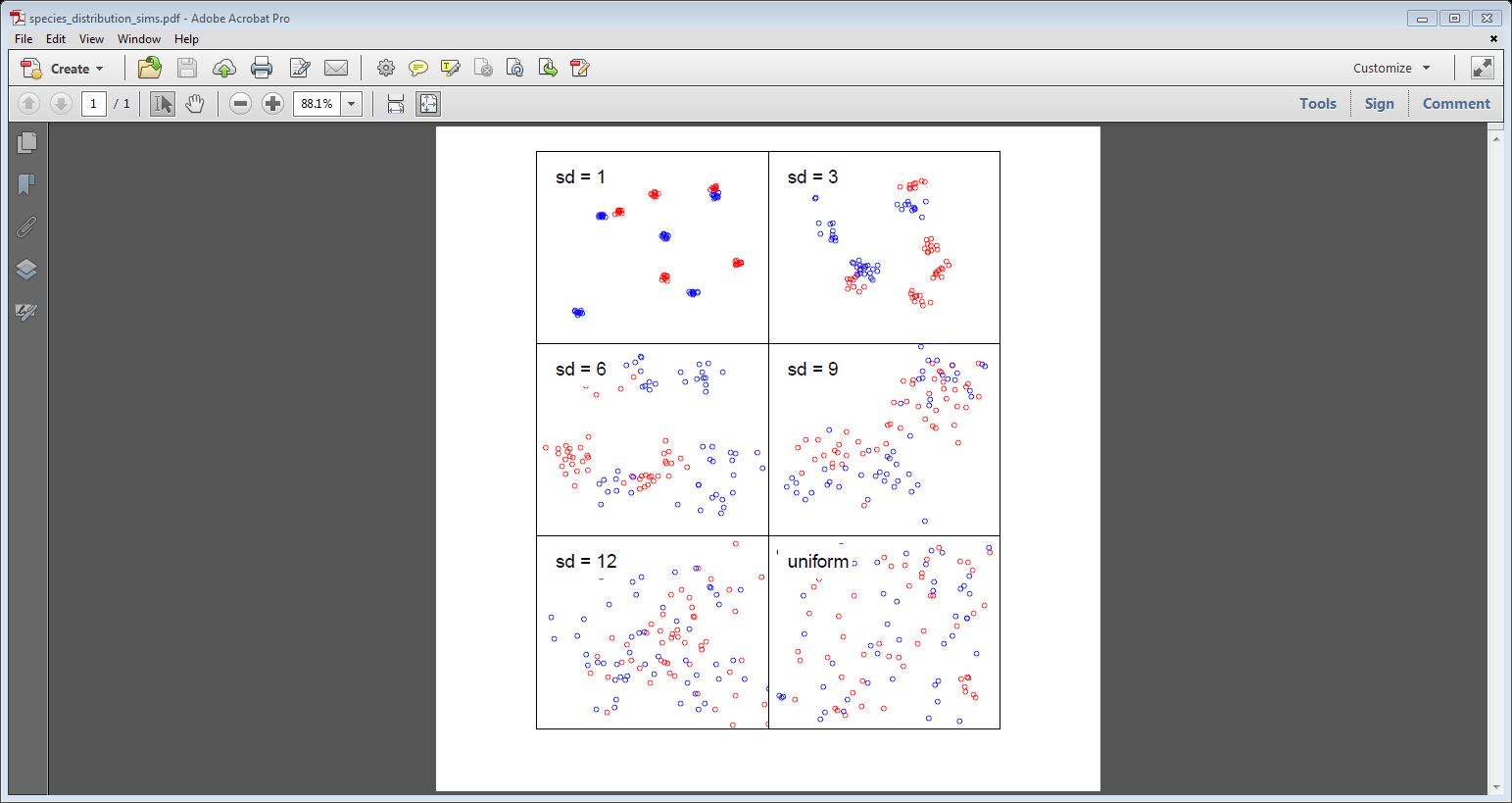
**Supporting Information**

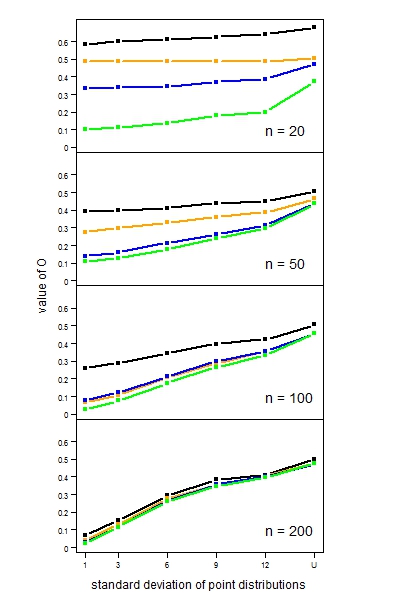
Appendix S1. Illustration of different levels of species overlap generated for tests of the point-proximity metric *O*.

To generate the distributions of two species, we began by randomly placing *n* focal points on a space. For each focal point, we placed *m* species occurrence points at a distance from the focal point drawn from a bivariate normal distribution with x and y coordinates of the focal point as the mean, and a standard deviation *σ*. This was repeated for each species. The value of σ determines the degree of clumping of species distributions, and hence the degree of overlap of two species. In this figure, we give an example of the degree of clumping for two species (red and blue), generated with *n* = 10, *m* = 20, and *σ* values of 1, 3, 6, 9, and 12, together with a set of points with xy coordinates drawn randomly from a uniform distribution.



Appendix S2. Simulation of values of the point-proximity metric *O*.

The plots show mean values of *O* from 500 simulations of two species’ occurrence points. The four boxes show simulations of different total numbers of points (n = 20, 50, 100, 200). Within each box, *O* values under different degrees of species clumping (σ= 1, 3, 6, 9, 12, plus uniformly-distributed points) are shown. The different coloured lines represent different levels of asymmetry in the distribution of points among the two species (green: equal numbers; blue, orange and black: one species has 75%, 90%, and 95% of points, respectively).



Comparison of lines of the same colour across the four panels gives an indication of the effect of random undersampling on the behaviour of *O*. Assuming there are 200 occurrences of the two species, the four panels represent 10%, 25%, 50% and 100% sampling, respectively. When numbers of occurrences are equal for the two species (green lines), undersampling has relatively little effect on the values of *O*. When occurrences are asymmetric and undersampling is severe, values of *O* are elevated, so that the degree of spatial proximity between two species is overestimated.