*Statistical analyses*. We used on-line worksheets provided by McDonald (2009) to perform various statistical tests with a probability of type I error <0.05 considered significant. Data within a site were not analyzed to determine variation among years but were pooled across years to increase sample sizes.

*Rarefaction*

We used rarefaction to compare species richness among sites. Rarefaction produces idealized species-accumulation curves that allow direct comparison of results among groups that differ in patterns of abundance or are sampled using different techniques (Gotelli and Colwell 2001). Rarefaction calculates the expected species richness of the different groups for a constant sampling effort, but does not provide an estimate of asymptotic richness. Rather, for each accumulation curve we calculated a Chao 1 non-parametric estimator of richness with its variance and 95% confidence interval (Chao 1984).

*Shannon Diversity, Dominance, Evenness,*

We compared species diversity with Shannon Diversity Index, and converted diversity to the “effective number of species” (Jost 2006) to represent true diversity with mathematical properties allowing comparison among groups. We calculated numerical dominance of species captured within each site, and an evenness index (Magurran 1988) of the distribution of individuals among taxa (absolute evenness = 1.0).

*Jaccard’s Index, Sorenson’s*

We used Jaccard’s index to compare the similarity of communities based on presence/absence of species, and Sorenson’s measure to compare the similarity of sites based on the proportional abundance of species recorded, and the proportional abundance of birds in each diet category.

*Capture rates*

We used a chi-square test of independence to examine captures among sites for all species with >30 mist net captures. Comparisons were based on actual numbers (not rates) of captures, with expected values based on net hours (Blake and Rougès 1997). Because of the large number of planned comparisons, we used the Dunn-Šidák method to decrease the level of α and reduce the probability of committing a Type I error.

*Site persistence*

We used a chi-square test to examine differences in site persistence among sites for species with sufficiently large sample sizes, or a Fisher’s exact test of independence when expected numbers were small.

*Age ratio, sex ratio, migratory status, diet, habitat*

We used these tests to examine differences in the proportion of adults and the proportion of males among sites, and to test for significant heterogeneity in the proportion of species and individuals among sites grouped by migratory status, diet category, and preferred habitat.

*Insect abundance*

All insect abundance data were analyzed with G-tests of Independence.

[General modeling approach]

Data from this study

To model sex ratios, age ratios, and etc…

Due to the nested structure of the data we used generalized linear mixed models (GLMMs) with site (4 or 5 levels), year (x), and bird species (x-y) as a random intercepts. Models were fit in *R* 3.4.4 (R Core Team 2017) using the *lme4* package (Bates et al. 2015). Because of the relatively low number of levels to our random effects (4-x) and the nested structure of the data we had to address convergence issues while fitting our model. To address convergence issues we used the *lme4* extension *blme* (Chung et al. 2013). This package facilitates model fitting by using a modified likelihood function (?) (penalized likelihood) that is mathematically equivalent to using a weakly informative prior in a Bayesian model. Though this approach borrows from Bayesian approaches, inference is carried out using standard frequentists methods (Chung et al. 2013, Chung et al. 2015).

Centered predictors

It is sometimes recommended to simplify mixed models if random terms are not significant and to improve convergence (Matuschek et al 2017; but see Bar et al 2013). We *had a priori* expectations for there to be correlations with sites and within years and therefore did not feel it was appropriate to simplify our models and so used the quasi-Bayesian approach advocated by Chung et al. (2013).

To further avoid optimization problems we used the *all\_fit()* function in the *afex* package (Singman et al 2018) to find a numeric optimizer that would not result in convergence warnings from the *lme4* package. Using the *Nelder\_Mead* and occasionally xxxx optimizers produced satisfactory results. We confirmed that our final models met convergence criteria by checking …. (gradient etc stuff; how to cite?).

To understand how bird communities changed as pastures aged we fit two sets of models. First, we modeled how characteristics of the bird communities changed over time as pastures aged at the La Cueva (2-x years), La Caoba (5- x year), Morelia (10 – x years), and El Corral (20-x) sites years using multilevel models with random pasture age slopes for each species. This approach provides a number of advantages relative to building separate models for each species, including accounting for correlations across species due to common study years and sites shared across species and joint modeling of shared characteristics of species (eg. migration behavior). These models can both improve power by sharing information across species (“borrowing strength”; ref) and reduce Type I errors without the need for corrections for multiple comparisons (Gelman xxxx). This approach is becoming increasing popular for multi-species studies (Jackson et al 2012, Brouwer et al. 2017).

[by species and by traits/characteristics?]

We calculated 95% confidence intervals around regression lines using the xxx function in the *merTools* package (Knowles and Frederick 2016) and built plots using *ggplot2* (Whickham 2009). This first set of models did not include the mature forest site (Aceitillar) because to our knowledge this site has never been significantly disturbed. To allow approximate comparisons with trends from these models we calculated an overall mean for Aceitillar using just time and species as random intercepts.

