Journal of Animal Ecology

BRITISH ECOLOGICAL SOCIETY

Land-use change alters the mechanisms assembling rainforest mammal communities in Borneo

Journal:	*Journal of Animal Ecology*
Manuscript ID	Draft
Manuscript Type:	Research Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Wearn, Oliver; Imperial College London, ; Zoological Society of London, Institute of Zoology Carbone, Chris; Zoological Society of London, Rowcliffe, J; Zoological Society of London, Institute of Zoology Pfeifer, Marion; Newcastle University, School of Biology Bernard, Henry; Universiti Malaysia Sabah, Institute for Tropical Biology and Conservation Ewers, Robert ; Imperial College London, Division of Biology
Key-words:	β-diversity, community assembly, environmental filtering, land-use change, mammals, oil palm agriculture, selective logging, species co-occurrence

Land-use change alters the mechanisms assembling rainforest mammal communities in Borneo

Submitted as a Research Article to Journal of Animal Ecology

Authors: Oliver R. Wearn[1,2*], Chris Carbone[1], J. Marcus Rowcliffe[1], Marion Pfeifer[2,3], Henry Bernard[4], Robert M. Ewers[2]

[1]Institute of Zoology, Zoological Society of London, Regent's Park, London NW1 4RY, UK

[2]Department of Life Sciences, Imperial College London, Silwood Park, Ascot SL5 7PY, UK

[3]School of Biology, Newcastle University, Newcastle NE1 7RU, UK

[4]Institute for Tropical Biology and Conservation, Universiti Malaysia Sabah, Jalan UMS, 88400 Kota Kinabalu, Sabah, Malaysia

*Corresponding author: Wearn, O. R. (oliver.wearn@ioz.ac.uk)

Abstract

1. The assembly of species communities at local scales is thought to be driven by environmental filtering, species interactions, and spatial processes such as dispersal limitation. Little is known about how the relative balance of these drivers of community assembly changes along environmental gradients, especially man-made environmental gradients associated with land-use change.

2. Using concurrent camera- and live-trapping, we investigated the local-scale assembly of mammal communities along a gradient of land-use intensity (old-growth forest, logged forest and oil palm plantations) in Borneo. We hypothesised that increasing land-use intensity would lead to an increasing dominance of environmental control over spatial processes in community assembly. Additionally, we hypothesised that competitive interactions among species might reduce in concert with declines in α-diversity (previously documented) along the land-use gradient.

3. To test our first hypothesis, we partitioned community variance into the fractions explained by environmental and spatial variables. To test our second hypothesis, we used probabilistic models of expected species co-occurrence patterns, in particular focussing on the prevalence of spatial avoidance between species. Spatial avoidance might indicate competition, but might also be due to divergent habitat preferences.

4. We found patterns that are consistent with a shift in the fundamental mechanics governing local community assembly. In support of our first hypothesis, the importance of spatial processes (dispersal limitation and fine-scale patterns of home-ranging) appeared to decrease from low to high intensity land-uses, whilst environmental control increased in importance

(in particular due to fine-scale habitat structure). Support for our second hypothesis was weak: whilst we found that the prevalence of spatial avoidance decreased along the land-use gradient, in particular between congeneric species pairs most likely to be in competition, few instances of spatial avoidance were detected in any land-use, and most were likely due to divergent habitat preferences.

5. The widespread changes in land-use occurring in the tropics might be altering not just the biodiversity found in landscapes, but also the fundamental mechanics governing the local assembly of communities. A better understanding of these mechanics, for a range of taxa, could underpin more effective conservation and management of threatened tropical landscapes.

Keywords: β-diversity, community assembly, environmental filtering, land-use change, mammals, oil palm agriculture, selective logging, species co-occurrence.

INTRODUCTION

An understanding of how species assemble into communities is a central goal of contemporary ecology (HilleRisLambers *et al.* 2012). Significant advances have been made since the turn of the century, with new conceptual and mathematical models of the mechanisms of assembly, including better integration of scale (Leibold *et al.* 2004), niche differentiation (Tilman 2004), dispersal (Gravel *et al.* 2006), and community drift (Hubbell 2001). At the same time, natural habitats across the globe, and in particular in the tropics, have been subjected to unprecedented rates of clearance and disturbance (Asner *et al.* 2009), but we know very little about the drivers of community assembly in these novel, man-made systems. In turn, this means we have a poor capacity to predict the consequences of land-use change on diversity, and to devise useful management interventions that deal directly with the fundamental mechanisms that create and maintain local-scale diversity in man-made landscapes.

Two contrasting drivers of community assembly are generally thought to be dominant at the local scale: niche assembly and dispersal assembly. Niche assembly mechanisms have a century-long pedigree in ecology (e.g. Grinnell 1917; Elton 1927) and involve selection of species according to their fundamental environmental niche (the “abiotic filter”), as well as small-scale interactions with competitors, mutualists and consumers (the “biotic filter”). Dispersal assembly refers to the stochastic assembly of a local community by dispersal, i.e. by the movement of organisms across space (Vellend 2010). Local communities are said to be “dispersal limited” whenever immigration is restricted and they are therefore isolated to some degree. Although dispersal was recognised in some of the earliest models of community assembly (e.g. Macarthur & Wilson 1967), it has since become associated with the unified neutral theory (Hubbell 2001), in which dispersal is conceptually from the metacommunity

(Leibold *et al.* 2004). Contemporary assembly theory recognises that niche and dispersal assembly are not mutually exclusive and that both may operate concurrently (Mutshinda & O'Hara 2011). The challenge, therefore, is to identify the relative importance of these two assembly mechanisms, and under what circumstances the balance might be altered.

Land-use change represents the principal threat to biodiversity in the tropics (Laurance, Sayer & Cassman 2014), and a vast body of literature has accumulated on changes in diversity along man-made environmental gradients. From this, we know that α-diversity (site-level species richness within a land-use) often declines along gradients of land-use intensity (Gibson *et al.* 2011), but β-diversity (site-to-site variation in community composition within a land-use) does not appear to change in a consistent direction (Dormann *et al.* 2007; Newbold *et al.* 2016). For tropical forests which are selectively logged or converted to agriculture, β-diversity may increase (Berry *et al.* 2008), decrease (Kitching *et al.* 2013; Solar *et al.* 2015), or show a grain-dependent response (Wearn *et al.* 2016). Whilst an understanding of β-diversity patterns can inform reserve design, little insight is gained about the underlying drivers of assembly. Indeed the same patterns in β-diversity can be produced by vastly different drivers of assembly (Myers *et al.* 2013). The drivers of assembly matter in the case of management because they might inform what steps are appropriate to restore biodiversity in a given degraded area. For example, if environmental control is the dominant driver of assembly, then steps to restore habitat quality may be important. Alternatively, if dispersal limitation dominates assembly then restoring landscape connectivity might be more important.

For a better understanding of the impacts of land-use change on species communities, it will be necessary to uncover the dominant drivers of assembly along gradients of land-use intensity. One hypothesis is that increasing levels of disturbance along gradients of land-use intensity may lead to the breakdown of spatial structure – created by dispersal limitation and broad gradients in, for example, climate or elevation – in the occurrence of species. At the same time, environmental control may assume a more dominant role in assembly as land-use intensity increases. For example, logging greatly increases the frequency and area of edge habitat in forests, a micro-habitat which forest species often show strong responses to (Pfeifer *et al.* 2017). The importance of environmental control may be even more pronounced in plantation habitats, which drastically differ from the forests that species are adapted to – in terms of structure, resources and microclimate – and are often subject to intense management (Styring *et al.* 2011; Luskin & Potts 2011).

It might be expected that the biotic filter, in particular the strength of species interactions, will also be affected by land-use change (Tylianakis *et al.* 2008). Reductions in α-diversity along gradients of land-use intensity may be hypothesised to weaken competitive interactions and, all else being equal, may lead to communities which are more assembled by neutral processes (Weiher *et al.* 2011). In tropical forest, logging and conversion to plantation habitats has previously been shown to reduce α-diversity in a range of taxonomic groups (e.g. Gibson *et al.* 2011), but the potential for a concomitant weakening in species interactions relative to intact habitat remains poorly explored.

The last decade has seen the development of new analytical tools to investigate the mechanisms of community assembly (e.g. Chase & Myers 2011; Dray *et al.* 2012). In particular, β-diversity can be dissected into its environmental and spatial components, using canonical ordination (Dray *et al.* 2012). The importance of niche assembly can be inferred from the explanatory power of environmental variables thought to control the occurrence of species, whilst the importance of spatially-contagious processes such as dispersal and home-ranging can be inferred from the explanatory power of surrogate spatial variables. β-diversity left unexplained by environmental control and space is likely due to random community drift, as well as any unmeasured environmental variables (which are not structured in space) and measurement error (Legendre *et al.* 2009). Although species interactions are generally thought to operate at smaller spatial scales than environmental filtering (Weiher *et al.* 2011), the two components to niche assembly are difficult to separate on the basis of spatial modelling alone (Kraft *et al.* 2015). Competitive interactions can however be expected to leave a signature of negative co-occurrence patterns between pairs of species (Veech 2006). Competition is especially likely among species which share more traits, such as congeneric species (assuming phylogenetic conservatism in traits). Negative co-occurrences, though, are also consistent with divergent habitat preferences, and therefore sound knowledge of a study system is needed in order for useful inferences to be made. In order to conclusively demonstrate competition, experimental work is needed.

Recent technological advances have opened up the possibility of collecting community-wide biodiversity data at higher spatial and temporal resolutions than has been possible before (Turner 2014), even for highly mobile taxa, such as mammals. This is an important development, since the majority of studies investigating community assembly have focussed on sessile organisms (e.g. Vellend *et al.* 2007; De Cáceres *et al.* 2012; Siefert *et al.* 2013; Myers *et al.* 2013). High mobility likely has a number of effects on community dynamics, for example by allowing species to more effectively partition themselves in space and time, and may therefore be expected to alter the dominant drivers of community assembly. It is now acknowledged that contemporary theory in community ecology must be confronted with empirical data from a wide a range of organisms in order to advance the field (Weiher *et al.* 2011).

Here we investigate the local-scale drivers assembling mammal communities along a gradient of land-use intensity which is ubiquitous in Southeast Asia: old-growth forest, logged forest and oil palm (*Elaeis guineensis*) plantations. Much of the remaining forest in Southeast Asia has been logged or degraded (Margono *et al.* 2014; Gaveau *et al.* 2014), and conversion to oil palm has been occurring at an unprecedented rate over recent decades (Wilcove *et al.* 2013). Mammals are highly mobile and rare, making them a challenge to sample. To overcome this, we used concurrent networks of camera traps and live traps to sample almost the entire terrestrial mammal assemblage, and expended much higher sampling efforts than would be typical for other taxonomic groups. We have previously shown that mammal communities in our study sites exhibit a decline in α-diversity along the land-use gradient, and are assembled in a significantly non-random manner (Wearn *et al.* 2016). In this study, we test two specific hypotheses about the drivers of community assembly: 1) that increased land-use intensity results in an increasing dominance of environmental control over spatial processes, and 2) that reductions in α-richness along a gradient of land-use intensity result in a reduced role for competitive interactions.

MATERIALS AND METHODS

Sampling design across the land-use gradient

We sampled mammals along a gradient of land-use intensity in Sabah, Malaysian Borneo, taking advantage of the experimental design of the Stability of Altered Forest Ecosystems (SAFE) Project (Ewers *et al.* 2011). The gradient consists of old-growth forest within the Maliau Basin Conservation Area, repeatedly-logged forest within the Kalabakan Forest

Reserve and two adjacent oil palm plantations straddling the Kalabakan Forest Reserve boundary (Wearn *et al.* 2016).

We employed a clustered hierarchical sampling design, with individual sampling points clustered together into 1.75 ha plots, and 3 to 6 plots in turn clustered into blocks (Fig. 1). This multi-scale approach allowed for the investigation of fine-scale drivers of assembly, such as competition between species, whilst also allowing for the investigation of larger-scale gradients in community composition within a study site. Plots consisted of 48 potential sampling points, separated by 23 m, of which a random subset were chosen for sampling (mean number of points per plot = 14). Separation distances between plots (170 to 290 m) and between blocks (0.6 to 3 km) were similar across the land-uses. The SAFE Project has been deliberately designed to minimise confounding factors across the land-use gradient, including latitude, slope and elevation (Ewers *et al.* 2011), and this applied equally to our sampling design for mammals.

