Supporting Information

Gotelli et al. 10.1073/pnas.0914089107

Null-Model Algorithms

Null-model analysis has been controversial, in part because the results depend on the assumptions of the specific null-model test, which often are difficult to evaluate (1, 2). Most null-model analyses have been based on a modified version of Connor and Simberloff's (3) original strategy of preserving observed row and column sums in the matrix. Although this algorithm originally was criticized as allegedly being too conservative (4), extensive benchmark tests with artificial matrices suggest it has good statistical properties (5–7). However, to ensure our analyses were robust and not unduly influenced by the performance of a single test, we used a suite of four null-model algorithms. These models all use the C-score (8) as the index for measuring species segregation or aggregation. Related indices, such as Stone and Robert's togetherness index (9) could be used also, although these indices have not yet been subject to benchmark testing.

Fixed-Fixed Model. The fixed-fixed model creates null matrices in which the row and column totals of the matrix are preserved (5). In the absence of additional biological or geographic information, the fixed row and column sums account for observed heterogeneity in site suitability and differences among species in colonization potential (3). To create such a matrix, we used an algorithm (5) which swaps the elements of randomly chosen 2 × 2 submatrices of the form [01 | 10] or [10 | 01]. Although the pattern of ones and zeros is randomized, each null community has the same number of species (column totals) and occupied cells (row totals) as the real avifaunal community. We created each matrix with a total of 30,000 consecutive swaps or mn swaps (where m = the number of rows in the matrix and n = the number of columns), whichever was larger. These numbers ensured that, in each randomly generated matrix, every swappable submatrix was reshuffled at least once. A unique, independent swap sequence was used for each of the 1,000 null matrices. The fixed-fixed model, when used with the C-score, has been subjected to extensive benchmark testing with artificial matrices that contain specified amounts of randomness and structure (5–7). The swapping algorithm that we have used to create null matrices is slightly less likely to detect segregation of species than is a more recent algorithm that samples all matrices with the same row and column totals equiprobably (10). However, this bias is small for large matrices of the size we have analyzed here (11).

Habitat Model. A potential weakness of the fixed-fixed model is that it does not directly simulate a random colonization process. To address this deficiency, we used a model in which the row totals of the matrix (the occurrence frequency of species) were fixed, but the column totals (the number of species per cell) were not. Most importantly, species were assigned randomly and independently to cells with the probability of occurrence set proportional to the measured index of habitat diversity (HD) for each cell.

Population Model. To reflect the natural differences in habitat diversity among cells and colonization potential among species, we constructed a population-null model in which the total number of species occurrences in the matrix was preserved, but where row and column total were allowed to vary randomly (5). The probability of an occurrence of species i in cell j was proportional to both the total breeding population size of species i in Denmark and the HD value of cell j. Thus, for the placement of the first species occurrence in the matrix, the cell most likely to be chosen would occur at the intersection of the row with the largest sum (species with largest population) and column with the largest sum (grid cell with the highest HD value). The least likely cell to be chosen would be the one with the smallest row and smallest column sum.

Biomass Model. The biomass model was identical to the population model, except that biomass (total biomass of breeding individuals in Denmark) was substituted for population size. The rational for this model is that total biomass reflects the total energy that has been sequestered by the species in Denmark, integrating the effects of both population size and body size. Because it is difficult to validate or parameterize null models for entire assemblages, our strategy was to test a suite of null models applied to different spatial scales and different levels of assemblage organization. Consistent results that emerge from such a battery of tests yield robust findings that are insensitive to the assumptions and restrictions that may apply to any particular null model or data partition.

Body masses were compiled from the *Handbook of the Birds of Europe, the Middle East, and North Africa* (12–20), with a preference for data from Danish, Dutch, and northern German populations. A priori guild assignments were made by C. Rahbek and J. Fjeldså before co-occurrence analyses were performed.

Census Data for Danish Avifauna

Species occurrence records for the 197 breeding birds of Denmark were derived from data in the Danish atlas of breeding birds in Denmark, 1993–1996 (21). Denmark was divided into 2,169 atlas-cells (5 km × 5 km). More than 99% of the cells were surveyed for breeding birds: 1,465 were well surveyed, 640 were reasonably well surveyed, 50 were incompletely surveyed, and only 14 cells were not surveyed. Atlas surveys were conducted by *ca*. 750 observers. The total number of observations (cells × species occurrences) equaled 141,865. Each cell was visited 5–10 times each year for a quantitative census of all breeding species. Field work was conducted between February and August during each of the 4 years (1993–1996).