To allow comparison between the four sites undergoing succession and the mature forest site at Aceitillar we fit one-way ANOVA-style models which treated time since disturbance (age) as a categorical variable. The two youngest sites had shared values for age and so we removed the overlapping years. We did an omnibus likelihood ratio test to determine if there was evidence for any difference between sites and the did a focused test for a linear or quadratic trend across sites using (cite ). Trend tests were conducted using the *multcomp* package (Hothorn et al 2008). Linear and quadratic trneds

Log transformed species richness and logit transformed eveness. All test statistics are reported for transformed data; analyses we re-run on untransformed data for plotting.

To determine if body condition varied among sites we first calculated the Scaled Mass Index (Peig and Green 2009) based on log-transformed body mass and wing chord measurements, using the *smatr* package (Warton et al 2012) to calculate the scaling exponent *bSMA*. SMI was only calculated for individuals that could be classified as site persistent, either because they were resighted or recaptured within a single season or were observed in subsequent years. If a bird was captured more than once we only used mass and length data from its first captures. For each species we modeled whether (log-transformed?) SMA varied with pasture age using one-way ANOVA style models with site as a random effect and only considered species that occurred both in the pasture sites and the mature forest.

Continuous numeric response variables (eg Simpson’s diversity, body codition?) was log transformed to improve normality. We analyzed binary data age and sex ratios using logistic GLMMs and counts (species richness, capture numbers) with poisson GLMMs. To account for overdispersion in counts we included and individual-level random effect (CITE).

To assess how insect availability changed between sites we calculated richness of taxonomic groups and diversity of taxonomic groups and analyzed change between sites using one-way ANOVA type models with site as a random effect. We investigated changes in community composition using permutational MANOVA (PERMANOVA; Anderson 2001) implemented in the *R* package *vegan* using site as a random effect. Insect samples were collected during the first two years of the study (see xxx above) and so there is no overlap in ages between the two youngest sites.

### [A new method for non‐parametric multivariate analysis of variance](http://onlinelibrary.wiley.com/doi/10.1111/j.1442-9993.2001.01070.pp.x/full)

[MJ **Anderson**](https://scholar.google.com/citations?user=vVvah5EAAAAJ&hl=en&oi=sra) - Austral **ecology**, 2001

Anderson, MJ 2001. A new method for non-parameteric multivariate analysis of variance. Austral Ecology 26:32 – 46.

Bates, D, M Maechler, B Bolker, Steve Walker (2015). Fitting

Linear Mixed-Effects Models Using lme4. Journal of Statistical Software,

67(1), 1-48. doi:10.18637/jss.v067.i01.

Chung, Y, S Rabe-Hesketh, V Dorie. 2015, A Gelman and J Liu. 2013. A nondegenerate penalized likelihood estimate for variance parameters in multilevel models. Pyschometrika 78: 685-709.

Chung Y, Rabe-Hesketh S, Dorie V, Gelman A and Liu J (2013). “A

nondegenerate penalized likelihood estimator for variance parameters in

multilevel models.” \_Psychometrika\_, \*78\*(4), pp. 685-709. <URL:

<http://gllamm.org/>>.

Chung, Y, A Gelman, S Rabe-Hesketh, J Liu and V. Dorie. 2015. Weakly informative prior for point estimation of covariance matrices in hierarchical models. Journal of Educational and Behavioral Statistics 40: 136-157.

Knowles, JE and C Frederick (2016). merTools: Tools for Analyzing

Mixed Effect Regression Models. R package version 0.3.0.

https://CRAN.R-project.org/package=merTools

Hothorn,T, Frank Bretz and Peter Westfall (2008). Simultaneous

Inference in General Parametric Models. Biometrical Journal 50(3),

346--363.

Jackson, MM, MG Turner, SM Pearson, AR Ives. 2012. Seeing the forest and the trees: multilevel models reveal both species and community patterns. Ecosphere 3: Article 79.

Oksanen, J, FG Blanchet, M Friendly, R Kindt, P Legendre, D McGlinn, PR

Mincin, RB O'Hara, GLSimpson, P Solymos, MHH Stevens, E Szoecs and H Wagner (201). vegan: Community Ecology Package. R package version 2.4-5.

https://CRAN.R-project.org/package=vegan

R Core Team (2017). R: A language and environment for statistical

computing. R Foundation for Statistical Computing, Vienna, Austria. URL

<https://www.R-project.org/>.

Singmann, H, B Bolker, J Westfall and F Aust (2018). *afex*:

Analysis of Factorial Experiments. R package version 0.19-1.

<https://CRAN.R-project.org/package=afex>

Warton, DI., Duursma, Remko A., Falster, Daniel S. and Taskinen, Sara

(2012) smatr 3 - an R package for estimation and inference about

allometric lines Methods in Ecology and Evolution, 3(2), 257-259

Wickham, H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag

New York, 2009.