Across the study sites, 430 points were sampled using both camera-trapping and live-trapping. These were nested within 31 plots and 8 blocks (9 plots in 3 blocks for old-growth forest; 16 plots in 3 blocks for logged forest, and 6 plots in 2 blocks for oil palm). We excluded 12 points which had been camera-trapped for less than seven days, giving a total sampling effort of 9,430 live trap nights and 19,116 camera trap nights (after correcting for camera failures). The sampling intensity was similar across land-uses (mean trap nights per sampling point: 60 in old-growth forest; 78 in logged forest, and 58 in oil palm).

Mammal community sampling

Medium and large mammals were sampled using camera traps (Reconyx HC500), deployed strictly within 5 m of each randomly chosen sampling point (following methods previously outlined in Wearn *et al.* 2013). Our deployment of cameras at random locations, not just along trails, is a departure from traditional camera-trapping protocols, and uniquely allows us to uncover fine-scale patterns of β-diversity in medium and large mammals, driven by species habitat-use and occupancy. Since camera traps are continuous-time detectors, we considered photographic capture events to be independent if they a) contained different individuals or b) were separated by > 12 hours, which matched the approximate minimum separation between live trap events. Camera traps were active between May 2011 and April 2014, during which most plots (28 of 31) were sampled multiple times (mean effort per plot = 617 trap nights).

Small mammal trapping was conducted at the plot level, with two locally-made steel-mesh traps (18 x 10-13 x 28 cm) placed near ground level (0 - 1.5 m) within 10 m of each of the 48 sampling points and baited with oil palm fruit (see Wearn *et al.* 2016 for further details on the trapping and handling protocols). Here we only use data from the 418 points which were also sufficiently sampled using camera traps. Each session consisted of seven consecutive trapping days and some plots (15 of 31) were sampled for multiple sessions over the course of the study (mean effort per plot in this study = 304 trap nights). Trapping was carried out between May 2011 and July 2014, during which there were no major mast-fruiting events.

Environmental and spatial components of β-diversity across land-use

We dissected mammal β-diversity within each land-use into its environmental and spatial components, using distance-based redundancy analysis (Peres-Neto *et al.* 2006), hereafter RDA. This involves partitioning community variance into the fractions explained by environmental and spatial variables, using the model $Y = f[E + S] + R$, where Y is the community response matrix, E and S are matrices of environmental and spatial covariates, respectively, and R is a matrix of residuals (Dray *et al.* 2012).

Environmental variables were composed of three sets: 1) fine-scale habitat structure variables (canopy closure, canopy height, mean tree diameter-at-breast height (DBH), tree density, a habitat disturbance score, presence of a logging road, vegetation cover in four height strata, and deadwood volume); 2) topographical variables (slope, elevation and flow accumulation), and 3) landscape context within a 500 m radius (above-ground live tree biomass, forest cover and distance from the nearest forest). Polynomial functions of the DBH, habitat score and topographical variables were also included, to allow for non-linear responses. Full details of the environmental variables used are provided in Appendix S1 (see Supporting Information).

The spatial component of β-diversity was modelled using surrogate variables representing positive spatial correlation. This positive spatial correlation is the signature of the dispersal limitation and home-ranging we were interested in. We emphasise that we are not here investigating the large-scale dispersal limitation that determines species ranges, but rather the dispersal limitation (and ranging) occurring at small-scales that determines the stochastic, "lottery-like" occurrence patterns within a species range. We thought dispersal limitation would be weak for the largest and best-dispersing mammal species in our dataset, but that barriers to large-scale movement might nonetheless exist for these species (e.g. caused by steep mountain tops, rivers, large land-slides, or highly degraded areas that some species might be less likely to cross). We thought dispersal limitation might be more important for many of the medium and small-sized mammal species in our dataset, likely with limited capacity to disperse at will across our study landscapes, and also for the rarer species in our dataset (producing few propagules).

Surrogate variables to model the positive spatial correlation were based on the geographic coordinates of sampling points, and included 1) distance-based Moran's eigenvector maps (db-MEMs), and 2) trend-surfaces (Dray *et al.* 2012). MEMs represent a spectral decomposition of the distances among sampling points, and can potentially model spatial patterns at all scales perceivable in the sampling design (Borcard *et al.* 2004). We used db-MEMs to model fine-scale spatial patterns present within sampling blocks (Appendix S2: Fig. S1), creating a separate set of db-MEMs for each sampling block (Declerck *et al.* 2011). We filled the widest gaps between sampling points in a block with a small number of supplementary points (between 5 and 10 per block; 12% of the total), which were then removed before RDA modelling (Borcard *et al.* 2004). This will have caused a slight loss of orthogonality between db-MEM variables, but allowed for the modelling of spatial patterns down to a scale of between 67 and 76 m. We selected only the db-MEMs with positive eigenvalues, which model positive spatial correlations. We visualised the fine-scale spatial patterns of community variation as a function of geographic distance using multivariate Mantel correlograms (Legendre & Legendre 2012). To model broader-scale spatial patterns occurring across sampling blocks, we created 1st- and 2nd-order polynomial functions of the geographic coordinates (i.e. trend surfaces), to represent both linear and saddle-shaped patterns in space (Borcard, Legendre & Drapeau 1992).

For the medium- and large-sized mammal species in our dataset, within-block patterns (i.e. those modelled using db-MEMs) will primarily reflect habitat-use and home-ranging, but may also represent dispersal limitation for the smaller species. Broader-scale patterns (i.e. those modelled using trend surfaces) will primarily reflect dispersal limitation, as well as home-ranging in the most mobile species in the dataset, such as the Sunda clouded leopard (*Neofelis diardi*).

We applied the modified forward-selection method (Blanchet, Legendre & Borcard 2008), based on the adjusted coefficient of multiple determination (R^2_{adj}), to select a parsimonious set of environmental and spatial variables for the final RDA models in each land-use (see Appendix S2: Table S1 and Fig. S2). This was done separately for each of the three sets of environmental variables (habitat structure, topography and landscape context variables) and two sets of spatial variables (db-MEMs and trend-surface variables). For the final, parsimonious RDA models in each land-use, we quantified the variance explained by environment, space, and jointly by both environment and space, using variation partitioning (Peres-Neto *et al.* 2006). We tested for significant differences between land-uses in the variation explained (overall, by environment and by space) using a bootstrap procedure (Peres-Neto *et al.* 2006).

We also repeated the above steps (selection of parsimonious sets of variables, and variation partitioning) for the combined dataset across all three land-uses, fitting a ‘global’ RDA model. This allowed us to add land-use into the RDA as a categorical variable, and determine if there was important community variation across land-use which our environmental and spatial variables had failed to capture.

RDA models were fitted to Hellinger-transformed community matrices, with mammal detections per seven days summed over the camera- and live-trapping protocols. The Hellinger transformation divides by the total abundance at a site and then takes the square-root (therefore dampening the effect of extremely abundant species), and has previously been shown to have desirable properties in the context of RDA (Legendre & Gallagher 2001). Detection probabilities likely vary across species and across the two protocols, but we do not expect that this will substantially affect the relative sizes of the variance fractions explained by environment and space, although imperfect detection could inflate the unexplained variance in all cases. The Hellinger transformation is asymmetrical, meaning that species absences, which could be a result of non-detection rather than lack of presence, have a lower influence on the coefficient than presences. Our sampling points were not independent, in the sense that individual animals may have been detected across multiple points. This will mean that the precision associated with our environmental RDA models will likely be inflated, but we here focus on the variation explained, rather than on the results of specific hypothesis tests using these models (which would require unbiased estimates of uncertainty). We also note that variation partitioning represents a very effective way of assessing environmental control after controlling for spatial non-independence (Peres-Neto & Legendre 2010), as we have done.

Species co-occurrence

We investigated co-occurrence patterns among species within each land-use using probabilistic models (Veech 2013). These models use the hypergeometric distribution to calculate the probability that two species spatially co-occur either less or more often than expected based on their mean incidence probabilities (Griffith, Veech & Marsh 2014). Using the observed co-occurrence frequencies and a specified alpha level (in this case, $\alpha = 0.05$), species co-occurrences were classified as significantly positive or negative, or occurring at random. Random co-occurrence patterns can potentially be generated both by genuine non-association between species or by a lack of statistical power. We had a relatively large number of sites within each land-use (56 to 213) and high sampling effort, with simulations suggesting that in this case models would typically have very high power to detect deviations from random co-occurrence as small as 5% of the total number of sites. We did not analyse species pairs with expected co-occurrence frequencies < 1 (following Veech 2013).

We were particularly interested in negative species co-occurrences, which will primarily be due to either 1) competition driving spatial niche separation, or 2) divergent habitat preferences. We are unable to conclusively determine which of these mechanisms is responsible for negative co-occurrences, and capture both possible explanations under the term "spatial avoidance" (e.g. Davis *et al.* 2018). We did, however, investigate if the number of congeneric pairs among the negative co-occurrences was greater than expected by chance, which would lend more support to competition as an explanation. Spatial co-occurrence analyses are unable to detect temporal niche separation and we are here focussed only on the spatial assembly of communities.

All analyses were done in R version 3.1.0 (R Development Core Team 2014), using the additional packages vegan 2.0-10 (Oksanen *et al.* 2013), PCNM 2.1-2 (Legendre *et al.* 2013) and cooccur 1.0 (Griffith *et al.* 2014).

RESULTS

We obtained 1,237 captures of 20 species from live-trapping, and 10,464 photo-captures of 56 species from camera-trapping, giving a total of 61 mammal species detected (15 species were captured using both methods). Overall trapping rates (live-trapping and camera-trapping combined) were similar for old-growth and logged forest (0.47 and 0.38 captures per trap night, respectively), but lower in oil palm (0.19).

How much community variation is explained by environment and space?

In old-growth forest, space explained significantly more of the community variation than environment (95% CI of the difference: 0.05 – 0.14, $p < 0.001$). Most of the variation explained by environmental variables was spatially-structured (90%; Fig. 2), and only a small, albeit significant, fraction of the variation was explained by pure environmental control ($F_{(12, 112)} = 1.34$, $p < 0.01$). Mammal communities in logged forest and oil palm were less structured in space than old-growth forest (Fig. 2; 95% CI of the variance explained for old-growth forest: 0.26 – 0.37; 95% CI for logged forest: 0.05 – 0.15; 95% CI for oil palm: -0.05 – 0.24). In logged forest, just 14% (95% CI: 10 – 19%) of the total variance was explained by environmental and spatial variables, much lower than for old-growth forest (33%, 95% CI: 27 – 39%) and oil palm (30%, 95% CI: 15 – 46%). However, the independent environmental and spatial components were still significant (environment: $F_{(10, 179)} = 1.90$, $p < 0.001$; space: $F_{(16, 179)} = 1.91$, $p < 0.001$). The variance explained by environmental and spatial variables was not significantly different for logged forest communities (95% CI of the difference: -0.06 – 0.01, $p = 0.17$). In oil palm, 86% of the explained variation was related to environmental control (79% independently) and the spatial component was significantly smaller (95% CI of the difference: 0.02 – 0.30, $p = 0.02$). Despite the contrasting variation partitioning results across the land-use gradient, total community variation (i.e. site-to-site β-diversity within a land-use) was broadly similar across the land-uses: $\mathrm{Var}(Y_{\text{Old-growth}}) = 0.57$; $\mathrm{Var}(Y_{\text{Logged}}) = 0.58$, and $\mathrm{Var}(Y_{\text{Oil palm}}) = 0.53$.

The relative importance of habitat structure, topography and landscape context

The relative importance of each set of environmental variables for structuring communities within land-uses differed markedly (Fig. 2). Fine-scale habitat structure was an important independent component in all land-uses (old-growth forest: $F_{(6, 143)} = 2.35$, $p < 0.001$; logged forest: $F_{(7, 195)} = 2.83$, $p < 0.001$; oil palm: $F_{(6, 45)} = 3.08$, $p < 0.001$). At coarser-scales, landscape context was important in old-growth forest, both in combination with fine-scale habitat structure and also independently ($F_{(6, 143)} = 2.35$, $p < 0.001$), but was only important in oil palm in combination with topography ($F_{(4, 45)} = 1.72$, $p = 0.01$) and not independently ($F_{(3, 45)} = 1.29$, $p = 0.16$). Landscape context was found to be entirely unimportant for logged forest communities ($F_{(1, 204)} = 1.25$, $p = 0.31$). Topography was important mostly in combination with other variables, but was also important in its own right in old-growth forest ($F_{(5, 143)} = 3.03$, $p < 0.001$).

