Supplementary Figures for: Inferring continuous and discrete population genetic structure across space

Gideon S. Bradburd[1,*], Graham M. Coop[2,☯], Peter L. Ralph[3,☯],

1 Ecology, Evolutionary Biology, and Behavior Graduate Group Department of Integrative Biology, Michigan State University, MI 48824
2 Center for Population Biology, Department of Evolution and Ecology, University of California, Davis, CA 95616
3 Institute of Ecology and Evolution, Departments of Mathematics and Biology, University of Oregon, Eugene, OR 97403

*bradburd@msu.edu

☯These authors contributed equally to this work.

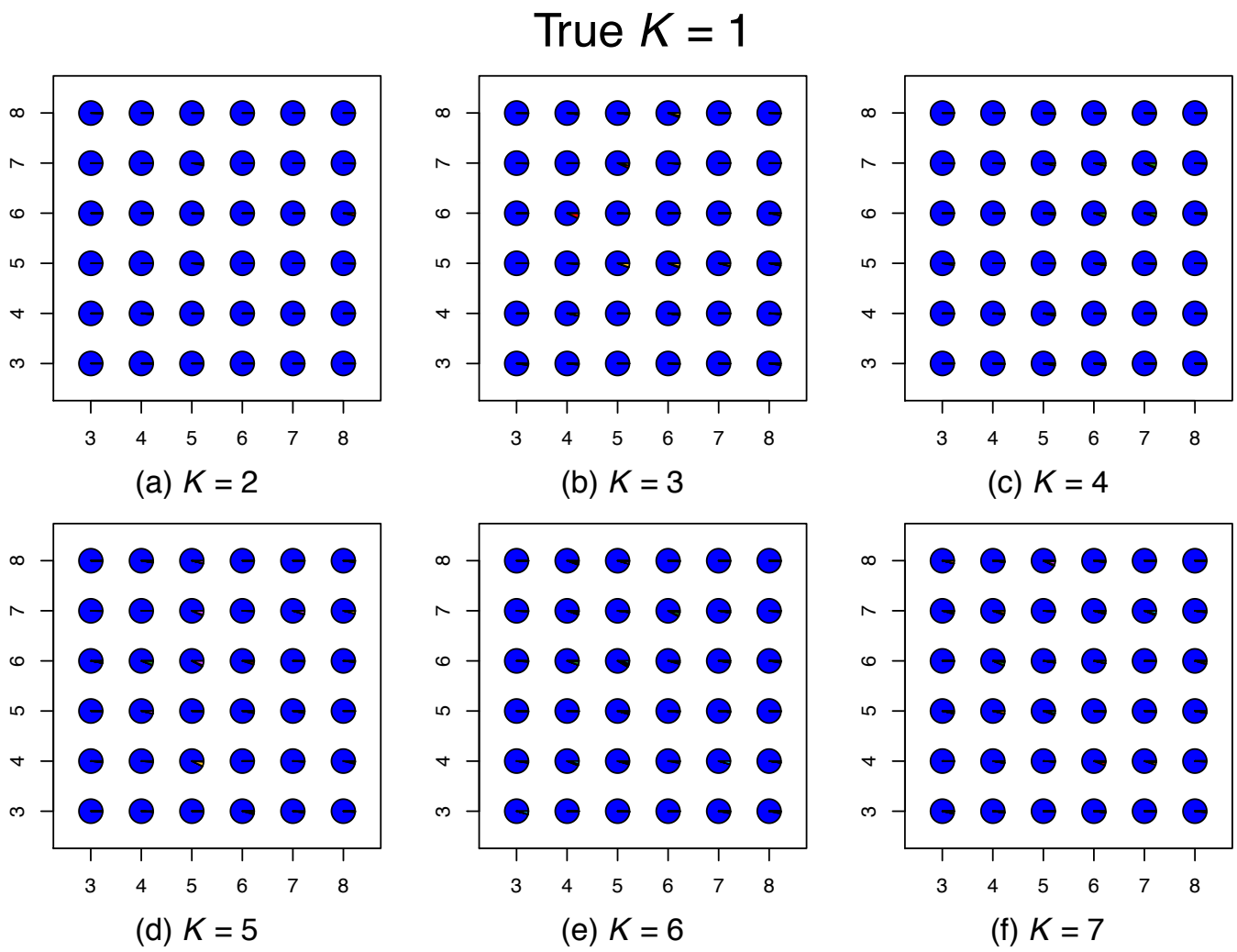

Figure S1 Map of admixture proportions estimated using a spatial `conStruct` model for $K = 2$ through 7. The data were simulated using one layer with nearest-neighbor symmetric migration between demes.

