THE ROLE OF POPULATION STRUCTURE IN PATHOGEN DIVERSITY IN WILD BAT POPULATIONS

TIM LUCAS. DECEMBER 15, 2015

1. Abstract

1.0.1. *One or two sentences providing a basic introduction to the field.* It is still unclear what factors determine the number of pathogens a wild species carries. But once understood, these factors could provide a way to prioritise surveillance of wild populations for zoonotic disease.

1.0.2. *Two to three sentences of more detailed background.* The pattern of contacts between individuals (i.e. population structure) has long been known to strongly affect epidemic processes. Theory suggests that population structure can promote pathogen richness while the ecological literature generally assumes it will decrease richness. Previous studies in wild populations have had contradictory results and the different measures of population structure have different shortcomings.

1.0.3. *One sentence clearly stating the general problem (the gap).* Here I use comparative data to test whether population structure influences pathogen richness in bats as they have been associated with a number of important, recent zoonotic outbreaks. Unlike previous studies I use two measures of population structure: a novel measure, number of subspecies, and a more careful application of genetic measures which have been used previously.

1.0.4. *One sentence summarising the main result.* I find that both of these measures are associated with pathogen richness but with effects in opposite directions suggesting that population structure has no clear effect.

1.0.5. *Two or three sentences explaining what the main result reveals in direct comparison to what was thought to be the case previously.* The results conflict with each other and with other studies which suggests that tests of population structure are sensitive to the exact measurements and data used. Given the conflicting results in the literature and unclear results here, it seems likely that population structure does not strongly affect pathogen richness in bats.

1.0.6. *One or two sentences to put the results into a more general context.* The use of larger datasets and and multiple measurements of population structure is therefore important to ensure the robustness of results. Given the weakness of any association between population structure and pathogen richness in bats, this is not a useful metric for prioritising zoonotic disease surveillance.

1.0.7. *Two or three sentences to provide a broader perspective,*

2. Introduction

2.1. **General Intro.** The number of pathogen species carried by a host species has important consequences for the ecology of the host and the probability that the host will be a reservoir of a zoonotic pathogen. However, the factors that affect pathogen richness are poorly understood.

2.2. **Specific Intro.**

2.3. **Theoretical background.** Single pathogen models show that increasing population structure simply slows disease spread and makes establishment less likely (). In the ecological literature this is often taken as predicting that increased population structure will decrease pathogen richness (). However, models of competition between multiple pathogens show that in unstructured populations a competitive exclusion process occurs but that splitting the population into two patches allows coexistence ().

2.4. **Previous Studies.** Three studies have used comparative data to test for an association between population structure and viral richness. A study on 15 African bats found a positive relationship between distribution fragmentation and viral richness (Maganga et al. 2014) while a study on 20 South-East Asian bats found the opposite relationship (Gay et al. 2014). A global study on 33 bats found a positive relationship between F_{ST} — a measure of genetic structure — and viral richness (Turmelle and Olival 2009). However, this study included measures using mtDNA which only measures female dispersal which may have biased the results many bat species show female philopatry (Kerth et al. 2002; Hulva et al. 2010). Furthermore, this study used measures of F_{ST} irrespective of the study scale with studies covering from tens (McCracken and Bradbury 1981) to thousands (Petit and Mayer 1999) of kilometers. As isolation by distance has been shown in a number of bat species (Burland et al. 1999; Hulva et al. 2010; O'Donnell et al. 2015; Vonhof et al. 2015) this could bias results further. Finally, when a global F_{ST} value is not given they use the mean of all pairwise F_{ST} between sites. It is not clear that this is correct as from global F_{ST} we expect migration rates of $M = \frac{1-F_{ST}}{8F_{ST}}$ while from F_{ST} between pairs of populations we expect migration rates of $M = \frac{1-F_{ST}}{16F_{ST}}$ where M is the absolute number of diploid inviduals dispersing per generation (Slatkin 1995). As it is in fact the movement of individuals that is epidemiologically relavent, using these studies is probably not correct without attempting to correct for these difference.

Studies on single pathogens, notably rabies, have also shown that for virulent pathogens, space can allow persistence where a well mixed population ould experience a single, large epidemic then pathogen extinction (**blackwood2013resolving**; **colizza**; Pons-Salort et al. 2014).

2.5. **Rates.**

2.6. **Choice of measure of population structure.** A number of measurments of population structure have been used and each has it's own shortcomings. In particular, the better, more direct measurements tend to be very work intensive which consequently means data is available for few species.

2.6.1. *Direct dispersal measurements.* The ideal measurement of population structure is direct measurement of dispersal rates and distance. These are incredibly difficult to obtain, especially over large scales. Due white nose syndrom, some very large mark-recapture studies have been conducted, but recapture rates are low. Further, these large studies have been in species the live in a fe large colonies, so recapture rates should be higher than in less social species.

2.6.2. *Genetic measures.* As direct measurement of dispersal are difficult, genetic data is often used. Measurement such as F_{ST} are used to calculate migration. There are strong model assumptions under the conversion from F_{ST} to migration. However, the main issue with this measure is the effort required for each study and the subsequent low number of measurements. Further, there are differences in the scales of the studies and the genetic regions being sequenced. This differences should not be ignored.