Broad- and fine-scale structuring of communities in space

Broad- and fine-scale spatial processes were equally important in old-growth forest (95% CI of the difference: -0.04 – 0.08, p = 0.43) and oil palm (95% CI of the difference: -0.11 – 0.13, p = 0.90), but there was a trend for the dominance of fine-scale space in logged forest (95% CI of the difference: -0.06 – 0.00, p = 0.09). We calculated *post-hoc* that broad-scale space, independent of environment, explained 6%, 3% and 1% of the community variation in old-growth forest, logged forest and oil palm, respectively. The same portions for fine-scale space were 9%, 4% and 3%, respectively. Mantel correlations between community distances and geographic distances were weak in all cases ($r < 0.12$). However, in old growth forest, there was a signal of positive correlation at distances less than 100 m, which was apparently absent in the two anthropogenic land-uses (Appendix S2: Fig. S3).

Is the effect of land-use on community assembly explained by our covariates?

The RDA model for the combined dataset across all three land-uses (global model) showed that land-use explained just 1.8% of the variance independently (Appendix S2: Fig. S4), suggesting that our measured environmental variables successfully captured the environmental gradient. In addition, the ordination tri-plot of the environmental component of this model showed clear separation of sampling points across the three land-uses (Fig. 3), with the first and second RDA axes explaining 7.2% and 3.9% of the community variation, respectively. Environmental and spatial variables in the global model, each explaining a similar portion of the community variance (95% CI of the difference: -0.03 – 0.03, $p = 0.776$), together explained a total of 33% of the variance (Appendix S2: Fig. S4).

Co-occurrence patterns across the land-use gradient

Co-occurrence patterns became increasingly random along the land-use gradient, with 26%, 14% and 7% of analysed species pairs classified as non-random in old-growth forest ($n = 392$ species pairs), logged forest ($n = 627$ pairs) and oil palm ($n = 43$ pairs), respectively. Of these non-random associations, most were positive (Fig. 4) and only in old-growth forest was there any substantial evidence of negative co-occurrences, i.e. spatial avoidance (13% of non-random associations). However, it should be noted that Type II errors are a possibility for oil palm, in which 25% of associations classified as random represented deviations of > 3% of the total number of sites (this was the case for just 1% and 0% for old-growth and logged forest, respectively).

Three of the 13 instances of spatial avoidance in old-growth forest were of congeneric pairs. A simple randomisation test in which 13 species pairs were selected at random 10,000 times, showed that this pattern was highly unlikely ($p < 0.001$, 95% CI of the number of expected congeners: 0 – 1). There were too few instances of spatial avoidance in the other land-uses to run this test.

DISCUSSION

We found patterns that are consistent with a shift under land-use change in the fundamental mechanics governing local community assembly. We found that across the land-use gradient: 1) environmental control was less important (old-growth forest), equally as important (logged forest) and more important (oil palm) than spatial processes in explaining community variation; 2) mammal communities were increasingly less spatially-structured (the share of the explained variation within the pure spatial component, and in the component jointly explained with habitat, increased), and 3) the prevalence of spatial avoidance between species decreased, particular among congeneric species pairs. Results (1) and (2) lend support to the first of our initial hypotheses, that increases in land-use intensity lead to an increasing dominance of environmental control over spatial processes (spatially-correlated dispersal and ranging) in assembly. Result (3) is consistent with the second of our initial hypotheses, that competitive interactions play a reduced role with increasing land-use intensity, but the support is weak given that few instances of spatial avoidance were detected in any land-use. In addition, more evidence is required, for example from experimental work, in order to conclusively demonstrate competition. Our study adds to a small body of research which has begun to investigate the impacts of land-use change on community assembly (Vellend *et al.* 2007; Pakeman 2011; Ding *et al.* 2012; Döbert *et al.* 2017; Danneyrolles, Arseneault & Bergeron 2018). To our knowledge, this is the first time that land-use change has been associated with altering the drivers of community assembly in non-sessile organisms.

Increased environmental control under land-use change

The patterns in our data suggest that the importance of environmental filtering compared to spatial processes was stronger along the land-use gradient. These results support previous work on plant communities indicating an increasing role for environmental control under disturbance, as indicated by increased phylogenetic and functional trait clustering (Pakeman

2011; Ding *et al.* 2012) and by variance partitioning analyses similar to those undertaken here (Danneyrolles *et al.* 2018). Contrary to this trend, one study of understorey plant communities (also conducted within the SAFE experimental design) found evidence that environmental control at fine scales decreased due to selective logging, as indicated by increased phylogenetic and functional dispersion (Döbert *et al.* 2017). Comparisons across studies are hampered by differences in spatial scale and analytical methods, but more studies of non-sessile taxa are especially needed in order to make broader generalisations about the impacts of land-use change on assembly.

Of the three land-uses in our study, environmental control was apparently strongest in oil palm (both in absolute terms and relative to spatial processes), particularly due to variation in habitat at fine scales (Fig. 2). Oil palm plantations often exhibit substantial heterogeneity (Luskin & Potts 2011), for example in the age and height of palms, the amount of scrub vegetation in the understorey and the presence of access roads, and indeed we found these to be important fine-scale environmental filters (Table S3). This echoes findings more generally in plantations (e.g. Greenberg *et al.* 1997; Peh *et al.* 2006; Styring *et al.* 2011; Lantschner, Rusch & Hayes 2012) and may indeed be a robust pattern in these highly-modified habitats. This sensitivity to fine-scale environmental control could be because species in plantations are persisting closer to their biological tolerances (e.g. for food resources, microclimate or cover from predators). Our finding that mammal communities in oil palm were strongly assembled by environmental control suggests that there is at least the potential for ameliorating the impacts of oil palm by altering management practices. For example, the ground and understorey layer, typically cleared in plantations using herbicides, could be maintained in some areas (Foster *et al.* 2014).

Spatial processes assembling communities across the land-use gradient

Old-growth communities were much more spatially-structured than communities in logged forest and oil palm (Fig. 2). The spatial processes involved in assembling old-growth communities at the local scale are likely to have included broad-scale dispersal limitation between sampling blocks, and finer-scale home-ranging movements. The latter is the most likely explanation for the stronger positive correlation in old-growth communities seen at distances below 100 m (Appendix S2: Fig. S3). One hypothesis for the weaker positive correlation signal evident in the two anthropogenic habitats might be that these habitats favour species with high dispersal capacity, or facultative behaviours which mean that individuals range more widely, in order to access resources more effectively in a heterogeneous environment. In old-growth forest, there was also a large portion of the variance explained by space and environment together. This fraction likely represents a combination of spatially-structured environmental control, for example along topographical gradients (which we have shown were important in old-growth forest), and spatial processes which happen to correlate with environmental gradients.

Does land-use change weaken the biotic filter?

We found that the prevalence of spatial avoidance among species pairs decreased along the gradient of land-use intensity, which might suggest a weakening of the biotic filter. However, we consider that divergent habitat preferences, rather than competition, is the likely explanation for most of the instances of spatial avoidance. Indeed, 69% of the spatial avoidance pairs involved the greater mouse-deer (*Tragulus napu*), which our environmental RDA identified as an old-growth forest specialist (Fig. 3). We did detect three instances of spatial avoidance among congeneric species (Appendix S2: Fig. S5), a result which we found to be highly unlikely by chance. It is possible, therefore, that competition is driving spatial niche separation in two morphologically very similar and abundant *Maxomys* rats (*M. surifer* and *M. rajah*), the greater and lesser (*T. kanchil*) mouse-deer, and the thick-spined and Malay porcupines (*Hystrix crassispinis* and *H. brachyura*). Overall though, co-occurrence analyses provided only weak evidence that competition weakens along the gradient of land-use intensity.

Few instances of spatial avoidance were detected in any land-use, suggesting that competition is unlikely to be a dominant assembly process in our study system. This would perhaps be expected for a trophically-diverse mammal assemblage, able to avoid competition through resource-partitioning. It is also in agreement with a broader set of studies on a range of taxa, all suggesting the primacy of environmental control over competition in determining the occurrence and abundance of species (Veech 2006, 2013; Houlahan *et al.* 2007; Mutshinda, O'Hara & Woiwod 2009). However, it is also possible that species are segregated in time or vertically in space (e.g. in the forest canopy), neither of which our co-occurrence analyses would detect. We are also not suggesting that competition is absent, only that it does not appear to be a dominant force in assembly on the basis of our co-occurrence analyses.

Random assembly under selective logging

The overall explanatory power of our models for logged forest communities was low compared to the other land-uses. We have already noted the reduced spatial structure of community variation in logged forest, but environmental variables also explained only a small amount of the variation. In particular, from old-growth to logged forest, there was a marked reduction in the variance explained by the local landscape context and topographical variables. This could be due to a destruction by logging of the heterogeneity in forest structure and species composition which is ordinarily present across local landscapes and along topographical gradients in old-growth forests of the region (Newbery *et al.* 1996). We note that our logged forest sites, which had been subjected to a very high intensity of logging (Appendix S1), were uniformly dominated by a single pioneer tree species, *Macaranga pearsonii* (~10% of basal area; M. Khoo, personal communication). An alternative, but less likely, explanation for the reduced ability of our models to explain community variation in logged forest would be that the environmental variables we used were inappropriate for logged forest. We did specifically design the variables to capture the structural dimensions affected by logging (Appendix S1), but we did not have fine-scale information available to us on tree species composition or direct measures of resource abundance, such as the availability of fruit. It remains the case, however, that spatial variables performed poorly in logged forest, indicating that, if there were key environmental variables missing from the analysis, they were not spatially-structured.

With a reduction in the importance of environment and space, this may mean that random community drift plays a more important role in logged forest compared to old-growth forest, which could represent a pervasive, but as-yet-undocumented, legacy of logging. A similar finding of randomly-assembled communities was found in the same landscape for understorey plants, particularly at high logging intensities (Döbert *et al.* 2017). If this is a general finding across the region, this could call into question the prospects for large-scale, unassisted restoration of ecosystem processes in Southeast Asian logged-over forests, and may mean that more active management interventions, targeted at species of particular conservation concern, may be a more appropriate management target (Lamb, Erskine & Parrotta 2005).

CONCLUSION

The widespread creation of man-made environmental gradients in the tropics may be leading to novel mechanics governing the local assembly of communities. This is a poorly-appreciated facet of land-use change that needs further investigation: that it might alter not just the biodiversity found in a landscape (the focus of most studies), but the mechanisms which create and maintain this biodiversity at the local scale. More broadly, we found patterns that support the niche-neutrality continuum model of community assembly, in which dispersal, drift and environmental control all combine variously to create communities at the local scale (Mutshinda & O'Hara 2011). A better understanding of this continuum model along environmental gradients, which are arguably now a defining feature of tropical forest landscapes, could underpin more effective conservation and management of biodiversity in these highly-threatened systems.

ACKNOWLEDGEMENTS

We are grateful to Yayasan Sabah, Benta Wawasan, Sabah Softwoods, the Sabah Forestry Department and the Maliau Basin Management Committee for allowing access to field sites, and to the Economic Planning Unit of Malaysia and Sabah Biodiversity Council for providing research permission. Fieldwork would not have been possible without the efforts of a great number of people and institutions, in particular the SAFE Project field staff, the Royal Society South East Asia Rainforest Research Programme, Glen Reynolds, Edgar Turner, MinSheng Khoo, Leah Findlay, Jeremy Cusack, Robin Loveridge, Matthew Holmes, Faye Thompson, Jack Thorley and Jessica Haysom. The European Space Agency provided satellite imagery under a Category 1 user agreement to MP (C1P 13735).

AUTHOR CONTRIBUTIONS

ORW, CC, JMR and RME conceived and designed the study. ORW collected and analysed the biodiversity data. MP processed the satellite data and created the models of above-ground biomass. All authors contributed to the writing of the paper and gave final approval for publication.

DATA ACCESSIBILITY

The data used in this study are available at: https://doi.org/10.5281/zenodo.877217.

REFERENCES

Asner, G.P., Rudel, T.K., Aide, T.M., Defries, R. & Emerson, R. (2009) A contemporary assessment of change in humid tropical forests. *Conservation Biology*, **23**, 1386–1395.

Berry, N.J., Phillips, O.L., Ong, R.C. & Hamer, K.C. (2008) Impacts of selective logging on tree diversity across a rainforest landscape: the importance of spatial scale. *Landscape Ecology*, **23**, 915–929.

Blanchet, F., Legendre, P. & Borcard, D. (2008) Forward selection of explanatory variables. *Ecology*, **89**, 2623–2632.

Borcard, D., Legendre, P., Avois-Jacquet, C. & Tuomisto, H. (2004) Dissecting the spatial structure of ecological data at multiple scales. *Ecology*, **85**, 1826–1832.

Borcard, D., Legendre, P. & Drapeau, P. (1992) Partialling out the spatial component of ecological variation. *Ecology*, **73**, 1045–1055.

