Details on Background similarity test for Malagasy reptiles:

1. Using same climate data as in the RTR analysis for Madagascar.
2. Background climate for each species defined as the minimum convex polygon around the occurrence data for each species including a 5km buffer zone
3. Projecton of occurrence densities onto a gridded PCA environment with a resolution of 100 cells
4. Background similarity test: Compares the observed niche overlap between species 1 and species 2 to overlaps between species 1 and random niches as available in the range of species 2 (and vice-versa).
5. Statistical significance based on 100 iterations
6. If the p-value for the observed niche overlap (measured by the D statistic) is higher or equal than 0.05, then the observed niche overlap is not statistically significant (*'n.s'* notation in the plots). Significant observed niche overlaps are annotated with a *'**'* in the plots.
7. As with the RTR, we also measured the percentile of the observed value (i.e., its position) relative to the random distribution (high percentiles refer to high niche overlaps relative to random and thus tendency for niche conservatism and the opposite for niche divergence)
8. Because the background test involves a reciprocal comparison, we assessed 2 sets of results (species1 with background of species 2 and species 2 with background of specis1.

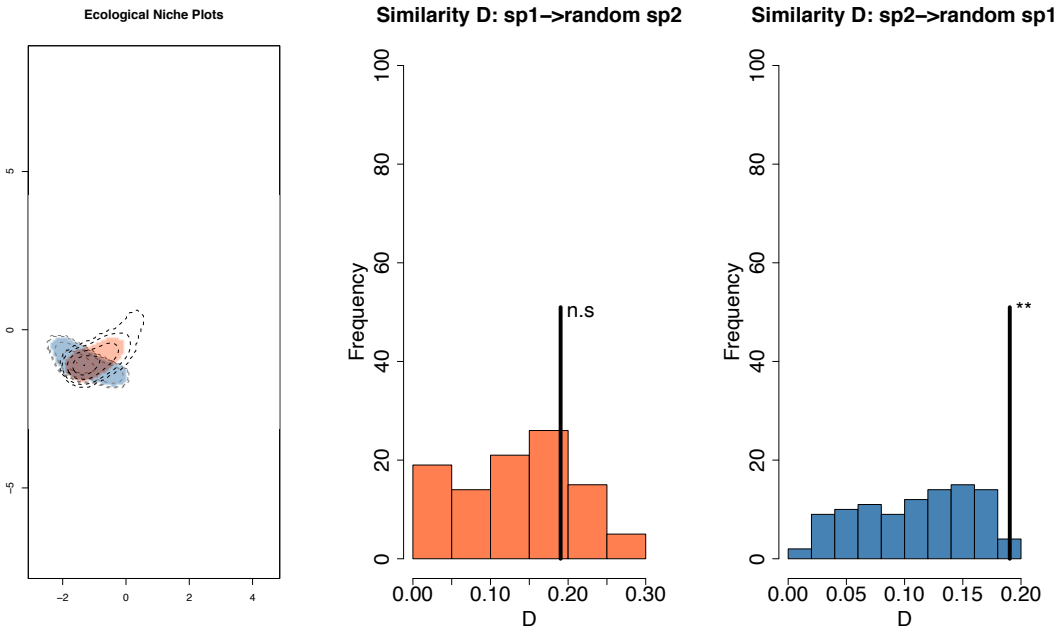

Figure S1: Outputs from background similarity test for the sister pair X X – Y Y. a, Ecological niche models for X X(species 1 in blue) and Y Y (species 2 in red) in environmental space for 2 species. **b**, Outputs of background similarity test assessing observed niche overlap with overlap between the niche of species 1 and simulated niches of species 2 and **c**, between niche of species 2 and to simulated niches of species 1. 'n.s' indicates a non-sigificant observed niche overlap and '**' refers to a significantly higher observed niche overlap, thus indicating PNC for this sister pair.