2.6.3. *Number of Subspecies.* For a population to evolve distinct phenotypic or genetic traits, such that they can be classed as a subspecies, there must be limited migration between populations. The number of subspecies a species has therefore reflects the level of population structure in that species. The value of this measurement is available for every bat species. However, it is likely biased, with well studied species being likely to have more recognised subspecies. Further, this is a very course measure and it is important to consider whether it is measuring migration at a timescale and rate that is epidemiologically relevant.

2.6.4. *Measures from range.* The final measurement that has been used is derived from the shape of the species' range, typically from IUCN (**iucn**) maps. The ratio between the perimeter of the range and the area (or similar values) are calculated. Range maps are very course for many species. Furthermore there is a potential bias with island living species being given sea based edges where continental species might be assumed to live everywhere in between locations where it is known to live, without considering the different terrestrial habitats in these areas.

2.7. **The gap.** There is a lack of studies using multiple measures of population structure and larger datasets to robustly estimate the importance of population structure. Furthermore, the

2.8. **What I did.** Here I use two measures of population structure — the number of subspecies and gene flow — to robustly test for an association between population structure and pathogen richness. Furthermore, I use a much larger dataset for one of these analyses, further promoting robustness of results.

2.9. **What I found.**

3. Methods

To measure pathogen richness I used data from (Luis et al. 2013). These simply include known infections of a bat species with a pathogen species. Only species with at least one pathogen were included in the analysis. Rows with host species that were not identified to species level were removed. Many viruses were not identified to species level or their identified species was not in the ICTV virus taxonomy (ICTV 2014). I counted a virus if it was the only virus, for that host species, in the lowest taxonomic level identified in the ICTV taxonomy. That is, if a host carries an unknown Paramyxoviridae virus, then it must carry at least one Paramyxoviridae virus. If a host carries an unknown Paramyxoviridae virus and a known Paramyxoviridae virus, then it is hard to confirm that the unknown virus is not another record of the known virus. In this case, this would be counted as one virus species.

I used two measures of population structure. F_{ST} and the number of subspecies. The number of subspecies was counted using the Wilson and Reeder taxonomy (Wilson and Reeder 2005). F_{ST} and other measures were collated from the literature. Studies are from a wide range of spatial scales, from local ($\sim 10\,\text{km}$) to continental. As F_{ST} inevitably increases with spatial scale I controlled for this by only using data from studies where a large proportion of the species range was studied. I used the ratio of the furthest distance between F_{ST} samples (measured with `http://www.distancefromto.net/` if not stated) to the width of the IUCN species range and only used studies if this ratio was greater than 0.2. To allow comparison between different

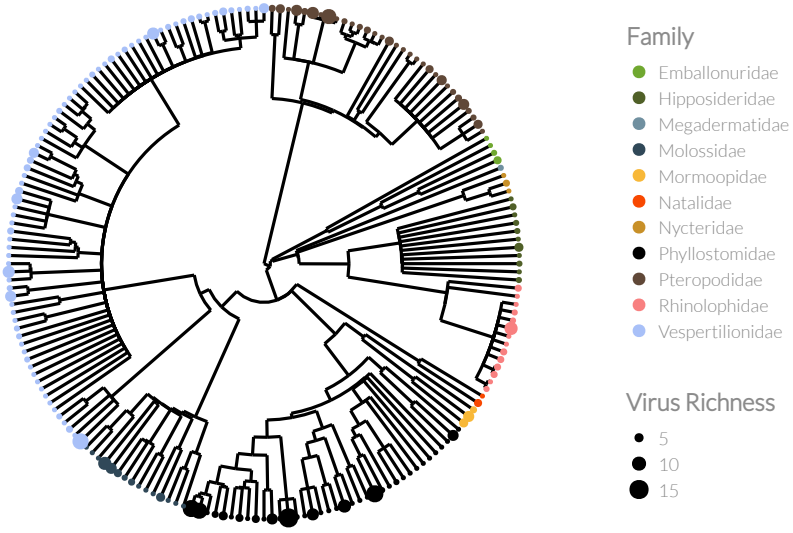

Figure 1. Pruned phylogeny with dot size showing number of pathogens and colour showing family.

measures (F_{ST}, ϕ_{ST}) and data from different molecular regions I converted all data to diploid gene flow. WILL ADD EXTRA METHODS LATER. These two measures of population structure were analysed separately as the number of subspecies has 196 data points while there is only F_{ST} data for ~ 30 bat species.

To control for study bias I collected the number of Pubmed and Google Scholar citations for each bat species including synonyms from ITIS (*Integrated Taxonomic Information System (ITIS)* n.d.) via the taxize package (Chamberlain and Szöcs 2013). The counts were scraped using the rvest package (Wickham 2015). I log transformed these variables as they were strongly right skewed. The log number of citations on Pubmed and Google scholar were highly correlated (pgls: $t = 19.32$, df $= 194$, $p = 0$). The results here are for analyses using only Google Scholar citations. See the appendix for analyses run using Pubmed citations.