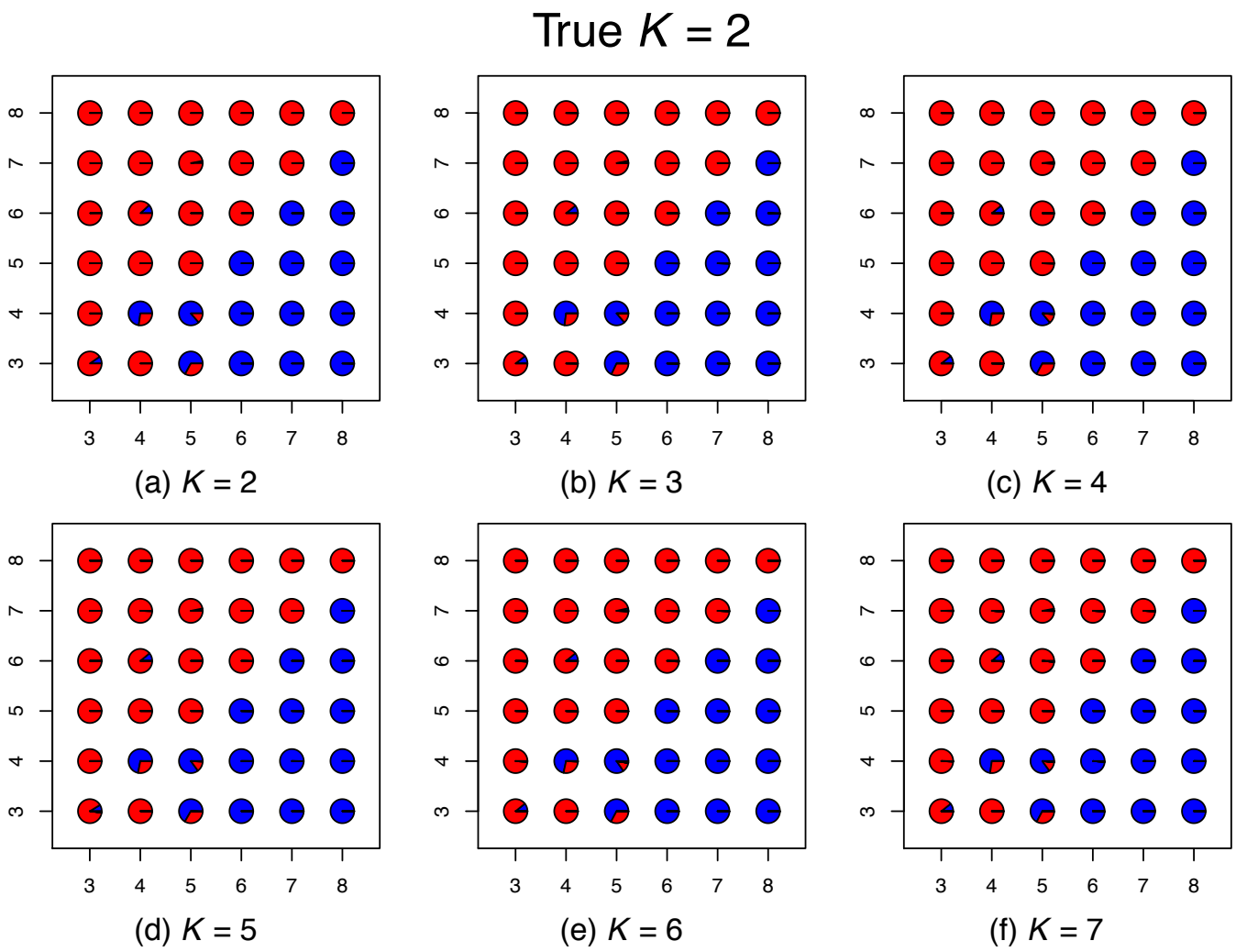

Figure S2 Map of admixture proportions estimated using a spatial `conStruct` model for $K = 2$ through 7. The data were simulated using two layers with nearest-neighbor symmetric migration between demes.