Results for comparison of species 1 with background of species 2

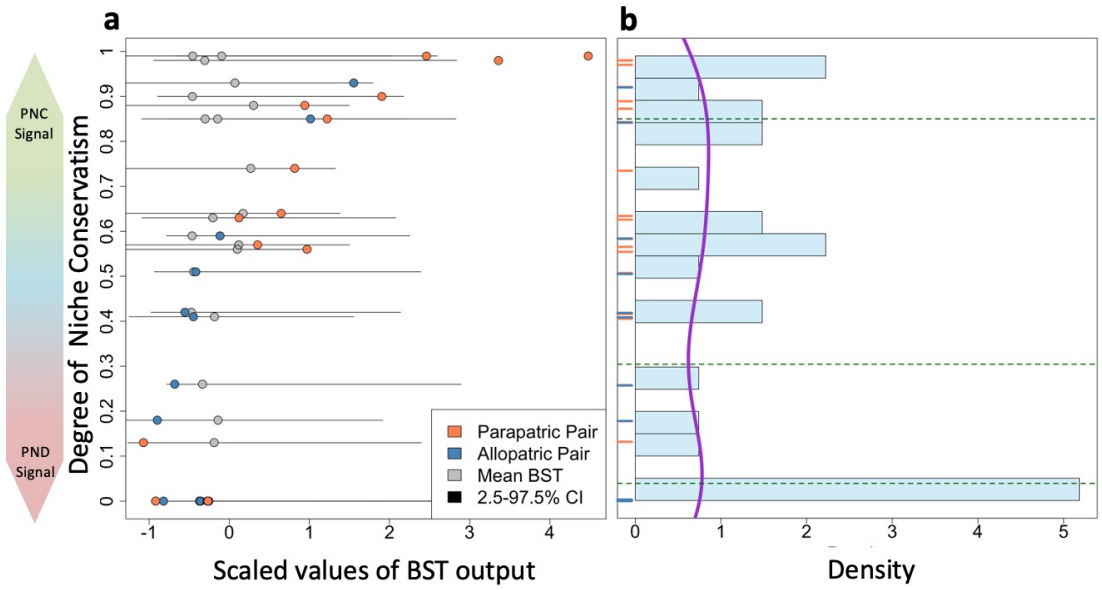

Figure S2. Bimodal distribution for the degree of niche conservatism for 28 sister pairs of endemic Malagasy lizards. a, Degree of niche conservatism against scaled values of background similarity test outputs, ordered by decreasing scaled niche overlap values for each sister pair. To illustrate the results for all sister pairs under one common scale, each set of outputs were scaled using their corresponding means and standard deviations using the scale function in R. The degree of niche conservatism of each sister pair is measured as the proportion of niche overlap replicates that are lower than the observed niche overlap (circles), so that higher values reflect a higher degree of phylogenetic niche conservatism (PNC) and lower values are associated with a lower degree of phylogenetic niche conservatism (PND). Cases where the observed niche overlap does not overlap with the 2.5-97.5% intervals of the replicate distributions have a significant ecological signal for either PNC (higher than 97.5%; 3 pairs) or PND (lower than 2.5%, 4 pairs). **b**, Density histogram showing the degree of niche conservatism for 28 pairs fitted with a kernel density line curve (purple). Hartingan's dip test supports a bimodal distribution (D = 0.98391, $P < 0.05$). Tick marks represent individual allopatric (blue) and parapatric (orange) sister pairs and dotted lines represent the upper and lower modes (0.858 and 0.038) respectively and the antimode (0.306) of the distribution. Full table of results in Supplementary Table S2.