Measures of body mass are taken from Pantheria (Jones, Bielby, et al. 2009) and primary literature (Canals et al. 2005; Arita 1993; López-Baucells et al. 2014; Orr and Zuk 2013; Lim and Engstrom 2001; Aldridge 1987; Ma et al. 2003; Owen et al. 2003; Henderson and Broders 2008; Heaney et al. 2012; Oleksy et al. 2015; Zhang et al. 2009). *Pipistrellus pygmaeus* was assigned the same mass as *P. pipistrellus* as they indistinguishable by mass. Body mass measurements were log transformed due to the strong right skew. Distribution size was estimated by downloading range maps for all species from IUCN () and were also logged due to right skew.

To control for phylogenetic nonindependance I used the best-supported phylogeny from (Fritz et al. 2009) which is the supertree from (Bininda-Emonds et al. 2007) with names updated to match the Wilson & Reeder taxonomy (Wilson and Reeder 2005). Phylogenetic manipulation was performed using the ape package (Paradis et al. 2004). The importance of the phylogeny on each variable separately was estimated using (Revell 2012).

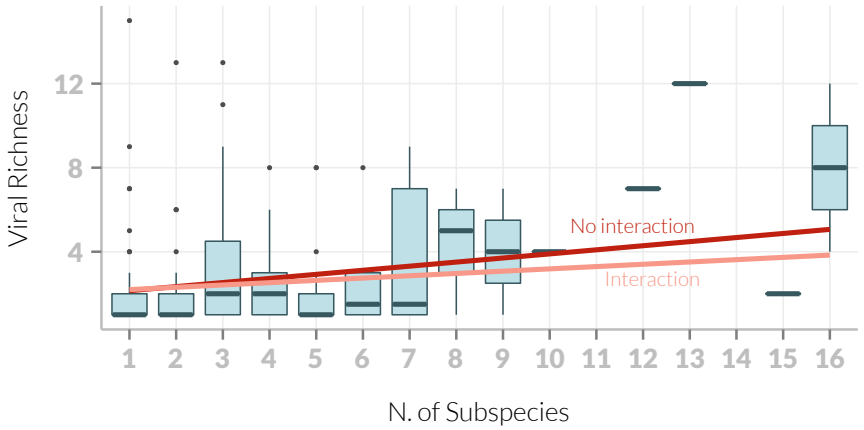

Figure 2. Number of virus species against number of subspecies. Data within a number of subspecies are plotted as boxplots with the dark bar showing the median, the box showing the interquartile range, vertical lines showing the range and outliers shown as seperate points. Regression lines are from multivariate phylogenetic models with other independant variables set at their median value. The models shown are thos with (pink) and without (red) an interaction between study effort and number of subspecies.

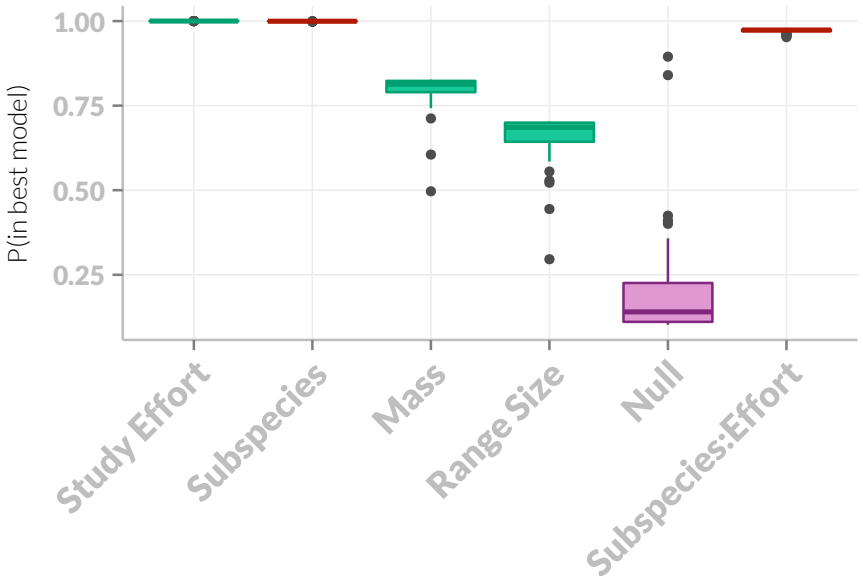

Figure 3. Akaika variable weights for number of subspecies analysis. The probability that each variable will be in the best model if the data were recollected is shown for each of the bootstrap analyses. The purple "Null" box is a uniform random variable used as a null. Population structure (Number of subspecies) and the interaction between subspecies and study effort, shown in red, are more likely to be in the best model than this random variable.