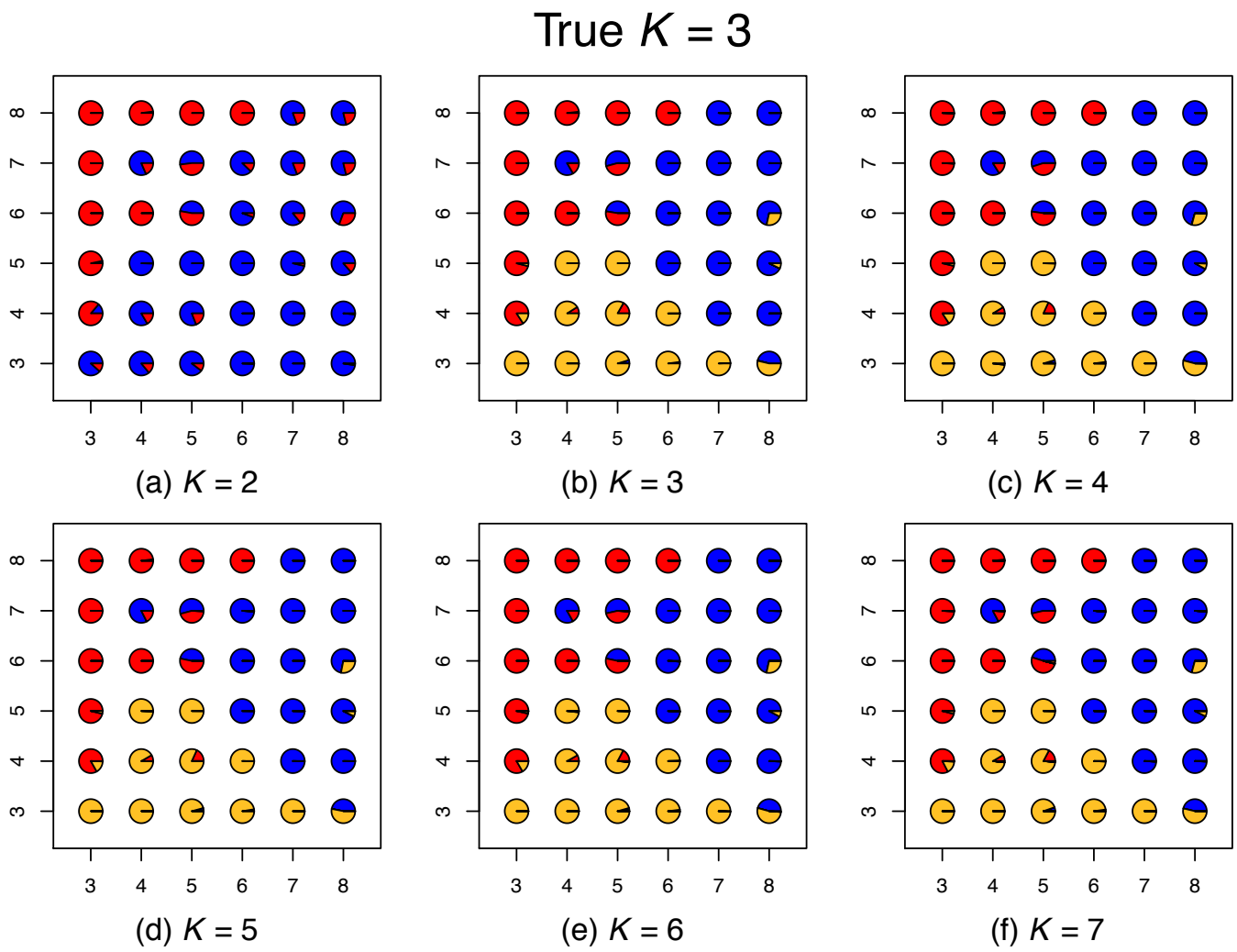

Figure S3 Map of admixture proportions estimated using a spatial `conStruct` model for $K = 2$ through 7. The data were simulated using three layers with nearest-neighbor symmetric migration between demes.