De Cáceres, M., Legendre, P., Valencia, R., Cao, M., Chang, L.-W., Chuyong, G., Condit, R., Hao, Z., Hsieh, C.-F., Hubbell, S., Kenfack, D., Ma, K., Mi, X., Supardi Noor, M.N., Kassim, A.R., Ren, H., Su, S.-H., Sun, I.-F., Thomas, D., Ye, W. & He, F. (2012) The variation of tree beta diversity across a global network of forest plots. *Global Ecology and Biogeography*, **21**, 1191–1202.

Chase, J.M. & Myers, J.A. (2011) Disentangling the importance of ecological niches from stochastic processes across scales. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, **366**, 2351–63.

Danneyrolles, V., Arseneault, D. & Bergeron, Y. (2018) Anthropogenic disturbances strengthened tree community-environment relationships at the temperate-boreal interface. *Landscape Ecology*, **33**, 213–224.

Davis, C.L., Rich, L.N., Farris, Z.J., Kelly, M.J., Di Bitetti, M.S., Blanco, Y. Di, Albanesi, S., Farhadinia, M.S., Gholikhani, N., Hamel, S., Harmsen, B.J., Wultsch, C., Kane, M.D., Martins, Q., Murphy, A.J., Steenweg, R., Sunarto, S., Taktehrani, A., Thapa, K., Tucker, J.M., Whittington, J., Widodo, F.A., Yoccoz, N.G. & Miller, D.A.W. (2018) Ecological correlates of the spatial co-occurrence of sympatric mammalian carnivores worldwide. *Ecology Letters*, **21**, 1401–1412.

Declerck, S.A.J., Coronel, J.S., Legendre, P. & Brendonck, L. (2011) Scale dependency of processes structuring metacommunities of cladocerans in temporary pools of High-Andes wetlands. *Ecography*, **34**, 296–305.

Ding, Y., Zang, R., Letcher, S.G., Liu, S. & He, F. (2012) Disturbance regime changes the trait distribution, phylogenetic structure and community assembly of tropical rain forests. *Oikos*, **121**, 1263–1270.

Döbert, T.F., Webber, B.L., Sugau, J.B., Dickinson, K.J.M. & Didham, R.K. (2017) Logging increases the functional and phylogenetic dispersion of understorey plant communities in tropical lowland rain forest. *Journal of Ecology*, **105**, 1235–1245.

Dormann, C.F., Schweiger, O., Augenstein, I., Bailey, D., Billeter, R., de Blust, G., DeFilippi, R., Frenzel, M., Hendrickx, F., Herzog, F., Klotz, S., Liira, J., Maelfait, J.-P., Schmidt, T., Speelmans, M., van Wingerden, W.K.R.E. & Zobel, M. (2007) Effects of landscape structure and land-use intensity on similarity of plant and animal communities. *Global Ecology and Biogeography*, **16**, 774–787.

Dray, S., Pélissier, R., Couteron, P., Fortin, M.-J., Legendre, P., Peres-Neto, P.R., Bellier, E., Bivand, R., Blanchet, F.G., De Caceres, M., Dufour, A.-B., Heegaard, M., Jombart, T., Munoz, F., Oksanen, J., Thioulouse, J. & Wagner, H.H. (2012) Community ecology in the age of multivariate multiscale spatial analysis. *Ecological Monographs*, **82**, 257–275.

Elton, C. (1927) *Animal Ecology*. Sedgwick and Jackson, London.

Ewers, R.M., Didham, R.K., Fahrig, L., Ferraz, G., Hector, A., Holt, R.D., Kapos, V., Reynolds, G., Sinun, W., Snaddon, J.L. & Turner, E.C. (2011) A large-scale forest fragmentation experiment: the Stability of Altered Forest Ecosystems Project. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **366**, 3292–3302.

Foster, W.A., Snaddon, J.L., Advento, A.D., Agung, A.A., Barclay, H., Caliman, J.P., Kumiawan, C., Kurz, D.J., Mann, D.J., Naim, M. & Pujianto, D.P. (2014) The Biodiversity and Ecosystem Function in Tropical Agriculture (BEFTA) project. *The Planter*, **90**, 581–591.

Gaveau, D.L.A., Sloan, S., Molidena, E., Yaen, H., Sheil, D., Abram, N.K., Ancrenaz, M., Nasi, R., Quinones, M., Wielaard, N. & Meijaard, E. (2014) Four decades of forest persistence, clearance and logging on Borneo. *Plos One*, **9**, e101654.

Gibson, L., Lee, T.M., Koh, L.P., Brook, B.W., Gardner, T.A., Barlow, J., Peres, C.A., Bradshaw, C.J.A., Laurance, W.F., Lovejoy, T.E. & Sodhi, N.S. (2011) Primary forests are irreplaceable for sustaining tropical biodiversity. *Nature*, **478**, 378–381.

Gravel, D., Canham, C.D., Beaudet, M. & Messier, C. (2006) Reconciling niche and neutrality: the continuum hypothesis. *Ecology Letters*, **9**, 399–409.

Greenberg, R., Bichier, P., Angon, A.C. & Reitsma, R. (1997) Bird populations in shade and sun coffee plantations in central Guatemala. *Conservation Biology*, **11**, 448–459.

Griffith, D.M., Veech, J.A. & Marsh, C.J. (2014) cooccur: Probabilistic species co-occurrence analysis in R. Available from http://cran.r-project.org/package=cooccur.

Grinnell, J. (1917) Field tests of theories concerning distributional control. *American Naturalist*, **51**, 115–128.

HilleRisLambers, J., Adler, P.B., Harpole, W.S., Levine, J.M. & Mayfield, M.M. (2012) Rethinking community assembly through the lens of coexistence theory. *Annual Review of Ecology, Evolution, and Systematics*, **43**, 227–248.

Houlahan, J.E., Currie, D.J., Cottenie, K., Cumming, G.S., Ernest, S.K.M., Findlay, C.S., Fuhlendorf, S.D., Gaedke, U., Legendre, P., Magnuson, J.J., McArdle, B.H., Muldavin, E.H., Noble, D., Russell, R., Stevens, R.D., Willis, T.J., Woiwod, I.P. & Wondzell, S.M. (2007) Compensatory dynamics are rare in natural ecological communities. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 3273–3277.

Hubbell, S.P. (2001) *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press.

Kitching, R.L., Ashton, L.A., Nakamura, A., Whitaker, T. & Khen, C.V. (2013) Distance-driven species turnover in Bornean rainforests: homogeneity and heterogeneity in primary and post-logging forests. *Ecography*, **36**, 675–682.

Kraft, N.J.B., Adler, P.B., Godoy, O., James, E.C., Fuller, S. & Levine, J.M. (2015) Community assembly, coexistence and the environmental filtering metaphor. *Functional Ecology*, **29**, 592–599.

Lamb, D., Erskine, P.D. & Parrotta, J.A. (2005) Restoration of degraded tropical forest landscapes. *Science*, **310**, 1628–32.

Lantschner, M.V., Rusch, V. & Hayes, J.P. (2012) Habitat use by carnivores at different spatial scales in a plantation forest landscape in Patagonia, Argentina. *Forest Ecology* *and Management*, **269**, 271–278.

Laurance, W.F., Sayer, J. & Cassman, K.G. (2014) Agricultural expansion and its impacts on tropical nature. *Trends in Ecology and Evolution*, **29**, 107–116.

Legendre, P., Borcard, D., Blanchet, F.G. & Dray, S. (2013) PCNM: MEM spatial eigenfunction and principal coordinate analyses. R package version 2.1-2. Available from http://R-Forge.R-project.org/projects/sedar/.

Legendre, P. & Gallagher, E. (2001) Ecologically meaningful transformations for ordination of species data. *Oecologia*, **129**, 271–280.

Legendre, P. & Legendre, L. (2012) *Numerical Ecology*, 3rd ed. Elsevier Science BV, Amsterdam.

Legendre, P., Mi, X., Ren, H., Ma, K., Yu, M., Sun, I.-F. & He, F. (2009) Partitioning beta diversity in a subtropical broad-leaved forest of China. *Ecology*, **90**, 663–74.

Leibold, M.A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J.M., Hoopes, M.F., Holt, R.D., Shurin, J.B., Law, R., Tilman, D., Loreau, M. & Gonzalez, A. (2004) The metacommunity concept: a framework for multi-scale community ecology. *Ecology Letters*, **7**, 601–613.

Luskin, M.S. & Potts, M.D. (2011) Microclimate and habitat heterogeneity through the oil palm lifecycle. *Basic and Applied Ecology*, **12**, 540–551.

Macarthur, R.H. & Wilson, E.O. (1967) *The Theory of Island Biogeography*. Princeton University Press.

Margono, B.A., Potapov, P. V., Turubanova, S., Stolle, F. & Hansen, M.C. (2014) Primary forest cover loss in Indonesia over 2000–2012. *Nature Climate Change*, **4**, 730–735.

Mutshinda, C.M. & O'Hara, R.B. (2011) Integrating the niche and neutral perspectives on community structure and dynamics. *Oecologia*, **166**, 241–51.

Mutshinda, C.M., O'Hara, R.B. & Woiwod, I.P. (2009) What drives community dynamics?

Proceedings of the Royal Society B: Biological Sciences, **276**, 2923–2929.

Myers, J.A., Chase, J.M., Jiménez, I., Jørgensen, P.M., Araujo-Murakami, A., Paniagua-Zambrana, N. & Seidel, R. (2013) Beta-diversity in temperate and tropical forests reflects dissimilar mechanisms of community assembly. *Ecology Letters*, **16**, 151–7.

Newbery, D.M., Campbell, E.J.F., Proctor, J. & Still, M.J. (1996) Primary lowland dipterocarp forest at Danum Valley, Sabah, Malaysia. Species composition and patterns in the understorey. *Vegetatio*, **122**, 193–220.

Newbold, T., Hudson, L.N., Hill, S.L.L., Contu, S., Gray, C.L., Scharlemann, J.P.W., Börger, L., Phillips, H.R.P., Sheil, D., Lysenko, I. & Purvis, A. (2016) Global patterns of terrestrial assemblage turnover within and among land uses. *Ecography*, 1–13.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Henry, M., Stevens, H. & Wagner, H. (2013) vegan: Community ecology package. R package version 2.0-10. Available from http://cran.r-project.org/package=vegan.

Pakeman, R.J. (2011) Functional diversity indices reveal the impacts of land use intensification on plant community assembly. *Journal of Ecology*, **99**, 1143–1151.

Peh, K.S.-H., Sodhi, N.S., de Jong, J., Sekercioglu, C.H., Yap, C. a.-M. & Lim, S.L.-H. (2006) Conservation value of degraded habitats for forest birds in southern Peninsular Malaysia. *Diversity and Distributions*, **12**, 572–581.

Peres-Neto, P.R. & Legendre, P. (2010) Estimating and controlling for spatial structure in the study of ecological communities. *Global Ecology and Biogeography*, **19**, 174–184.

Peres-Neto, P., Legendre, P., Dray, S. & Borcard, D. (2006) Variation partitioning of species data matrices: estimation and comparison of fractions. *Ecology*, **87**, 2614–2625.

Pfeifer, M., Lefebvre, V., Peres, C.A., Banks-Leite, C., Wearn, O.R., Marsh, C.J., Butchart, S.H.M., Arroyo-Rodríguez, V., Barlow, J., Cerezo, A., Cisneros, L., D'Cruze, N., Faria, D., Hadley, A., Harris, S.M., Klingbeil, B.T., Kormann, U., Lens, L., Medina-Rangel, G.F., Morante-Filho, J.C., Olivier, P., Peters, S.L., Pidgeon, A., Ribeiro, D.B., Scherber, C., Schneider-Maunoury, L., Struebig, M., Urbina-Cardona, N., Watling, J.I., Willig, M.R., Wood, E.M. & Ewers, R.M. (2017) Creation of forest edges has a global impact on forest vertebrates. *Nature*, **551**.

R Development Core Team. (2014) R: A language and environment for statistical computing.

Siefert, A., Ravenscroft, C., Weiser, M.D. & Swenson, N.G. (2013) Functional beta-diversity patterns reveal deterministic community assembly processes in eastern North American trees. *Global Ecology and Biogeography*, **22**, 682–691.

Solar, R.R. de C., Barlow, J., Ferreira, J., Berenguer, E., Lees, A.C., Thomson, J.R., Louzada, J., Maués, M., Moura, N.G., Oliveira, V.H.F., Chaul, J.C.M., Schoereder, J.H., Vieira, I.C.G., Mac Nally, R. & Gardner, T.A. (2015) How pervasive is biotic homogenization in human-modified tropical forest landscapes? *Ecology Letters*, **18**, 1108–1118.