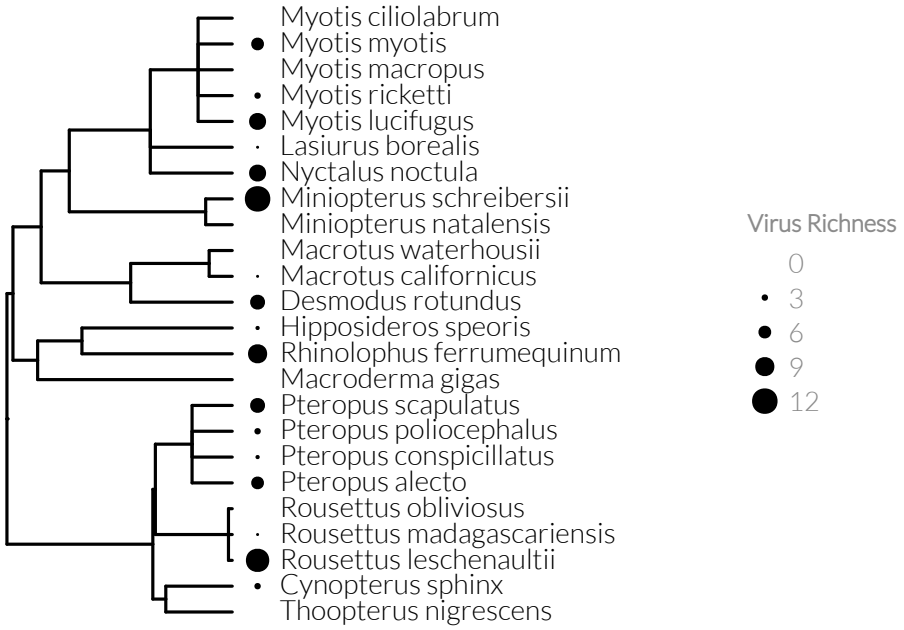

Figure 4. Pruned phylogeny with dot size showing number of pathogens and colour showing family.

F_{ST} studies are conducted at a range of spatial scales, but F_{ST} often increases with distance studied (). To minimise the effects of this I only used data from studies that cover 20% of the diameter of the species range. This is a largely arbitrary value that could be considered to reflect a 'global' estimate of F_{ST} while keeping a reasonable number of datapoints available. I calculated the diameter of the species range by finding the furthest apart points in the IUCN species range () even if the range is split into multiple polygons. The width covered by each study was the distance between the most distant sampling sites. When this was not explicit in the paper, the centre of the lowest level of geographic area was used.

3.1. **Statistical analysis.** Statistical analysis for both dependant variables was conducted using a information theory/model averaging approach (Burnham and Anderson 2002) specifically following (Whittingham, Swetnam, et al. 2005; Whittingham, Stephens, et al. 2006). I chose a credible set of models including all combinations of independent variables. In the analysis using the number of subspecies dependant variable I also included an interaction term between study effort and number of subspecies as I believe *a priori* that this interaction may be present. The interaction was only included in models with both study effort and number of subspecies as an individual term.

I fitted phylogenetic regressions using nlme (Pinheiro et al. 2015) to all models. In each case I simultaneously fitted the λ parameter as this avoids mispecifying the model (Revell 2010). κ and δ were constrained to one.

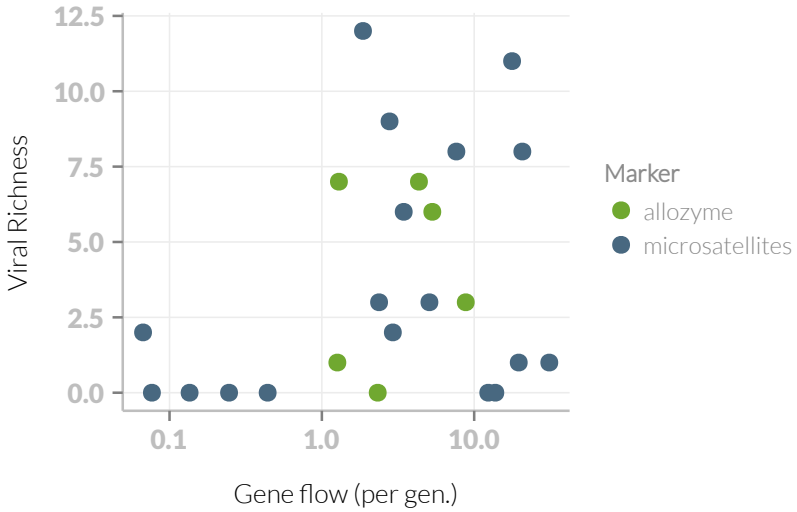

Figure 5. Gene flow per generation (on a log scale) against viral richness with the genetic marker used shown with colour.

I calculated AICc for each model and then calculated Akaiki weights. This value can be interpreted as the probability that a model would be the best model if the data were recollected. For each variable, the sum of the Akaiki weights of models containing that variable are summed. This value can be interpreted as the probability that the given variable is in the best model. Following (Whittingham, Swetnam, et al. 2005) I included a uniformally random variable as a null variable as even unimportant variables can have Akaiki weights notably greater than zero. The whole analysis was run 50 times, resampling the random variable each time. I calculated the average AICc, $\bar{\text{AIC}}$c, by averaging AICc scores within models. ΔAICc was calculated as $\min(\bar{\text{AIC}}\text{c}) - \bar{\text{AIC}}\text{c}$, not the mean of the individual ΔAICc scores, to guarantee that the best model has $\Delta\text{AICc} = 0$.

