 2952.852    | 0.001001001 |
```

### STD dataset

mvabund commands

```r
lepm1.nb <- manyglm(lepmvabund~altitude*type, family="negative.binomial")
lepmvabund.anova1.pit <- anova(lepm1.nb, p.uni="adjusted", nBoot=299, resamp="pit.trap", show.time=TRUE)
lepmvabund.anova1.pit$table
```
### Relationship between altitude and type

|             | Res.Df | Df.diff | Dev     | Pr(>Dev)  |
|-------------|--------|---------|---------|-----------|
| (Intercept) | 38     | NA      | NA      | NA        |
| altitude    | 37     | 1       | 4916.371| 0.003344482 |
| type        | 36     | 1       | 2596.276| 0.003344482 |

```r
lepmvabund.anova2.pit <- anova(lep2.nb, p.uni = "unadjusted", resamp = "pit", nBoot = 999, show.time = TRUE)
lepmvabund.anova2.pit$table
```

### Thetford MBC dataset (Arthropoda OTUs)

|             | Res.Df | Df.diff | Dev     | Pr(>Dev)  |
|-------------|--------|---------|---------|-----------|
| (Intercept) | 38     | NA      | NA      | NA        |
| altitude    | 37     | 1       | 4916.371| 0.001001001 |
| type        | 36     | 1       | 2596.276| 0.001001001 |

```r
arthB.nb <- manyglm(arthmvabundB ~ arthropods.env$treatment, family = "binomial")
arthmvabundB.summ.pit <- summary(arthB.nb, test = "wald", p.uni = "adjusted", resamp = "pit", nBoot = 999, show.time = TRUE)
arthmvabundB.summ.pit$coefficients
pvaluesB = arthmvabundB.summ.pit$coefficients[c(1:8), 2]
pvaluesB.corr.fdr = p.adjust(pvaluesB, method = "fdr", n = length(pvaluesB))
pvaluesB.corr.fdr
```

|             | wald value | Pr(>wald) | sig | fdr-corrected p-values |
|-------------|------------|-----------|-----|------------------------|
| (Intercept) | 25.196541  | 0.001001  |     | 0.004004004             |
| arthropods.env$treatmentAgriPlough | 10.635074  | 0.01101101 | *   | 0.017617618             |
| arthropods.env$treatmentDisc      | 9.754609   | 0.1011011 |     | 0.115544116             |
STD dataset (ants, spiders, carabid beetles)

```
arth.nb <- manyglm(arthmvabund~arthropods.env$treatment, family="negative.binomial")
arthmvabund.summ.pit <- summary(arth.nb,test="wald",p.uni="adjusted",resamp="pit.trap", nBoot=999, show.time=TRUE)
arthmvabund.summ.pit$coefficients
pvalues=arthmvabund.summ.pit$coefficients[c(1:8),2]
pvalues.corr.fdr<-p.adjust(pvalues, method = "fdr", n = length(pvalues))
```

|                      | wald value | Pr(>wald) | sig | fdr-corrected p-values |
|----------------------|------------|-----------|-----|------------------------|
| (Intercept)          | 33.72605   | 0.01601602|     | 0.03803804             |
| arthropods.env$treatmentAgriPlough | 11.648017 | 0.01201201| *   | 0.03803804             |
| arthropods.env$treatmentDisc    | 9.464827  | 0.15315315|     | 0.16716717             |
| arthropods.env$treatmentForageHarvest | 10.020655 | 0.04104104|     | 0.06566567             |
| arthropods.env$treatmentForestPlough | 10.235845 | 0.11511512|     | 0.15348682             |
| arthropods.env$treatmentHeath   | 14.054133 | 0.01501502| *   | 0.03803804             |
| arthropods.env$treatmentSwipe   | 8.347123  | 0.16716717|     | 0.16716717             |
| arthropods.env$treatmentTurfStrip | 12.848741 | 0.01901902| *   | 0.03803804             |

Danum Valley

MBC dataset (Arthropoda OTUs)