Styring, A.R., Ragai, R., Unggang, J., Stuebing, R., Hosner, P.A. & Sheldon, F.H. (2011) Bird community assembly in Bornean industrial tree plantations: effects of forest age and structure. *Forest Ecology and Management*, **261**, 531–544.

Tilman, D. (2004) Niche tradeoffs, neutrality, and community structure: a stochastic theory of resource competition, invasion, and community assembly. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 10854–61.

Turner, W. (2014) Sensing biodiversity. *Science*, **346**, 301–302.

Tylianakis, J.M., Didham, R.K., Bascompte, J. & Wardle, D.A. (2008) Global change and species interactions in terrestrial ecosystems. *Ecology Letters*, **11**, 1351–1363.

Veech, J.A. (2006) A probability-based analysis of temporal and spatial co-occurrence in grassland birds. *Journal of Biogeography*, **33**, 2145–2153.

Veech, J.A. (2013) A probabilistic model for analysing species co-occurrence (ed P Peres-

Neto). *Global Ecology and Biogeography*, **22**, 252–260.

Vellend, M. (2010) Conceptual synthesis in community ecology. *The Quarterly Review of Biology*, **85**, 183–206.

Vellend, M., Verheyen, K., Flinn, K.M., Jacquemyn, H., Kolb, A., Van Calster, H., Peterken, G., Graae, B.J., Bellemare, J., Honnay, O., Brunet, J., Wulf, M., Gerhardt, F. & Hermy, M. (2007) Homogenization of forest plant communities and weakening of species-environment relationships via agricultural land use. *Journal of Ecology*, **95**, 565–573.

Wearn, O.R., Carbone, C., Rowcliffe, J.M., Bernard, H. & Ewers, R.M. (2016) Grain-dependent responses of mammalian diversity to land-use and the implications for conservation set-aside. *Ecological Applications*, **26**, 1409–1420.

Wearn, O.R., Rowcliffe, J.M., Carbone, C., Bernard, H. & Ewers, R.M. (2013) Assessing the status of wild felids in a highly-disturbed commercial forest reserve in Borneo and the implications for camera trap survey design. *Plos One*, **8**, e77598.

Weiher, E., Freund, D., Bunton, T., Stefanski, A., Lee, T. & Bentivenga, S. (2011) Advances, challenges and a developing synthesis of ecological community assembly theory. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **366**, 2403–13.

Wilcove, D.S., Giam, X., Edwards, D.P., Fisher, B. & Koh, L.P. (2013) Navjot’s nightmare revisited: logging, agriculture, and biodiversity in Southeast Asia. *Trends in Ecology & Evolution*, **28**, 531–540.

FIGURES

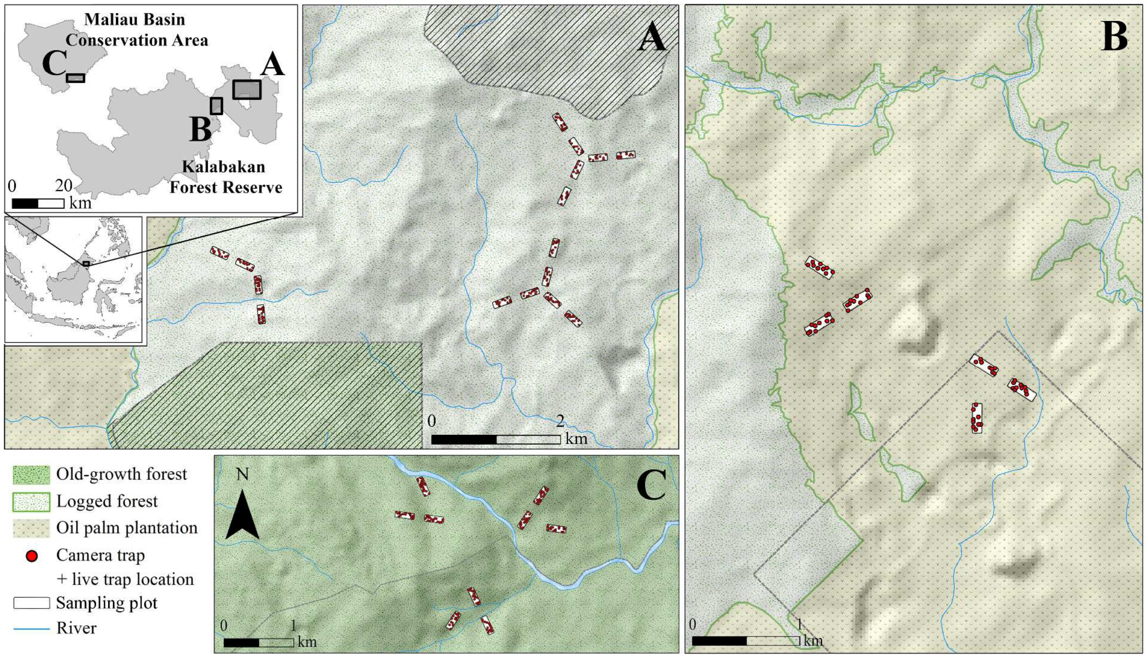

Figure 1. Sampling design across logged forest (A), oil palm (B) and old-growth forest (C), showing the sampling points (in red) sampled using both camera traps and live traps. Clusters of sampling plots, i.e. sampling blocks, were arranged in the same spatial configuration in old-growth forest and oil palm, but were arranged to coincide with future experimental forest fragments in logged forest. Separation between points, plots and blocks was nonetheless similar across land-uses. Shaded areas (in A) lie outside the Kalabakan Forest Reserve, consisting of a 2,200 ha Virgin Jungle Reserve (Brantian-Tatulit) to the south and an extensive (>1 million ha) area of logged forest to the north (Mount Louisa Forest Reserve and other connecting reserves).

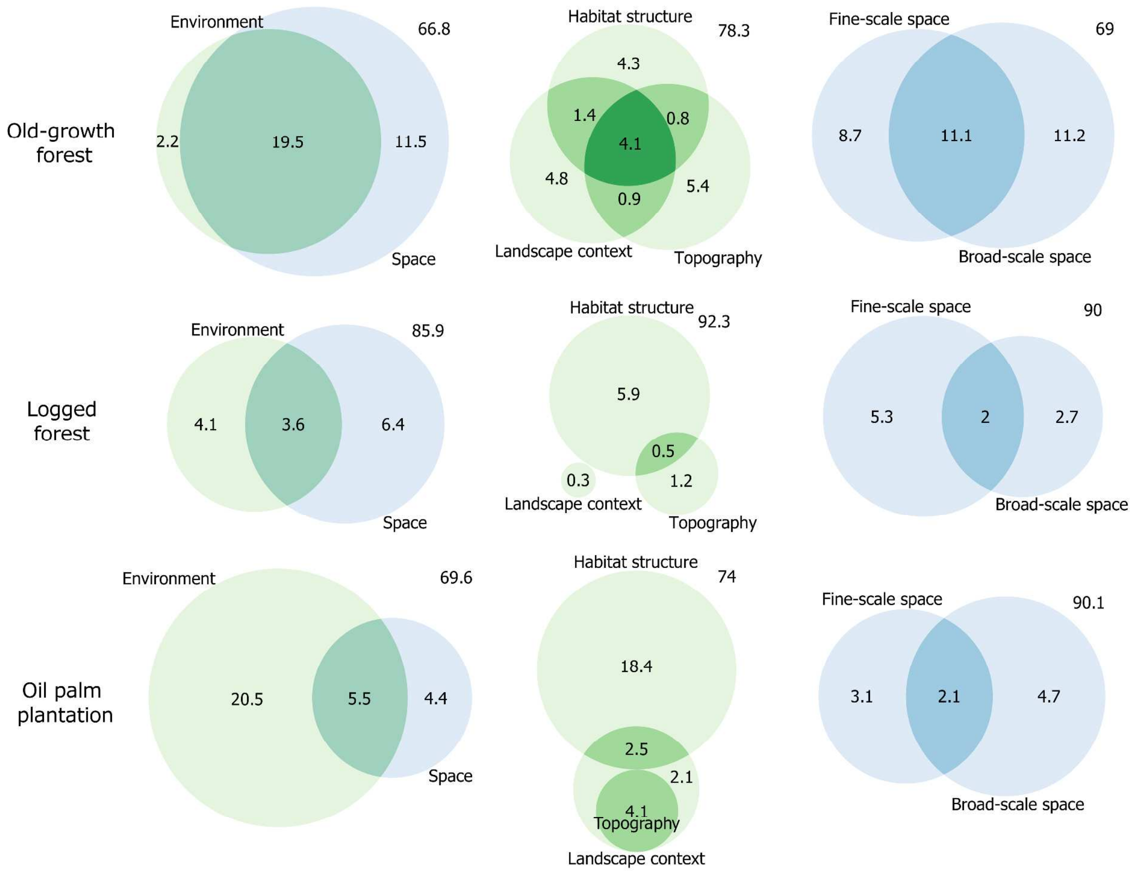

Figure 2. Variation partitioning of mammal community composition data across a tropical land-use gradient, represented with Euler diagrams. Community variation was partitioned using redundancy analyses (RDA) according to: three sets of environmental control variables (2nd column); broad- and fine-scale spatial processes (3rd column), and environmental control and space overall (1st column). A parsimonious set of environmental and spatial variables were chosen separately for each land-use. Percentage values represent the adjusted coefficient of multiple determination (R^2_{adj}) and values lying outside the area of the Euler diagrams represent the percentage variation left unexplained in each case. The landscape context RDA for oil palm could not be represented in full using a Euler diagram, and a small fraction (1.2%) shared between habitat structure and topography was omitted in order to allow for plotting.

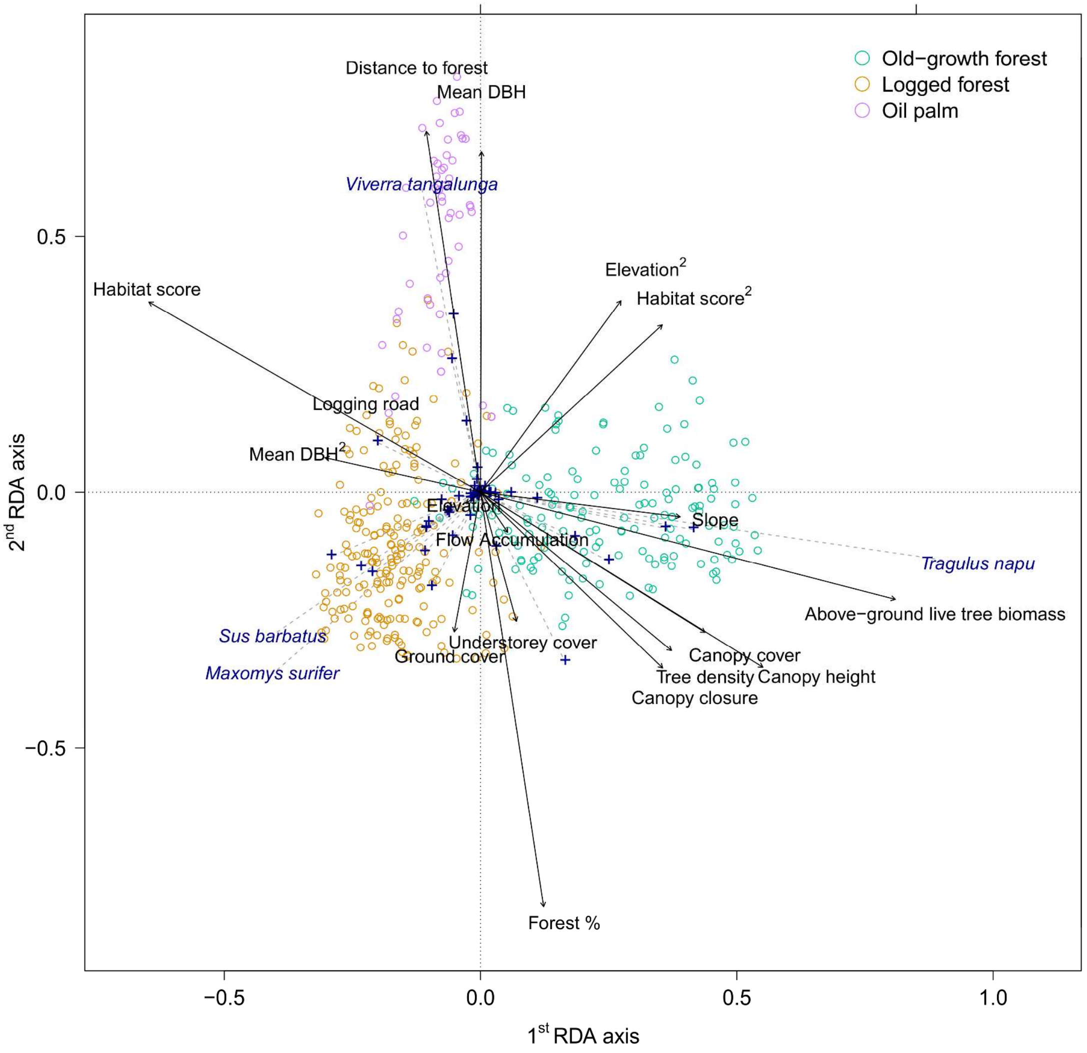

Figure 3. Ordination tri-plot depicting the position of sampling points (coloured circles), species (blue crosses) and environmental variables (black arrows) along the first two axes of a redundancy analysis (RDA) of the mammal community composition data (combined across land-use). Land-use was not included as a variable in this model. The first and second axes were both significant in permutation tests (1st axis: $F_{(1, 399)} = 65.1$, $p < 0.001$; 2nd axis: $F_{(1, 399)} = 35.2$, $p < 0.001$). Four species which were characteristic of old-growth forest (greater mouse-deer, *T. napu*), logged forest (bearded pig, *S. barbatus* and red spiny rat, *M. surifer*) and oil palm (Malay civet, *V. tangalunga*) are individually-named.