4. Results

4.1. **Number of Subspecies.**

4.1.1. *More descriptive.* After data cleaning there was data for 196 bat species in 11 families. The number of described virus species for a bat host ranged up to 15 viruses in *Carollia perspicillata*. Figure 1 shows the phylogeny used and the number of viruses for each species. The mean number of viruses across families is fairly constant with a lower range of 1.67 for Nycteridae. The highest mean is Mormoopidae with 5 virus species per bat species, but this is based on a sample size of 3. The Phyllostomidae have the second highest mean (n = 37) of 3.49.

The small change in mean pathogen richness across families and the lack of clear pattern in Figure 1 implies that viral richness is not strongly phylogenetic. This is corroborated by the small estimated size of λ ($\lambda = 0.04$, $p = 0.21$). This fact implies that other factors must control pathogen richness. It also implies that pathogens are not directly inherited down the phylogeny, although this is to be expected by the fast evolution of viruses.

Of the explanatory variables, the number of subspecies has no phylogenetic autocorrelation ($\lambda = 10^{-6}$, $p = 1$), study effort and distribution size have weak but significant autocorrelation (Study Effort: $\lambda = 0.1$, $p = 9.12 \times 10^{-3}$, Distribution size: $\lambda = 0.32$, $p = 8.82 \times 10^{-8}$) and mass is strongly phylogenetic ($\lambda = 0.93$, $p = 0$). The parameter λ is fitted and governs how

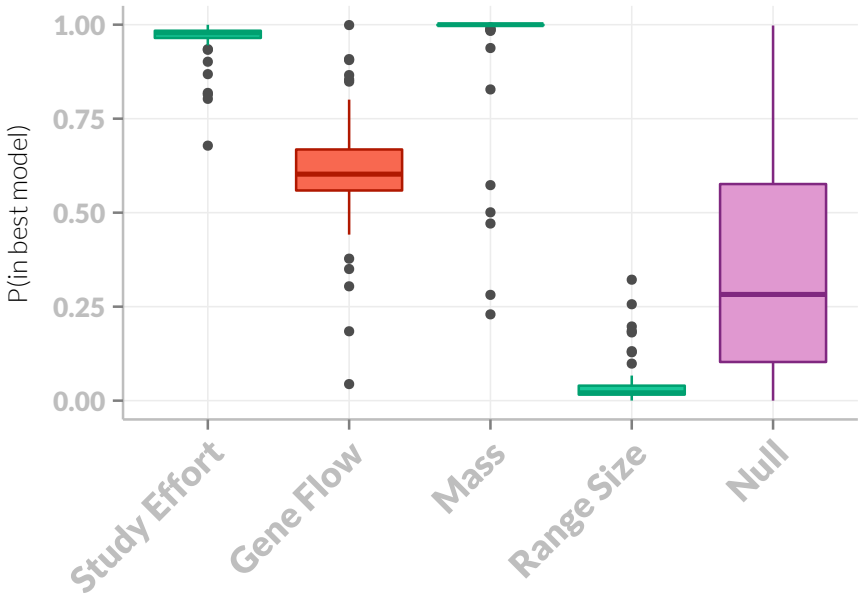

Figure 6. Akaika variable weights for F_{ST} analysis. The probability that each variable will be in the best model if the data were recollected is shown for each of the bootstrap analyses. The purple "Null" box is a uniform random variable used as a null. Population structure (F_{ST}), shown in red, is much less likely to be in the best model than this random variable.

important the phylogeny is in the model. Across all models the mean value of λ is 0.08 implying the residuals from the model are weakly phylogenetic. A small number of models (0.35%) had negatively phylogenetically distributed residuals.

4.1.2. *Model results.* Summing the Akaiki weights of all models that contain a given variable gives a probability, Pr, that the variable would be in the best model (Figure 3) if the data were recollected (Whittingham, Stephens, et al. 2006). The number of subspecies is very likely in the best model ($Pr > 0.99$) as is the interaction between number of species and study effort ($Pr = 0.97$) compared to the benchmark random variable which has $Pr = 0.2$. When models with the interaction term are removed, on average (mean weighted by Akaiki weights) there is a positive relationship between the number of subspecies and viral richness ($\beta = 0.19$, variance $= 1.24 \times 10^{-4}$). Models with an interaction between number of subspecies and study effort have a positive interaction term ($\beta = 0.14$, variance $= 3.94 \times 10^{-6}$) and a negative linear term ($\beta = -0.78$, variance $= 2.37 \times 10^{-4}$) which in total is a positive relationship for most of the range of the study effort data. This supports the hypothesis that population structure promotes pathogen richness. As seen in Figure 3, study effort is very likely in the best model ($\beta = 0.35$, $Pr > 0.99$) body mass and range size are also probably in the best model ($\beta = 0.38$, $Pr = 0.8$ and $\beta = 0.26$, $Pr = 0.66$).