Data derived from the atlas surveys were supplemented with information from census records from 2,500 large nature reserves, ongoing single-species surveys, and monitoring or research programs on rare and/or endangered species, wildfowl, and raptors. Incidental information on rare breeding species also was included from published maps on occurrences of Danish breeding birds.

The occurrence of each species in a cell was categorized as (*i*) confirmed breeding (e.g., observation of adults feeding chicks, occurrence of freshly used nests, and/or adult birds carrying food or excrement); (*ii*) probable breeding (e.g., territorial singing males observed in the breeding season, individuals observed defending territories, engaging in courtship, building nests, or carrying nesting materials); and (*iii*) presence observed (e.g., birds were observed in the breeding season but with no other evidence of breeding). In our analyses, we used only records from the first two categories to designate species occurrences.

1. Gotelli NJ, Graves GR (1996) *Null Models in Ecology* (Smithsonian Institution Press, Washington, DC).
2. Ladau J (2008) Validation of null model tests using Neyman-Pearson hypothesis testing theory. *Theoretical Ecology* 1:241–248.
3. Connor EF, Simberloff D (1979) The assembly of species communities: Chance or competition? *Ecology* 60:1132–1140.
4. Diamond JM, Gilpin ME (1982) Examination of the "null" model of Connor and Simberloff for species co-occurrences on islands. *Oecologia* 52:64–74.
5. Gotelli NJ (2000) Null model analysis of species co-occurrence patterns. *Ecology* 81: 2606–2621.
6. Ulrich W, Gotelli NJ (2007) Null model analysis of species nestedness patterns. *Ecology* 88:1824–1831.

7. Gotelli NJ, Ulrich W (2010) The empirical Bayes approach as a tool to identify non-random species associations. *Oecologia* 162:463–477.
8. Stone L, Roberts A (1990) The checkerboard score and species distributions. *Oecologia* 85:74–79.
9. Stone L, Roberts A (1992) Competitive exclusion, or species aggregation? An aid in deciding. *Oecologia* 91:419–424.
10. Miklós I, Podani J (2004) Randomization of presence–absence matrices: Comments and new algorithms. *Ecology* 85:86–92.
11. Lehsten V, Harmand P (2006) Null models for species co-occurrence patterns: Assessing bias and minimum iteration number for the sequential swap. *Ecography* 29:786–792.
12. Cramp S, Perrins CM (1994) *Handbook of the Birds of Europe, the Middle East and North Africa: The Birds of the Western Palearctic. Buntings and New World Warblers* (Oxford Univ Press, Oxford, U.K.), Vol 9.
13. Cramp S, Perrins CM (1994) *Handbook of the Birds of Europe, the Middle East and North Africa: The Birds of the Western Palearctic. Crows to Finches* (Oxford Univ Press, Oxford, U.K.), Vol 8.
14. Cramp S, Perrins CM (1993) *Handbook of the Birds of Europe, the Middle East and North Africa: The Birds of the Western Palearctic. Flycatchers to Shrikes* (Oxford Univ Press, Oxford, U.K.), Vol 7.
15. Cramp S (1992) *Handbook of the Birds of Europe, the Middle East and North Africa: The Birds of the Western Palearctic. Warblers* (Oxford Univ Press, Oxford, U.K.), Vol 6.
16. Cramp S (1988) *Handbook of the Birds of Europe, the Middle East and North Africa: The Birds of the Western Palearctic. Tyrant Flycatchers to Thrushes* (Oxford Univ Press, Oxford, U.K.), Vol V.
17. Cramp S (1985) *Handbook of the Birds of Europe, the Middle East and North Africa: The Birds of the Western Palearctic. Terns to Woodpeckers* (Oxford Univ Press, Oxford, U.K.), Vol 4.
18. Cramp S, Simmons KEL (1983) *Handbook of the Birds of Europe, the Middle East and North America: The Birds of the Western Palearctic. Waders to Gulls* (Oxford Univ Press, Oxford, U.K.), Vol 3.
19. Cramp S, Simmons KEL (1980) *Handbook of the Birds of Europe, the Middle East and North Africa: The Birds of the Western Palearctic Hawks to Bustards* (Oxford Univ Press, Oxford, U.K.), Vol 2.
20. Cramp S, Simmons KEL (1977) *Handbook of the Birds of Europe, the Middle East and North Africa: The Birds of the Western Palearctic. Ostrich to Ducks* (Oxford Univ Press, Oxford, U.K.), Vol 1.
21. Grell MB (1998) *Fuglenes Danmark* (Gads Forlag, Copenhagen, Denmark).