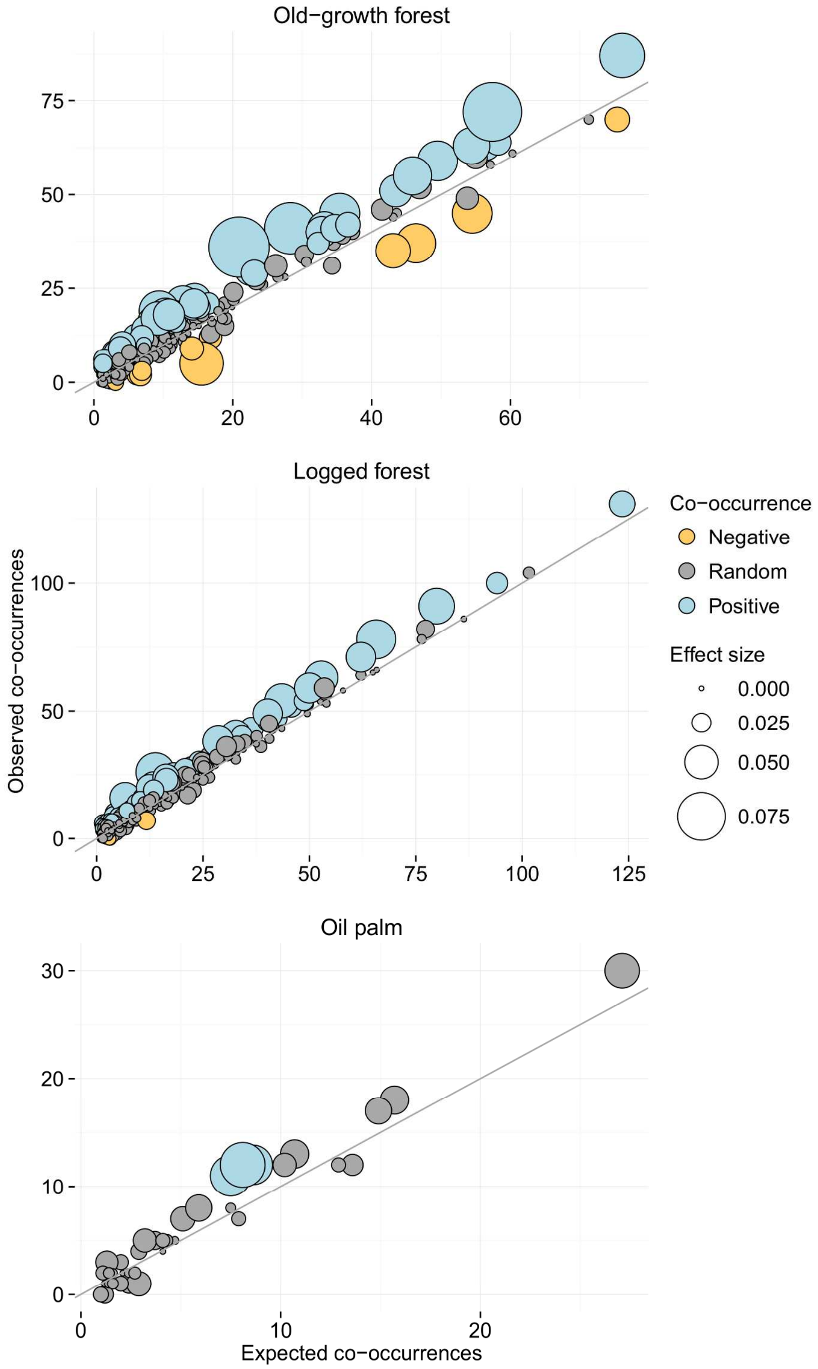

Figure 4. Observed and expected species co-occurrences between species for each land-use type. Significant positive co-occurrences (blue points) lie above the 1:1 line and significant negative co-occurrences (orange points) lie below it. Effects sizes were calculated by standardising the difference between observed and expected co-occurrences by the number of sampling points in each land-use.

Appendix S2 - Supplementary results

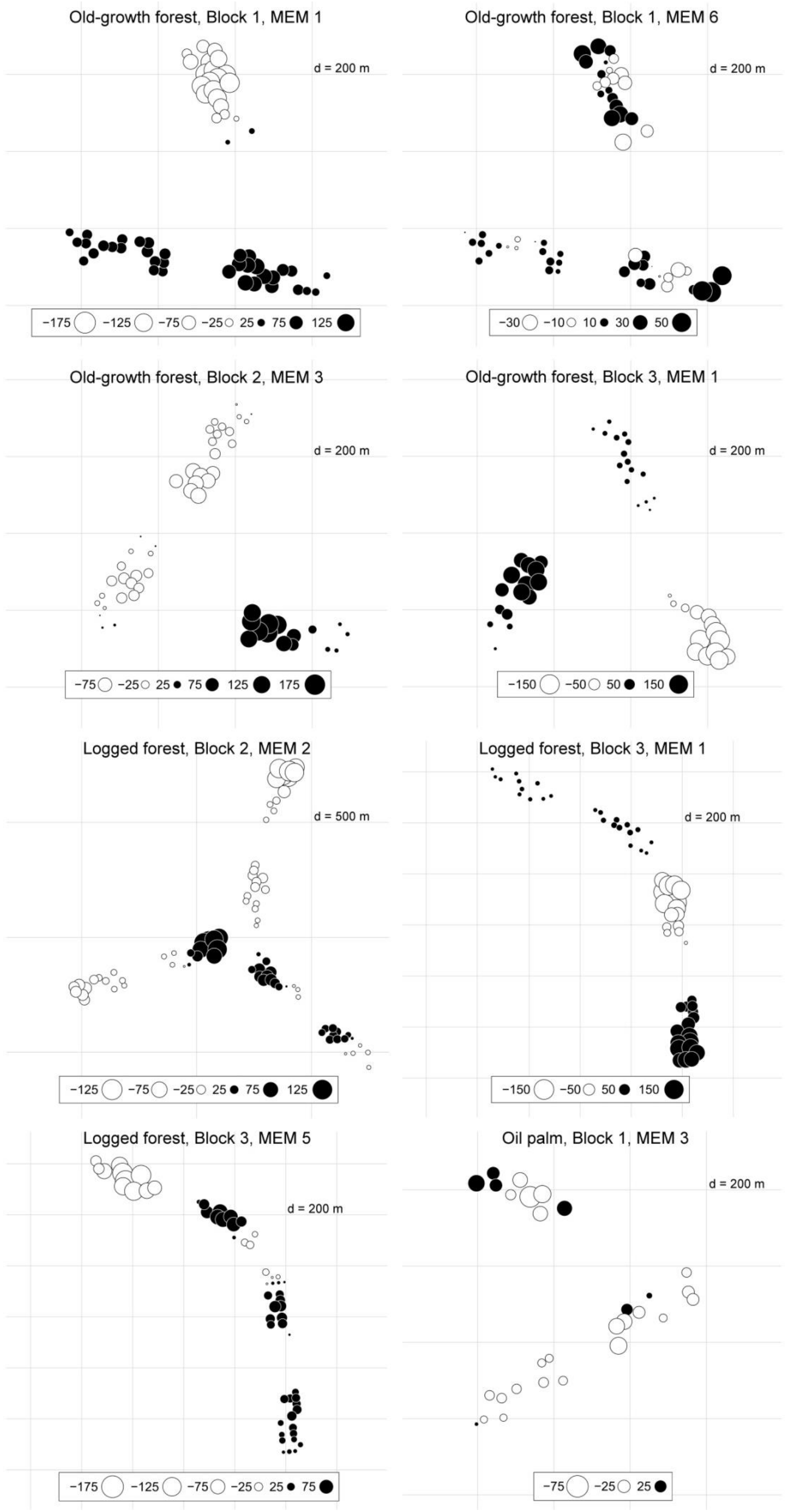

Figure S1. Example distance-based Moran’s eigenvector maps (db-MEMs), identified as significant in explaining β-diversity patterns using redundancy analysis.

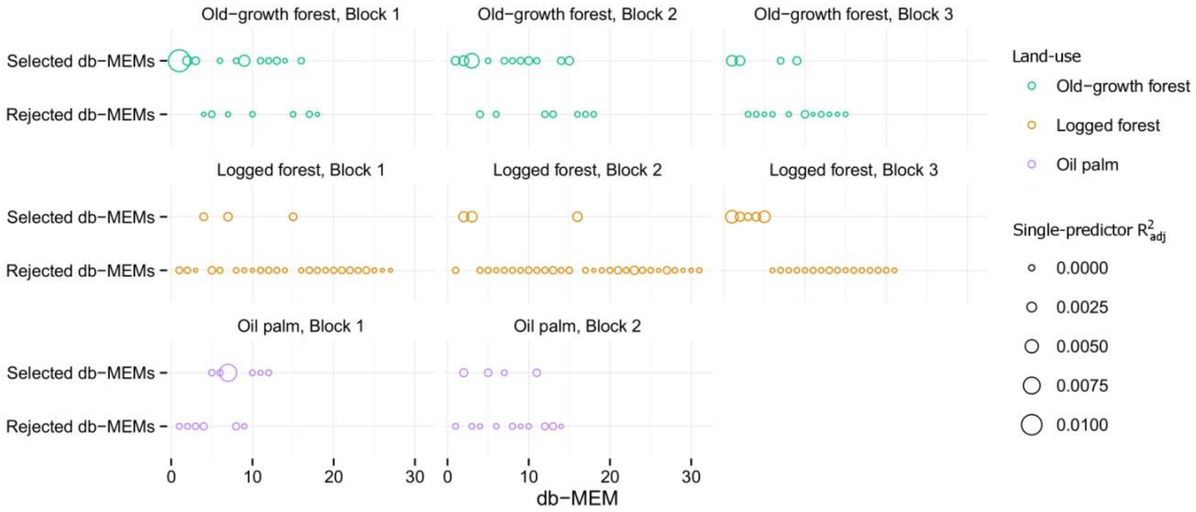

Figure S2. Distance-based Moran's eigenvector map (db-MEM) variables identified as significant in explaining β-diversity patterns using redundancy analysis. Variables are ordered from the coarsest to the finest scale within each block. Of the 156 db-MEMs generated, 47 were identified as significant in forward selection. Single-predictor R^2_{adj} values were derived from separate redundancy analyses for each db-MEM.