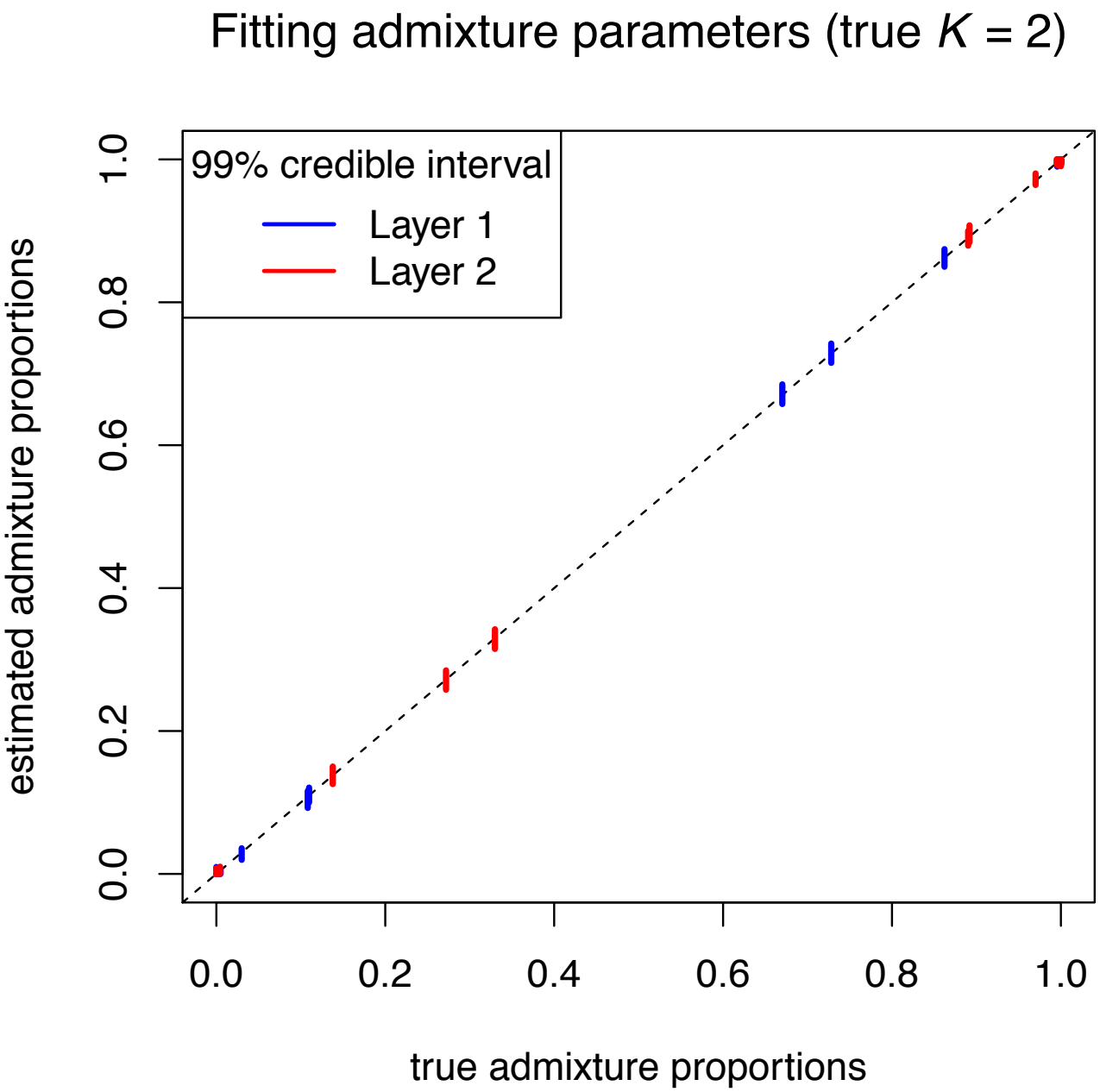

Figure S4 Plot of `conStruct`'s ability to correctly estimate admixture proportions on simulated data, from an analysis with a spatial model using $K = 2$. The horizontal axis shows the admixture proportions used to simulate the data, and the vertical axis shows the 99% credible intervals for those proportions as reported by conStruct.