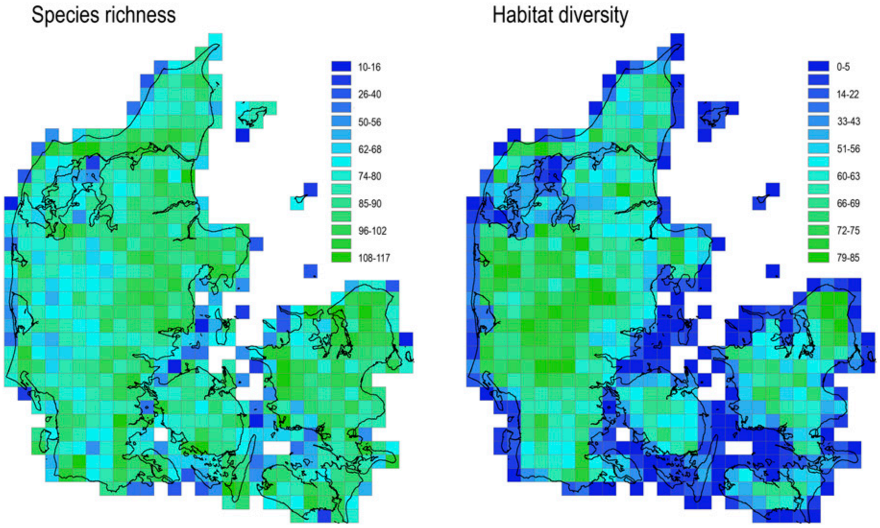

Fig. S1. Species richness and habitat diversity (100-km^2 grain size). Species richness of Danish breeding birds and spatial variation in habitat diversity (HD) of grid cells at a grain size of 10 km × 10 km (100 km^2). The HD score is the product of relative grid cell area and the probability that two points randomly chosen within a grid cell represent different habitat types (1). The HD score was used to parameterize null models of random species colonization independently. Species richness ranged from 10 to 117 species per cell (average = 81.45). The best-fitting power function was $S = 48.17259(HD)^{0.1468}$, $r^2 = 0.4233$.

1. Grell MB (1998) *Fuglenes Danmark* (Gads Forlag, Copenhagen, Denmark).

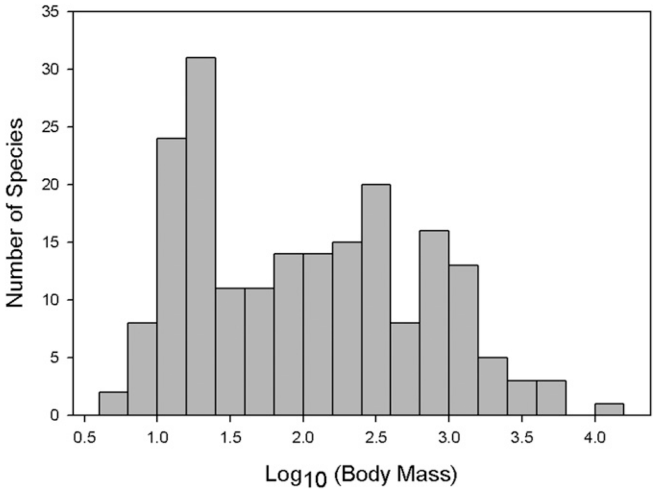

Fig. S2. Individual body mass. Distribution of body masses of the Danish avifauna (*n* = 197 species).

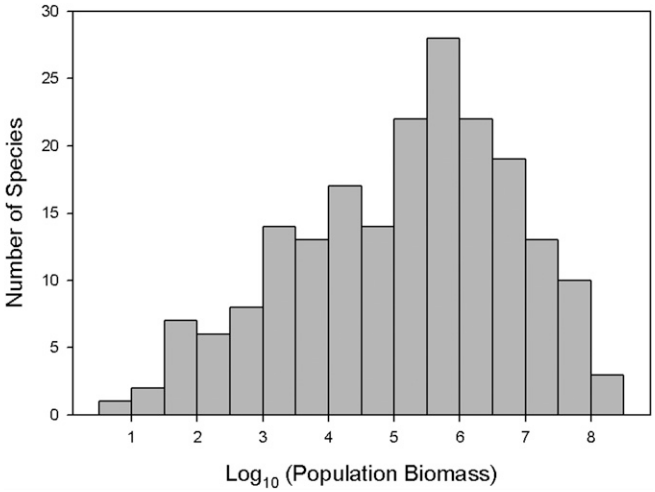

Fig. S3. Population biomass. Distribution of population biomass (individual body mass × population size) for the Danish avifauna (*n* = 197 species).

Other Supporting Information

- Table S1 (PDF)
- Table S2 (PDF)
- Table S3 (PDF)
- Table S4 (PDF)
- Table S5 (PDF)
- Table S6 (PDF)
- Table S7 (PDF)