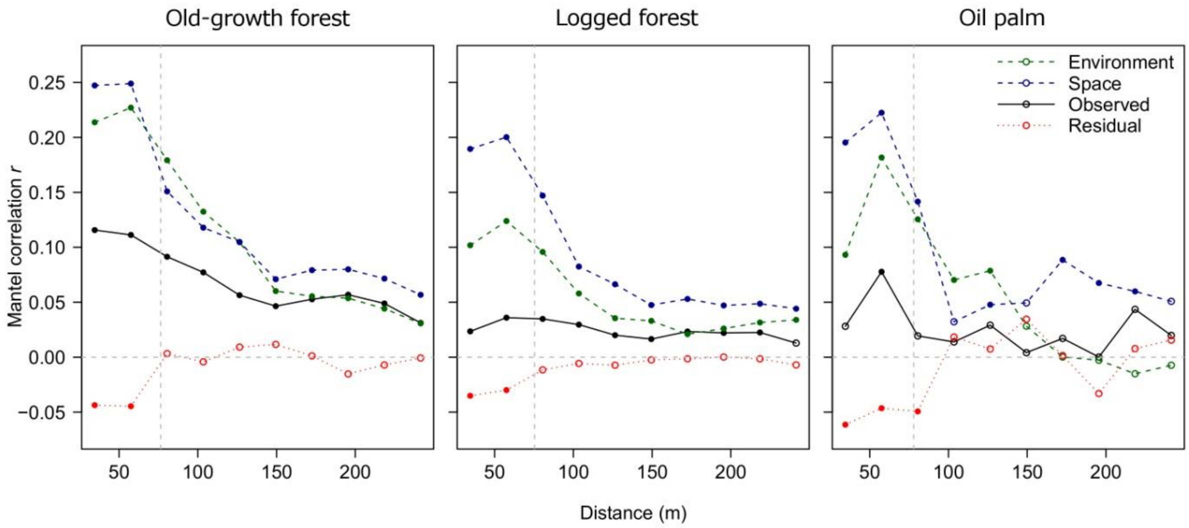

Figure S3. Mantel correlograms of observed community composition data (black) compared to fitted and residual values from redundancy analyses (RDA), across three land-use types. Fitted values in each case are shown for the environmental control RDA (green) and spatial RDA (blue). Residuals (red) are shown for the overall RDA for each land-use, containing both environmental and spatial variables. Filled points indicate significant correlations, as deduced using permutation tests and progressive Holm correction for multiple testing ($\alpha = 0.05$ for the first distance class and $\alpha < 0.05$ thereafter). Grey dashed vertical lines represent the finest scale resolvable by the distance-based Moran’s eigenvector map (db-MEM) variables in each land-use.

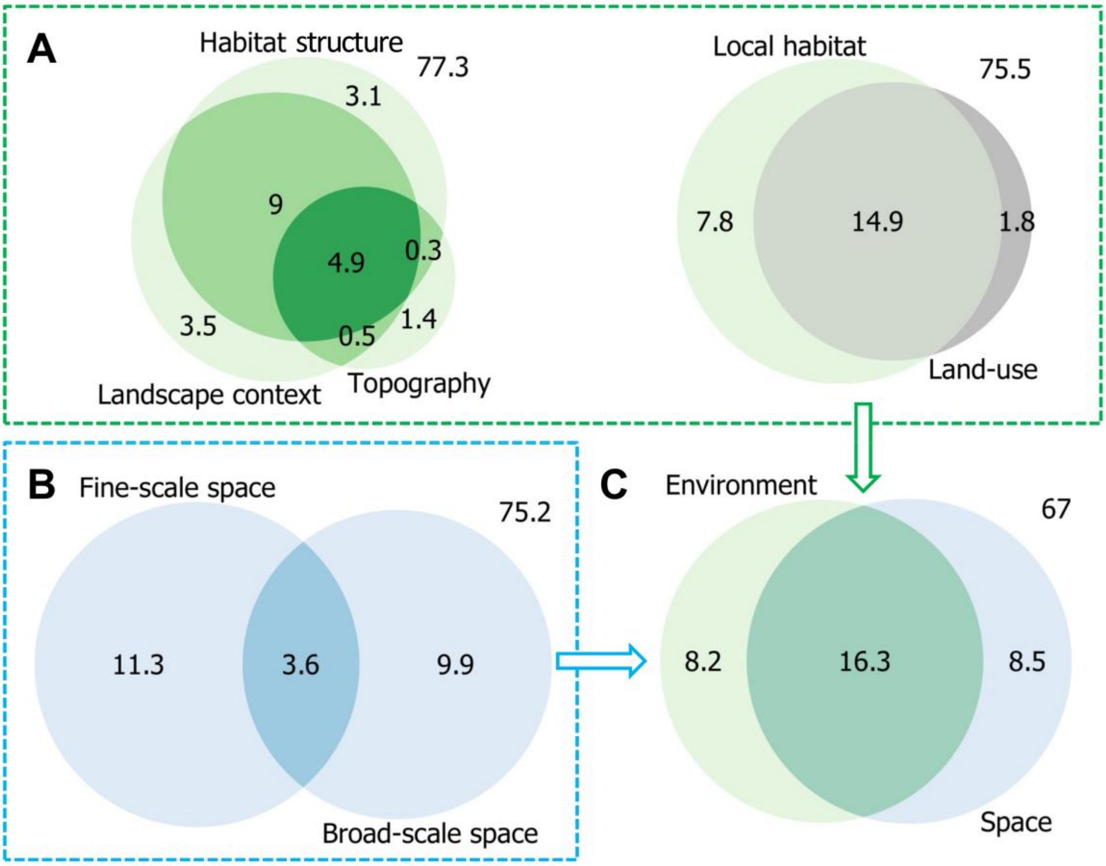

Figure S4. Variation partitioning of the mammal community composition data combined across land-use ('global' RDA model), separated into variation explained by a) local habitat characteristics only, and local habitat in combination with land-use, b) spatial processes (represented by spatial surrogate variables) and c) environmental control (local habitat and land-use) and spatial processes together. Percentage values represent the adjusted coefficient of multiple determination (R^2_{adj}) calculated using redundancy analyses. Values lying outside the area of the Euler diagrams represent the percentage variation left unexplained in each case.

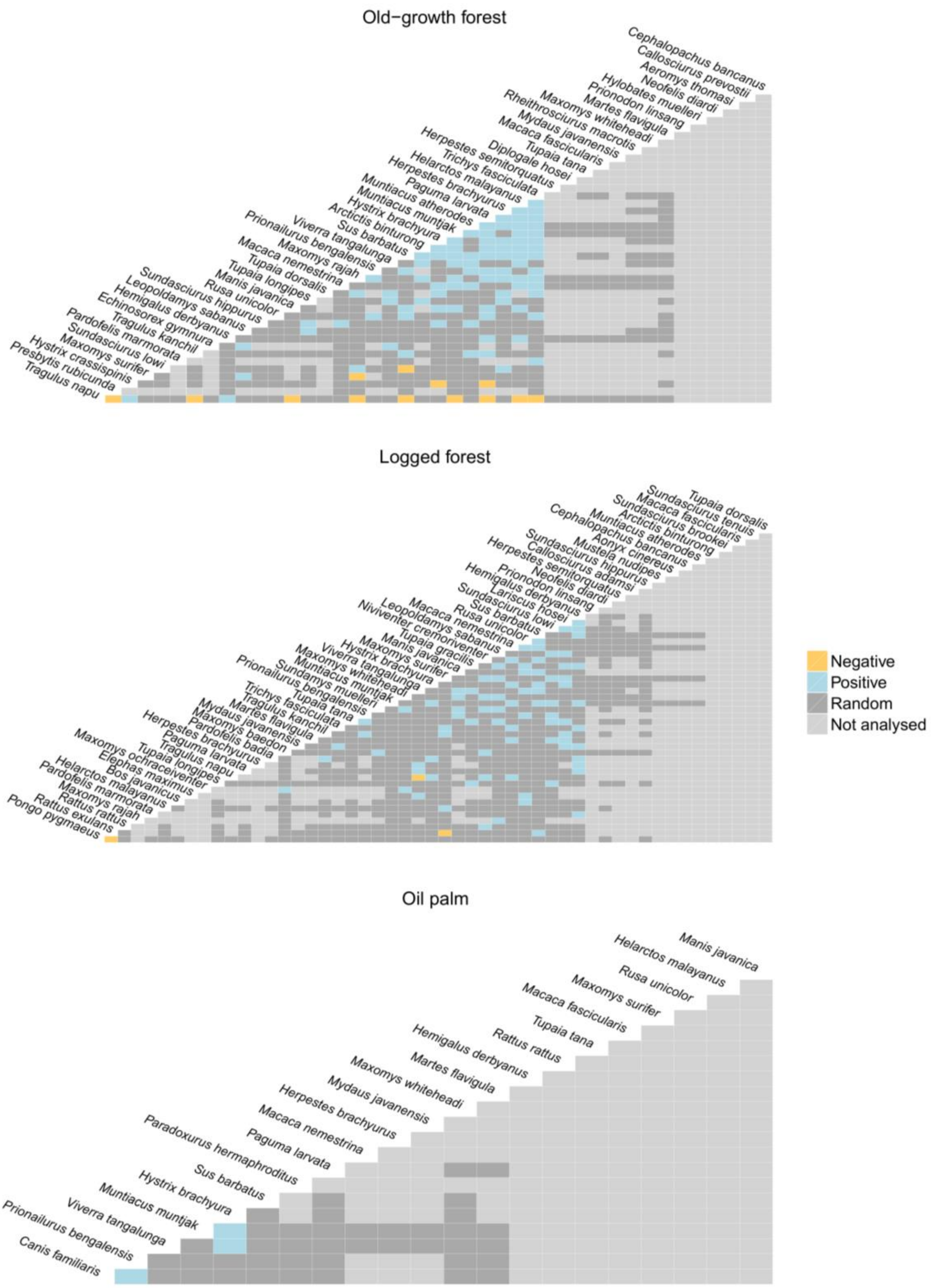

Figure S5. Species co-occurrence matrices across land-use types. Negative and positive co-occurrences indicate where two species occurred together significantly less or more often than expected by chance, respectively. Species pairs with expected co-occurrence frequencies < 1 were not analysed.

Table S1. Environmental variables chosen by a modified forward-selection procedure based on the adjusted coefficient of multiple determination (R^2_{adj}), for the global model and models specific to land-use types (old-growth forest, logged forest and oil palm) and species groups (large mammals and small mammals).

	Redundancy analysis (RDA) model			
Environmental variable	Global	Old-growth forest	Logged forest	Oil palm
Habitat structure[a]				
Ground cover	✓			
Understorey cover	✓		✓	
Midstorey cover				
Canopy cover	✓			
Canopy closure (arcsine-transformed %)	✓		✓	✓
Habitat score	✓	✓	✓	✓
Habitat score^2	✓		✓	
Logging road (binary)	✓			✓
Maximum tree height (m) (quadrat-based)[c]	✓	✓	✓	✓
Tree density (quadrat-based)[c]	✓	✓	✓	✓
Mean DBH (cm) (quadrat-based)[c]	✓	✓		
Mean DBH^2 (quadrat-based)[c]	✓	✓		✓
Deadwood volume (m^3) (quadrat-based)[c]		✓	✓	
Topography[b]				
Elevation (m)	✓	✓	✓	✓
Elevation^2	✓	✓	✓	
Flow accumulation	✓	✓		
Flow accumulation^2		✓		
Slope (degrees)	✓	✓		
Slope^2				
Local landscape context[b]				
Above-ground live tree biomass (Mg/ha) (500 m radius)	✓	✓		
Forest cover (%) (500 m radius)	✓			✓
Distance from forest (m)	✓			✓

[a] All habitat structure variables were measured in the field.

[b] All topographical and landscape variables were derived from satellite data.

[c] "Quadrat-based" variables were measured within 25 m^2 quadrats, located 0 - 75 m from sampling points, at a density of two quadrats per plot.

Appendix S1 – Detailed description of methods used to measure environmental variables

S1. 1. Field-based measurement of environmental variables

All habitat structure variables were quantified directly in the field. Eight variables were measured directly at sampling points by a single observer (O.R.W.), whilst a further five were measured in intensively-sampled 25 m^2 vegetation plots as part of ongoing monitoring at the Stability of Altered Forest Ecosystems (SAFE) Project (Ewers et al. 2011). We hereafter refer to these vegetation plots as 'quadrats' to distinguish them from the larger plots used to sample mammal communities. Two quadrats were located within each sampling plot, spaced evenly within the plot area and 150 m apart (centre-to-centre distance). Quadrats were mostly < 45 m from individual sampling points (range: 0 to 75 m).

Vegetation cover within a 5 m radius of each sampling point was estimated in four height strata (ground: 0-0.5 m; understorey: 0.5-3 m; midstorey: 3-20 m, and canopy: above 20 m) and placed into one of five broad classes for each stratum (1: 0-25%; 2: 25-50%; 3: 50-75%, and 4: 75-100%). Canopy cover was quantified using a spherical densiometer (Lemmon 1957), held at waist height and recorded as an average percentage across four measurements (one for each cardinal direction). Before analysis, this percentage was arcsine-transformed, owing to the strong negative skew apparent across all measurements. The intensity of habitat disturbance ("habitat score") within a 5 m radius of each point was recorded on a 1 to 5 scale (definitions provided in Table S1), following a similar methodology to previous studies (Ewers et al. 2011; Cusack et al. 2015). Higher habitat scores represent more disturbed sites, and sites within the oil palm plantation crop itself were never scored below 4, though areas in the margins of plantations were sometimes assigned scores of 3. We also noted whether the sampling point was on a logging road or not, due to the strong influence these features have on the occurrence of some species in our study sites (Wearn et al. 2013). Old logging roads, often following ridge-lines, were apparent due to their graded surface, the poor regeneration state of understorey vegetation, and the lack of canopy cover. Habitat scores and vegetation cover estimates were treated as ordered categorical variables during analysis. We also included in the analysis a 2nd-order polynomial term for habitat disturbance ("habitat score^2"), due to the non-linear responses to this variable we have previously found for some species (Cusack et al. 2015).