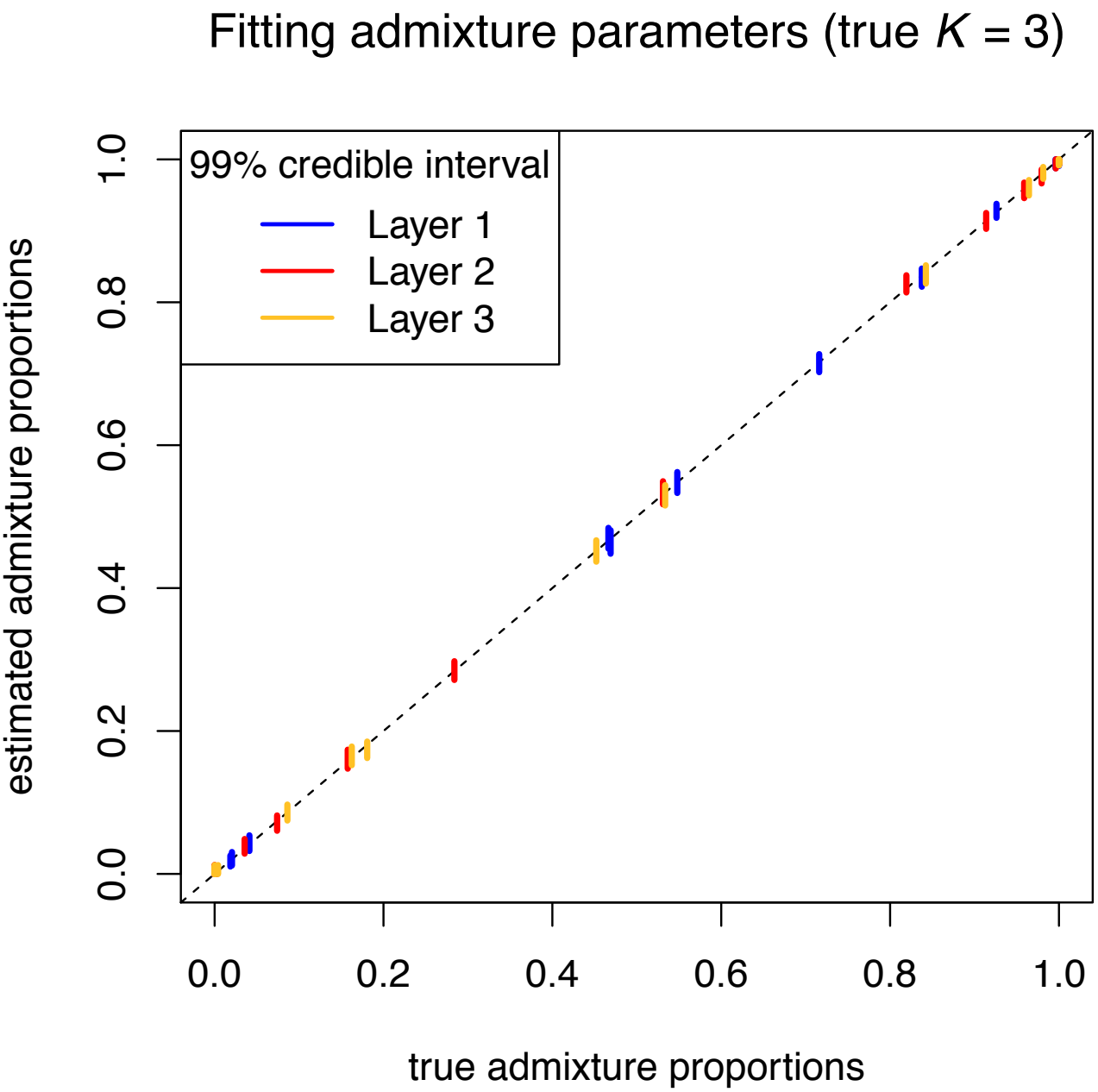

Figure S5 Plot of `conStruct` ability to correctly estimate admixture proportions on simulated data, from an analysis with a spatial model using $K = 3$. The horizontal axis shows the admixture proportions used to simulate the data, and the vertical axis shows the 99% credible intervals for those proportions as reported by conStruct.