4.2. **Gene Flow.**

Model	$\bar{\text{AIC}}$c	ΔAICc	w_i	$\sum w_i$
log(Scholar)*NSubspecies + rand	882	0.00	0.38	0.38
log(Scholar)*NSubspecies	884	1.39	0.19	0.57
log(Scholar)*NSubspecies + rand + log(Mass)	885	2.24	0.12	0.70
log(Scholar)*NSubspecies + log(Mass)	885	3.14	0.08	0.78
log(Scholar)*NSubspecies + log(Mass) + log(RangeSize)	886	3.18	0.08	0.86
log(Scholar)*NSubspecies + rand + log(RangeSize)	886	3.94	0.05	0.91
log(Scholar)*NSubspecies + log(RangeSize)	886	3.95	0.05	0.96
log(Scholar) + NSubspecies	889	6.93	0.01	0.97
log(Scholar) + NSubspecies + log(Mass)	890	7.80	0.01	0.98

Table 1. Model selection results for number of subspecies analysis. $\bar{\text{AIC}}$c is the mean AICc score across 50 resamplings of the null random variable. ΔAICc is the $\bar{\text{AIC}}$c score minus the lowest score. w_i is the Akaike weight and can be interpreted as the probability that the model is the best model (of those in the plausible set). $\sum w_i$ is the cumulative sum of the Akaike weights.

4.2.1. *More Descriptive.* Due to the low number of studies and the restrictive requirements imposed on study design, there are only data for 24 bat species in 7 families. The number of described virus species for a bat host ranged up to 12 viruses in *Miniopterus schreibersii*. Figure 4 shows the phylogeny used and the number of viruses for each species.

4.2.2. *Model results.* As with the Number of Subspecies dataset, there is no phylogenetic signal in the number of virus species ($\lambda = 10^{-6}$, $p = 1$). Gene flow also has no phylogenetic autocorrelation ($\lambda = 10^{-6}$, $p = 1$). Due to the low sample size, significance tests are unlikely to have much power. However, study effort has some phylogenetic autocorrelation ($\lambda = 0.15$, $p = 0.56$) while distribution size and mass seem to show phylogenetic signal (Distribution size: $\lambda = 0.67$, $p = 0.53$, Mass: $\lambda = 0.79$, $p = 2.69 \times 10^{-3}$).

While less strongly supported than the number of subspecies, gene flow is likely in the best model ($Pr = 0.61$) compared to the benchmark random variable which has $Pr = 0.37$ as shown in Figure 6. On average (mean weighted by Akaiki weights) there is a negative relationship between gene flow and viral richness (β = NA, variance = NA) despite the apparent positive relationship in Figure 5. This supports the hypothesis that population structure promotes viral richness. Possibly due to the smaller sample size, or a weaker relationship, this coefficient is much more varied than the number of subspecies coeficient with 24.25% of models estimating a positive relationship. As in the number of subspecies analysis, study effort is very likely in the best model ($Pr = 0.96$) as is body mass ($Pr = 0.94$). In contrast to the number of subspecies analysis, range size is almost certainly not in the best model with $Pr = 0.05$ being much less than the random variable.

Across all models the mean value of λ is -2.88 and a large number of individual models (66.58%) had negatively phylogenetically distributed residuals implying the residuals from the model are strongly negatively phylogenetic. Due to the small sample size this is probably due to a small number of data points with large residuals being distant on the tree.

Model	$\bar{\text{AIC}}$c	ΔAICc	w_i	$\sum w_i$
log(Scholar) + log(Gene flow) + log(Mass)	70	0.00	0.60	0.60
log(Scholar) + log(Mass)	71	0.94	0.38	0.98
log(Mass) + log(Range size)	77	6.73	0.02	1.00
log(Scholar) + log(Gene flow) + log(Mass) + rand	94	24.15	0.00	1.00

Table 2. Model selection results for gene flow analysis. The top two models are considerably better than the other models; there is a 98% chance that one of them is the best model. As body mass and study effort are in both these models these variables are very likely in the best model. Gene flow is in the top model, implying it may well be in the best model. $\bar{\text{AIC}}$c is the mean AICc score across 50 resamplings of the null random variable. ΔAICc is the $\bar{\text{AIC}}$c score minus the lowest score. w_i is the Akaike weight and can be interpreted as the probability that the model is the best model (of those in the plausible set). $\sum w_i$ is the cumulative sum of the Akaike weights.