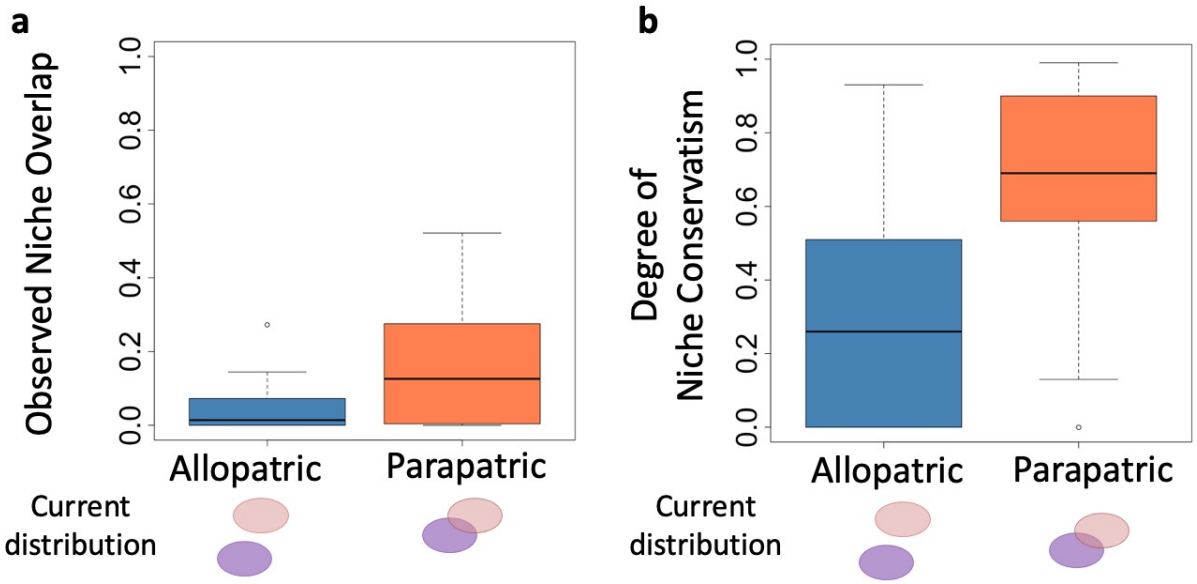

Figure S3. Comparison of observed niche overlap and degree of niche conservatism between allopatric (n=14) and parapatric sister pairs (n=14) of endemic Malagasy lizards. a, Allopatric sister pairs have significantly lower observed niche overlap compared to parapatric sister pairs ($P = 0.035$). **b**, The degree of niche conservatism in allopatric pairs tends to be higher than in parapatric pairs but differences are not statistically significantly ($P = 0.964$).