Table S1. Definition of habitat disturbance scale.

Habitat score	Definition
1	*Undisturbed forest*. Dominated by old-growth dipterocarps. High, continuous canopy with sparsely-vegetated understorey. Unlogged, with little recent disturbance evident.
2	*Disturbed forest*. Mostly pioneer tree species (typically *Macaranga* species), but some old-growth dipterocarp species may be present. Discontinuous canopy. Lower intensity of logging or natural disturbance.
3	*Heavily-disturbed forest*. High scrub or dense understorey layer (typically with vines and *Dinochloa* climbing bamboo species), with a low, heavily-broken canopy layer (< 20 m). Possibly some large isolated trees (> 20 m). Intensively-logged area or large gap disturbance.
4	*Herbaceous scrub*. Dominated by herbs (typically Zingiberaceae), vines and shrubs, with no trees > 3 m in height (except oil palm *Elaeis guineensis*). Typically representing secondary re-growth from clear-felling, or large gaps due to landslides.
5	*Open area*. Dominated by grasses and small shrubs (< 1 m in height). Typically on logging roads or old log landing areas.

Quadrat-based variables were measured during the course of long-term vegetation monitoring using RAINFOR protocols (Malhi et al. 2002; Turner et al. 2012). This involved mapping, measuring and tagging all trees ≥ 10 cm diameter-at-breast height (DBH) inside each quadrat. Tree heights were estimated by field teams on the ground, and were not significantly different from model-based estimates obtained using DBH measurements (M. Pfeifer, unpublished data). We also included a 2nd-order polynomial term for DBH during the analysis. This was because we expected hump-shaped responses to this variable in at least some species, owing to the fact that the largest mean DBH values were observed within oil palm quadrats. Total deadwood volumes in each quadrat were obtained by summing the volumes of all coarse woody debris pieces (≥ 10 cm diameter), including standing, hanging and fallen deadwood. Volumes of each piece were estimated using the formula for a truncated cone, following (Baker et al. 2007), which required measuring the diameter of each piece at both ends, as well as the length. For standing deadwood, the top-most diameter was estimated using the taper function (Chambers et al. 2000).

S1. 2. Satellite-based measurement of environmental variables

Topographical variables were all derived from the Advanced Spaceborne Thermal Emission and Reflection Radiometer Global Digital Elevation Model (DEM) version 2 (https://lpdaac.usgs.gov), jointly developed by the Ministry of Economy, Trade and Industry of Japan and the United States National Aeronautics and Space Administration, and which was provided at 30 m horizontal resolution. We extracted elevation data from this DEM at our sampling points, applying bilinear interpolation. Slope was estimated as the maximum rate of change from each cell in the DEM and was calculated in degrees (taking values between 0 and 90). To calculate flow accumulation, we 1) created a depressionless DEM by filling in ‘sink’ artefacts in the data, 2) created a flow direction map from this corrected DEM, and then 3) for each cell, summed the number of upstream cells. Areas of zero or low flow accumulation represented ridges, whilst high flow accumulation areas represented gullies, stream heads, streams and rivers. For the purposes of analysis, we log-transformed flow accumulation, owing to the strong positive skew in the values, with rivers otherwise represented by very large values. All topographical variables were calculated in ArcMap version 10 (ESRI, Redlands, California, USA). During analysis, we also included 2nd-order polynomial terms for each topographical variable, owing to the non-linear, and possibly hump-shaped, species responses we expected for these variables.

Ground-based measurements of above-ground live tree biomass (AGB) were taken from all vegetation quadrats (n = 193) sampled across the SAFE Project, which amounts to intensive-sampling of > 12 ha in total. AGB was calculated for each quadrat using Chave et al.'s (2014) pan-tropical algorithm. Spectral data were derived from sensors onboard the RapidEye satellite (European Space Agency Earth Observation Portal: https://earth.esa.int), which were provided at 5 m resolution. All pre-processing and atmospheric correction of the images, taken in 2012 and 2013, followed the steps outlined in Pfeifer et al. (2016). Within 20 m radius buffers centred on each quadrat, we extracted the spectral intensity values for each of the five bands present in the RapidEye images (blue, green, red, red-edge and near-infrared) and calculated a spectral vegetation index, the Modified Soil-Adjusted Vegetation Index 2 ($MSAVI_2$) from the red and near-infrared bands (Qi et al. 1994). Note, we did not take the red and near-infrared band spectral intensities forward into modelling, since these were used in the calculation of $MSAVI_2$. We transformed $MSAVI_2$ by taking its exponent, because of the saturating response observed at high levels of AGB. We also calculated a measure of image texture (dissimilarity), within 9 x 9 pixel moving windows, for each band. 58 quadrats were covered by cloud or cloud shadow in our images and were excluded. Using linear models of AGB as a function of each possible combination of the nine covariates ($MSAVI_2$, three spectral intensity covariates and five dissimilarity covariates), we then obtained a candidate set of “best” models based on information-theoretic criteria (i.e. models for which $\Delta AICc < 4$) and calculated model-averaged estimates for each parameter (Table S2) based on the model selection weights in this set. The pseudo-R^2 (explained deviance) of this final model was 0.53.

Based on the model-averaged parameters, we made AGB predictions within 500 m, 1 km and 2 km radius buffers surrounding each of our sampling points, at a resolution of 25 m^2, which matched the resolution of our ground-based measurements. Finally, we calculated the mean AGB (excluding cloud-covered pixels) within each buffer size. All steps in the analysis of AGB were done in R version 3.1.0 (R Development Core Team, 2014), using the packages raster 2.3-0 (Hijmans 2014), rgeos 0.3-8 (Bivand & Rundel 2014), glcm 1.2 (Zvoleff 2015) and MuMIn (Barton 2015).

Table S2. Model-averaged parameter estimates for linear models of field-based above-ground live tree biomass (AGB) measurements, as a function of satellite-derived measures of vegetation, spectral intensity and image texture.

Parameter	Estimate	Standard error (adjusted)	z-value	p-value	Relative variable importance[a]
Intercept	145.03	38.21	3.80	< 0.001	-
Band 2 (green) intensity	-17.44	3.56	4.91	< 0.001	1
Band 2 (green) dissimilarity	20.99	5.92	3.55	< 0.001	1
Band 4 (red-edge) intensity	8.28	2.61	3.18	0.001	1
exp(MSAVI$_2$)	-29.97	12.05	2.49	0.013	0.97
Band 3 (red) dissimilarity	6.41	5.00	1.28	0.20	0.42
Band 1 (blue) dissimilarity	5.69	5.11	1.12	0.27	0.34
Band 5 (near-infrared) dissimilarity	-1.15	1.34	0.86	0.39	0.31
Band 4 (red-edge) dissimilarity	-1.89	3.38	0.56	0.58	0.26
Band 1 (blue) intensity	-0.13	2.81	0.05	0.96	0.19

[a] Calculated as the sum of the AICc weights for the models in which the given parameter appears.

AGB values across buffer sizes were highly correlated (Pearson's $r > 0.98$), so we fitted redundancy analysis (RDA) models (with the vegan 2.0-10 package in R; Oksanen et al. 2013) for each buffer size and used the buffer size explaining the largest share of the community variation (calculated using the adjusted coefficient of multiple determination, R^2_{adj}; Blanchet et al. 2008) in further analyses. This selected the buffer with a 500 m radius (R^2_{adj} = 11.1%), although there were not large differences in the variation explained by the different buffer sizes (1 km: 10.8 %; 2 km: 10.7 %).

In order to calculate landscape forest cover and distances from forest for each sampling point, we first created a digitised forest cover map in ArcMap using visual interpretation of RapidEye satellite images, in combination with cloud-free Landsat 7 and 8 images (30 m resolution) released by Hansen et al. (2013). We distinguished natural forest from mature oil palm and *Acacia mangium* plantations by observing the dynamics of vegetation gain and loss over multiple years (1999-2013), as well as using our detailed knowledge of the study sites, but it is possible that some older plantation areas may have been included in our forest cover map (if they were already > 5 m in height before the year 2000 and were not harvested after that time). Euclidean distances from forest were calculated in ArcMap and percentage forest cover was quantified in buffers with 500 m, 1 km and 2 km radii using the rgeos package in R. As for AGB, we fit RDA models for each of the buffer sizes and selected the radius which explained the largest percentage of the community variation. This resulted in the 500 m radius being chosen (R^2_{adj} = 6.2%), though similar percentage variances were explained by the other buffer sizes (1 km: 5.8%; 2 km: 5.3%).

References

Baker, T. R., E. N. Honorio Coronado, O. L. Phillips, J. Martin, G. M. F. Van Der Heijden, M. Garcia, and J. Silva Espejo. 2007. Low stocks of coarse woody debris in a southwest Amazonian forest. Oecologia **152**:495–504.

Barton, K. 2015. MuMIn: Multi-Model Inference. R package version 1.13-4. Available from http://cran.r-project.org/package=MuMIn.

Bivand, R., and C. Rundel. 2014. rgeos: Interface to Geometry Engine - Open Source (GEOS). R package version 0.3-8. Available from http://cran.r-project.org/package=rgeos.

Blanchet, F., P. Legendre, and D. Borcard. 2008. Forward selection of explanatory variables.

Ecology **89**:2623–2632.

Chambers, J. Q., N. Higuchi, and J. M. Melack. 2000. Decomposition and carbon cycling of dead trees in tropical forests of the central Amazon. Oecologia **122**:380–388.

Chave, J. et al. 2014. Improved allometric models to estimate the aboveground biomass of tropical trees. Global Change Biology **20**:3177–3190.

Cusack, J. J., O. R. Wearn, H. Bernard, and R. M. Ewers. 2015. Influence of microhabitat structure and disturbance on detection of native and non-native murids in logged and unlogged forests of northern Borneo. Journal of Tropical Ecology **31**:25–25.

Ewers, R. M. et al. 2011. A large-Scale forest fragmentation experiment: the Stability of Altered Forest Ecosystems Project - DRAFT. Philosophical transactions of the Royal Society of London. Series B, Biological sciences **366**:3292–302.

Hansen, M. C. et al. 2013. High-resolution global maps of 21st-century forest cover change. Science **342**:850–853.

Hijmans, R. J. 2014. raster: Geographic data analysis and modeling. R package version 2.3-0. Available from http://cran.r-project.org/package=raster.

Lemmon, P. E. 1957. A new instrument for measuring forest overstory density. Journal of Forestry **55**:667–668.

Malhi, Y. et al. 2002. An international network to monitor the structure, composition and dynamics of Amazonian forests (RAINFOR). Journal of Vegetation Science **13**:439–450.

Oksanen, J. et al. 2013. vegan: Community ecology package. R package version 2.0-10. Available from http://cran.r-project.org/package=vegan.

Pfeifer, M., L. Kor, R. Nilus, E. Turner, J. Cusack, I. Lysenko, M. Khoo, V. Chey, A. Y. C. Chung, and R. M. Ewers. 2016. Mapping the structure of Borneo’s tropical forests across a degradation gradient. Remote Sensing of Environment **176**:84–97.

Qi, J., A. Chehbouni, A. R. Huete, Y. H. Kerr, and S. Sorooshian. 1994. A modified soil adjusted vegetation index. Remote Sensing of Environment **48**:119–126.

R Development Core Team. 2014. R: A language and environment for statistical computing. Vienna, Austria.

Turner, E. C., Y. Z. Abidin, H. Barlow, M. Tom, A. Nainar, G. Reynolds, Y. B. I. N. Yusof, M. I. N. S. Khoo, and R. M. Ewers. 2012. The Stability of Altered Forest Ecosystems Project: Investigating the Design of Human-Modified Landscapes for Productivity and Conservation. The Planter **88**:453–468.

Wearn, O. R., J. M. Rowcliffe, C. Carbone, H. Bernard, and R. M. Ewers. 2013. Assessing the status of wild felids in a highly-disturbed commercial forest reserve in Borneo and the implications for camera trap survey design. Plos One **8**:e77598.

Zvoleff, A. 2015. glcm: Calculate textures from grey-level co-occurrence matrices (GLCMs) in R. R package version 1.2. Available from http://cran.r-project.org/package=glcm.