5. Discussion

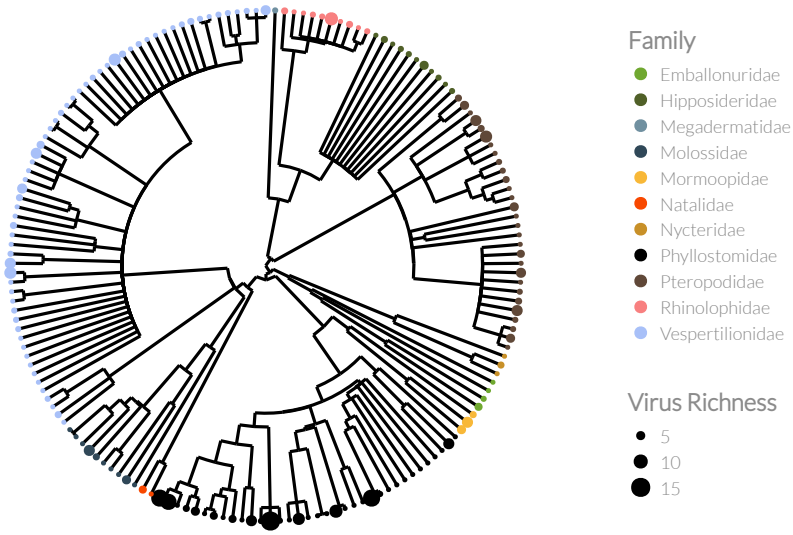

Figure 7. Pruned phylogeny (Jones, Bininda-Emonds, et al. 2005) with dot size showing number of pathogens and colour showing family.

6. Appendix

References

Maganga, G. D. et al. (2014). "Bat distribution size or shape as determinant of viral richness in african bats". In: *PloS one* 9.6, e100172.

Gay, N. et al. (2014). "Parasite and viral species richness of Southeast Asian bats: Fragmentation of area distribution matters". In: *International Journal for Parasitology: Parasites and Wildlife* 3.2, pp. 161–170. doi: doi:10.1016/j.ijppaw.2014.06.003.

Turmelle, A. S. and K. J. Olival (2009). "Correlates of viral richness in bats (order Chiroptera)". In: *EcoHealth* 6.4, pp. 522–539.

Kerth, G., F. Mayer, and E. Petit (2002). "Extreme sex-biased dispersal in the communally breeding, nonmigratory Bechstein's bat (*Myotis bechsteinii*)". In: *Molecular Ecology* 11.8, pp. 1491–1498.

Hulva, P. et al. (2010). "Mechanisms of radiation in a bat group from the genus Pipistrellus inferred by phylogeography, demography and population genetics". In: *Molecular ecology* 19.24, pp. 5417–5431.

McCracken, G. F. and J. W. Bradbury (1981). "Social organization and kinship in the polygynous bat *Phyllostomus hastatus*". In: *Behavioral Ecology and Sociobiology* 8.1, pp. 11–34.

Petit, E. and F. Mayer (1999). "Male dispersal in the noctule bat (Nyctalus noctula): where are the limits?" In: *Proceedings of the Royal Society of London B: Biological Sciences* 266.1430, pp. 1717–1722.

Burland, T. et al. (1999). "Population genetic structure and gene flow in a gleaning bat, Plecotus auritus". In: *Proceedings of the Royal Society of London B: Biological Sciences* 266.1422, pp. 975–980.

O'Donnell, C. F. et al. (2015). "Genetic diversity is maintained in the endangered New Zealand long-tailed bat (Chalinolobus tuberculatus) despite a closed social structure and regular population crashes". In: *Conservation Genetics*, pp. 1–12.

Vonhof, M. J., A. L. Russell, and C. M. Miller-Butterworth (2015). "Range-Wide Genetic Analysis of Little Brown Bat (Myotis lucifugus) Populations: Estimating the Risk of Spread of White-Nose Syndrome". In: *PloS one* 10.7.

Slatkin, M. (1995). "A measure of population subdivision based on microsatellite allele frequencies." In: *Genetics* 139.1, pp. 457–462.

Pons-Salort, M. et al. (2014). "Insights into persistence mechanisms of a zoonotic virus in bat colonies using a multispecies metapopulation model". In: *PloS one* 9.4, e95610.

Luis, A. D. et al. (2013). "A comparison of bats and rodents as reservoirs of zoonotic viruses: are bats special?" In: *Proceedings of the Royal Society B: Biological Sciences* 280.1756, p. 20122753.

ICTV (2014). *International Committee on Taxonomy of Viruses Master Species List*. url: `http://talk.ictvonline.org/files/ictv_documents/m/msl/5208/download.aspx`.

Wilson, D. E. and D. M. Reeder (2005). *Mammal species of the world: a taxonomic and geographic reference*. Vol. 12. JHU Press.

Integrated Taxonomic Information System (ITIS). `http://www.itis.gov`.

Chamberlain, S. A. and E. Szöcs (2013). "taxize: taxonomic search and retrieval in R". In: *F1000Research* 2.

Wickham, H. (2015). *rvest: Easily Harvest (Scrape) Web Pages*. R package version 0.2.0. url: `http://CRAN.R-project.org/package=rvest`.

Jones, K. E., J. Bielby, et al. (2009). "PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals: Ecological Archives E090-184". In: *Ecology* 90.9, pp. 2648–2648.

Canals, M. et al. (2005). "Relative size of hearts and lungs of small bats". In: *Acta Chiropterologica* 7.1, pp. 65–72.