```
MBC.nb <- manyglm(MBCmvabund~dbenv$Logcode, family="binomial")
MBCmvabund.summ5.pit <- summary(MBC.nb,test="wald",p.uni="adjusted",resamp="pit.trap", nBoot=999, show.time=TRUE)
MBCmvabund.summ5.pit$coefficients # 0L = unlogged, 1L = once-logged, 2L = twice-logged, aliased against 0L
```
```r
pvaluesB = MBCmvabund.summ5.pit$coefficients[c(1:3), 2]
pvaluesB.corr.fdr = p.adjust(pvaluesB, method = "fdr", n = length(pvaluesB))

| Wald value | Pr(>wald) | Sig | fdr-corrected p-values |
|------------|-----------|-----|------------------------|
| (Intercept)| 52.81046  | 0.001001 | 0.003003003 |
| dbenv$Logcode1L | 25.43188 | 0.38938939 | 0.389389389 |
| dbenv$Logcode2L | 28.54568 | 0.00900901 * | 0.013513514 |

### STD Birds dataset

birdnb <- manyglm(birdmvabund ~ birdsMNenv$Logcode, family = "negative.binomial")
birdmvabund.summ.pit <- summary(birdnb, test = "wald", p.uni = "adjusted", resamp = "pit.trap", nBoot = 999, show.time = TRUE)
birdmvabund.summ.pit$coefficients # 0L = unlogged, 1L = once-logged, 2L = twice-logged, aliased against 0L
pvalues = birdmvabund.summ.pit$coefficients[c(1:3), 2]
pvalues.corr.fdr = p.adjust(pvalues, method = "fdr", n = length(pvalues))

| Wald value | Pr(>wald) | Sig | fdr-corrected p-values |
|------------|-----------|-----|------------------------|
| (Intercept)| 12.308402 | 0.02402402 | 0.02402402 |
| birdsMNenv$Logcode1L | 9.02335 | 0.01801802 * | 0.02402402 |
| birdsMNenv$Logcode2L | 9.948525 | 0.01301301 * | 0.02402402 |

### STD Dung beetles dataset

dungbnb <- manyglm(dbmvabund ~ dbenv$Logcode, family = "negative.binomial")
dbmvabund.summ.pit <- summary(dungbnb, test = "wald", p.uni = "adjusted", resamp = "pit.trap", nBoot = 999, show.time = TRUE)
dbmvabund.summ.pit$coefficients # 0L = unlogged, 1L = once-logged, 2L = twice-logged, aliased against 0L
pvalues = dbmvabund.summ.pit$coefficients[c(1:3), 2]