Results for comparison of species 2 with background of species1

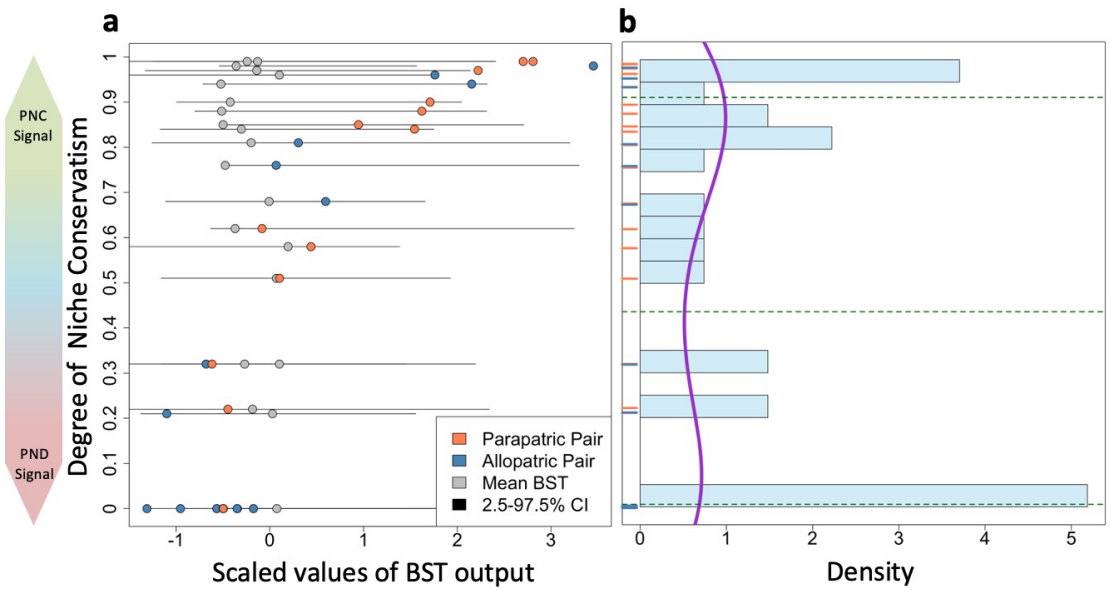

Figure S4. Bimodal distribution for the degree of niche conservatism for 28 sister pairs of endemic Malagasy lizards. a, Degree of niche conservatism against scaled values of background similarity test outputs, ordered by decreasing scaled niche overlap values for each sister pair. To illustrate the results for all sister pairs under one common scale, each set of outputs were scaled using their corresponding means and standard deviations using the scale function in R. The degree of niche conservatism of each sister pair is measured as the proportion of niche overlap replicates that are lower than the observed niche overlap (circles), so that higher values reflect a higher degree of phylogenetic niche conservatism (PNC) and lower values are associated with a lower degree of phylogenetic niche conservatism (PND). Cases where the observed niche overlap does not overlap with the 2.5-97.5% intervals of the replicate distributions have a significant ecological signal for either PNC (higher than 97.5%; 3 pairs) or PND (lower than 2.5%, 4 pairs). **b**, Density histogram showing the degree of niche conservatism for 28 pairs fitted with a kernel density line curve (purple). Hartingan's dip test supports a bimodal distribution (D = 0.12963, $P < 0.005$). Tick marks represent individual allopatric (blue) and parapatric (orange) sister pairs and dotted lines represent the upper and lower modes (0.916 and 0.006 respectively) and the antimode (0.437) of the distribution. Full table of results in Supplementary Table S2.

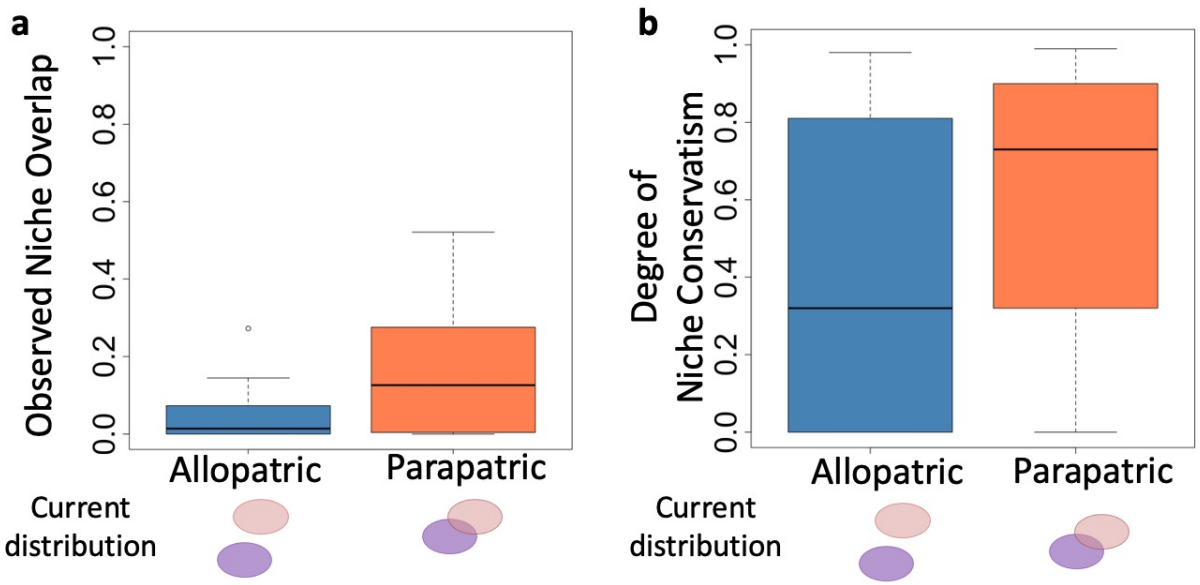

Figure S5. Comparison of observed niche overlap and degree of niche conservatism between allopatric (n=14) and parapatric sister pairs (n=14) of endemic Malagasy lizards. a, Allopatric sister pairs have significantly lower observed niche overlap compared to parapatric sister pairs ($P = 0.034$). **b**, The degree of niche conservatism in allopatric pairs tends to be higher than in parapatric pairs but differences are not statistically significantly ($P = 0.865$).