Arita, H. T. (1993). "Rarity in Neotropical bats: correlations with phylogeny, diet, and body mass". In: *Ecological Applications*, pp. 506–517.

López-Baucells, A. et al. (2014). "Echolocation of the big red bat *Lasiurus egregius* (Chiroptera: Vespertilionidae) and first record from the Central Brazilian Amazon". In: *Studies on Neotropical Fauna and Environment* 49.1, pp. 18–25.

Orr, T. J. and M. Zuk (2013). "Does delayed fertilization facilitate sperm competition in bats?" In: *Behavioral Ecology and Sociobiology* 67.12, pp. 1903–1913.

Lim, B. K. and M. D. Engstrom (2001). "Bat community structure at Iwokrama forest, Guyana". In: *Journal of Tropical Ecology* 17.05, pp. 647–665.

Aldridge, H. (1987). "Turning flight of bats". In: *Journal of Experimental Biology* 128.1, pp. 419–425.

Ma, J. et al. (2003). "Dietary analysis confirms that Rickett's big-footed bat (*Myotis ricketti*) is a piscivore". In: *Journal of Zoology* 261.03, pp. 245–248.

Owen, S. F. et al. (2003). "Home-range size and habitat used by the northern myotis (*Myotis septentrionalis*)". In: *The American midland naturalist* 150.2, pp. 352–359.

Henderson, L. E. and H. G. Broders (2008). "Movements and resource selection of the Northern Long-Eared Myotis (*Myotis septentrionalis*) in a forest–agriculture landscape". In: *Journal of Mammalogy* 89.4, pp. 952–963.

Heaney, L. R. et al. (2012). "*Nyctalus plancyi* and *Falsistrellus petersi* (Chiroptera: Vespertilionidae) from northern Luzon, Philippines: ecology, phylogeny, and biogeographic implications". In: *Acta Chiropterologica* 14.2, pp. 265–278.

Oleksy, R., P. A. Racey, and G. Jones (2015). "High-resolution GPS tracking reveals habitat selection and the potential for long-distance seed dispersal by Madagascan flying foxes *Pteropus rufus*". In: *Global Ecology and Conservation* 3, pp. 678–692.

Zhang, L. et al. (2009). "Recent surveys of bats (Mammalia: Chiroptera) from China. I. Rhinolophidae and Hipposideridae". In: *Acta Chiropterologica* 11.1, pp. 71–88.

Fritz, S. A., O. R. Bininda-Emonds, and A. Purvis (2009). "Geographical variation in predictors of mammalian extinction risk: big is bad, but only in the tropics". In: *Ecology letters* 12.6, pp. 538–549.

Bininda-Emonds, O. R. et al. (2007). "The delayed rise of present-day mammals". In: *Nature* 446.7135, pp. 507–512.

Paradis, E., J. Claude, and K. Strimmer (2004). "APE: analyses of phylogenetics and evolution in R language". In: *Bioinformatics* 20, pp. 289–290.

Revell, L. J. (2012). "phytools: An R package for phylogenetic comparative biology (and other things)." In: *Methods in Ecology and Evolution* 3, pp. 217–223.

Burnham, K. P. and D. R. Anderson (2002). *Model selection and multimodel inference: a practical information-theoretic approach*. Springer Science & Business Media.

Whittingham, M. J., R. D. Swetnam, et al. (2005). "Habitat selection by yellowhammers Emberiza citrinella on lowland farmland at two spatial scales: implications for conservation management". In: *Journal of applied ecology* 42.2, pp. 270–280.

Whittingham, M. J., P. A. Stephens, et al. (2006). "Why do we still use stepwise modelling in ecology and behaviour?" In: *Journal of animal ecology* 75.5, pp. 1182–1189.

Pinheiro, J. et al. (2015). *nlme: Linear and Nonlinear Mixed Effects Models*. R package version 3.1-122. url: `http://CRAN.R-project.org/package=nlme`.

Revell, L. J. (2010). "Phylogenetic signal and linear regression on species data". In: *Methods in Ecology and Evolution* 1.4, pp. 319–329.

Wickham, H. (2009). *ggplot2: elegant graphics for data analysis*. Springer New York. isbn: 978-0-387-98140-6. url: `http://had.co.nz/ggplot2/book`.

Lucas, T. (2015). *palettetown: Use Pokemon Inspired Colour Palettes*. R package version 0.1.0. url: `http://CRAN.R-project.org/package=palettetown`.

Solt, F. and Y. Hu (2015). *dotwhisker: Dot-and-Whisker Plots of Regression Coefficients from Tidy Data Frames*. Available at The Comprehensive R Archive Network (CRAN). url: `http://CRAN.R-project.org/package=dotwhisker`.

Yu, G. (2015). "ggtree: an R package for versatile annotation and visualization of phylogenetic tree". In:

Jones, K. E., O. R. Bininda-Emonds, and J. L. Gittleman (2005). "Bats, clocks, and rocks: diversification patterns in Chiroptera". In: *Evolution* 59.10, pp. 2243–2255.