| Wald value | Pr(>wald) | Sig | fdr-corrected p-values |
|------------|-----------|-----|------------------------|
| (Intercept)| 71.21403  | 0.001001 | 0.003003003 |
| dbenv$Logcode1L | 10.1894 | 0.02602603 * | 0.026026026 |
| dbenv$Logcode2L | 11.8707 | 0.00500501 * | 0.007507508 |
```
STD Ant dataset

ant.nb <- manyglm(antmvabund~antenv$Logcode, family="negative.binomial")
antmvabund.summ.pit <- summary(ant.nb, test="wald", p.uni="adjusted", resamp="pit.trap", nBoot=999, show.time=TRUE)
antmvabund.summ.pit$coefficients # 0L = unlogged, 1L = once-logged, 2L = twice-logged, aliased against 0L
pvalues=antmvabund.summ.pit$coefficients[c(1:3),2]
pvalues.corr.fdr <- p.adjust(pvalues, method = "fdr", n = length(pvalues))

|                | wald value | Pr(>wald) | sig      | fdr-corrected p-values |
|----------------|------------|-----------|----------|------------------------|
| (Intercept)    | 19.45681   | 0.02602603| 0.07807808|                        |
| antenv$Logcode1L | 11.23531   | 0.20220220| 0.20220220|                        |
| antenv$Logcode2L | 11.61287   | 0.06806807| 0.10210210|                        |
Supporting Information 4. Scatterplot correlations of species richness between the MBC and STD datasets across the eight Thetford restoration treatment levels.
Supporting Information 5. Histogram outputs from the RSW2 Monte Carlo R script. The red line indicates the number of observed matches between the MBC and STD datasets, at each budget level (15 = $15,000, 30 = $30,000, ...).
References

1. Basset Y., Novotny V., Miller S. & Springate N.E. (1998). Assessing the impact of forest disturbance on tropical invertebrates: some comments. *J Appl Ecol*, 35, 461-466.

2. Blacker N.C. & Collingwood C.A. (2002). Some significant new records of ants (Hymenoptera: Formicidae) from the Salisbury area, South Wiltshire, England, with a key to the British species of Lasius. *British Journal of Entomology and Natural History*, 15, 25-46.

3. Bolton B. & Collingwood A. (1975). *Hymenoptera: Formicidae. Handbooks for the Identification of British Insects 6.3c*. Royal Entomological Society London.

4. Ding Y., Ren G., Shi G., Gong P., Zheng X., Zhai P., Zhang D., Zhao Z., Wang S., Wang H., Luo Y., Chen D., Gao X. & Dai X. (2007). China’s national assessment report of climate change (I): climate change in China and its future trend. *Advances in Climate Change Research*, 3, 1-5.

5. Dolman P.M., Panter C.J. & Mossman H.L. (2012). The Biodiversity Audit Approach challenges regional priorities and identifies a mismatch in conservation. *J Appl Ecol*, 49, 986-997.

6. EC (1992). The Habitats Directive. In. European Community Brussels.

7. Edwards D.P., Hodgson J.A., Hamer K.C., Mitchell S.L., Ahmad A.H., Cornell S.J. & Wilcove D.S. (2010). Wildlife-friendly oil palm plantations fail to protect biodiversity effectively. *Cons Letts*, 3, 236-242.

8. Edwards D.P., Larsen T.H., Docherty T.D.S., Ansell F.A., Hsu W.W., Derhé M.A., Hamer K.C. & Wilcove D.S. (2011). Degraded lands worth protecting: The biological importance of Southeast Asia’s repeatedly logged forests. *Proc R Soc Lond B*, 278, 82-90.

9. Fisher B., Edwards D.P., Larsen T.H., Ansell F.A., Hsu W.W., Roberts C.S. & Wilcove D.S. (2011). Cost-effective conservation: calculating biodiversity and logging trade-offs in Southeast Asia. *Cons Letts*, 4, 443-450.

10. IPCC (2007). Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. In: (eds. Core Writing Team, Pachauri RK & Reisinger A) Geneva, Switzerland, p. 104 pp.

11. Kitching R.L., Boulter S.L., Vickerman G., Laidlaw M., Hurley K.L. & Grimbacher P.S. (2005). *The Comparative Assessment of Arthropod and Tree Biodiversity in Old-World Rainforests*. 2nd edn. Rainforest CRC & Earthwatch Institute, James Cook University, Cairns.

12. Larsen T.H. & Forsyth A. (2005). Trap spacing and transect design for dung beetle biodiversity studies. *Biotropica*, 37, 322-325.

13. Luff M.L. (2007). *The Carabidae (ground beetles) of Britain and Ireland*. Royal Entomological Society, St Albans.

14. Oksanen J., Blanchet F.G., Kindt R., Legendre P., Minchin R.B., O'Hara R.B., Simpson G.L., Solymos P., Stevens M.H.H. & Wagner H. (2012). vegan: Community Ecology Package. In: http://CRAN.R-project.org/package=vegan.

15. Pedley S.M., Franco A.M.A., Pankhurst T. & Dolman P.M. (2013). Physical disturbance enhances ecological networks for heathland biota: A multiple taxa experiment. *Biol Conserv*, 160, 173-182.

16. Roberts M.J. (1987). *The Spiders of Great Britain and Ireland*. Harley Books, Colchester.

17. Roberts M.J. (1996). *Spiders of Britain and Northern Europe*. HarperCollins Publishers Ltd, London.

18. Skinner G.J. & Allen G.W. (1996). *Ants (Naturalists’ Handbooks 24)* Richmond Publishing Co. Ltd, Slough.

19. Wilson E.O. (1987). The little things that run the world (the importance and conservation of invertebrates). *Conserv Biol*, 1, 344-346.

20. Woodcock P., Edwards D.P., Fayle T.M., Newton R.J., Khen C.V., Bottrell S.H. & Hamer K.C. (2011). The conservation value of South East Asia's highly degraded forests: evidence from leaf-litter ants. *Philos Trans R Soc Lond B*, 366, 3256-3264.

21. Yang Y., Tian K., Hao J., Pei S. & Yang Y. (2004). Biodiversity and biodiversity conservation in Yunnan, China. *Biodiversity and Conservation*, 13, 813-826.
Young S.S. & Wang Z.-J. (1989). Comparison of secondary and primary forests in the Ailao Shan region of Yunnan, China. *For Ecol Manag*, 28, 281-300.