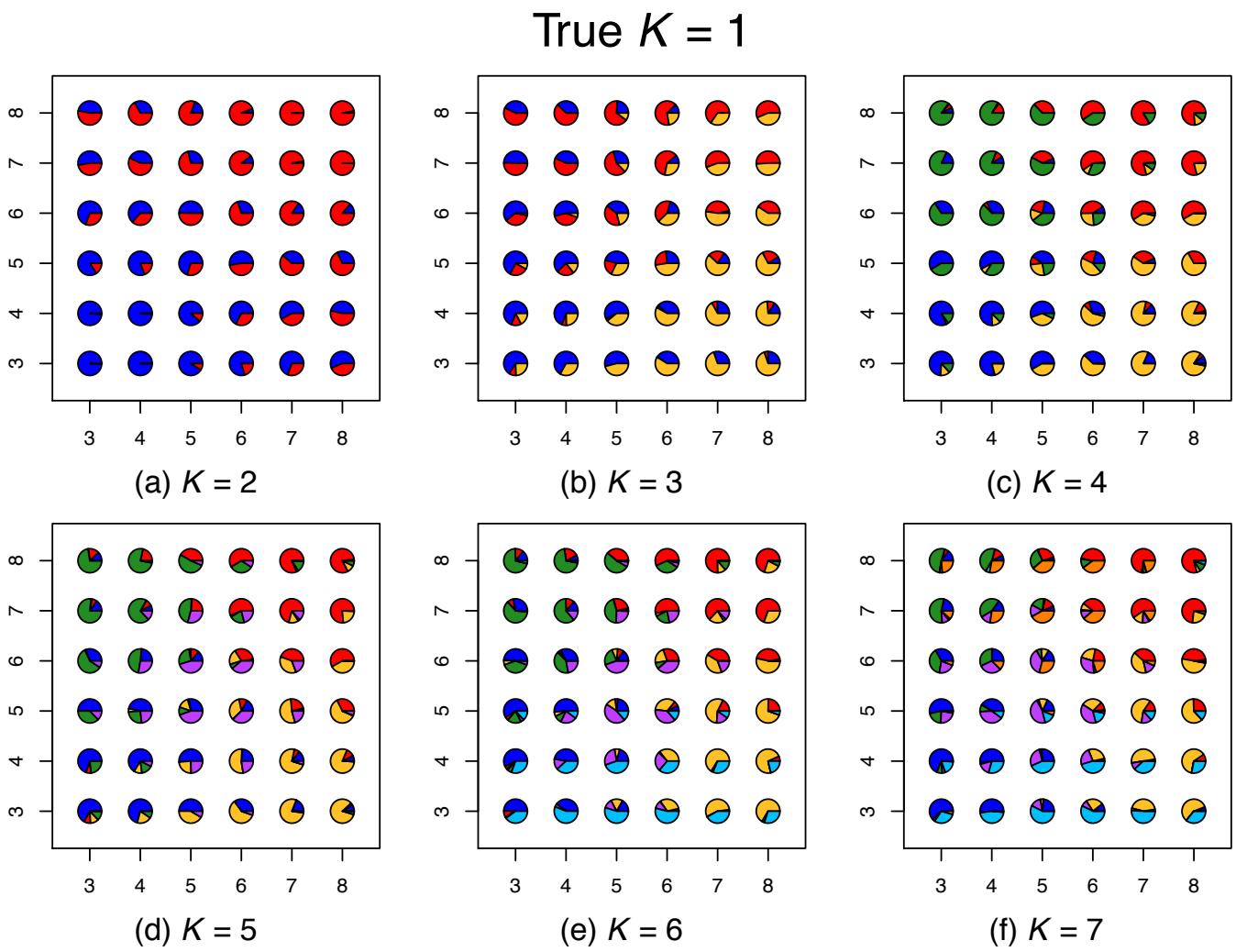

Figure S6 Map of admixture proportions estimated using a nonspatial `conStruct` model for $K = 2$ through 7. The data were simulated using one layer with nearest-neighbor symmetric migration between demes.

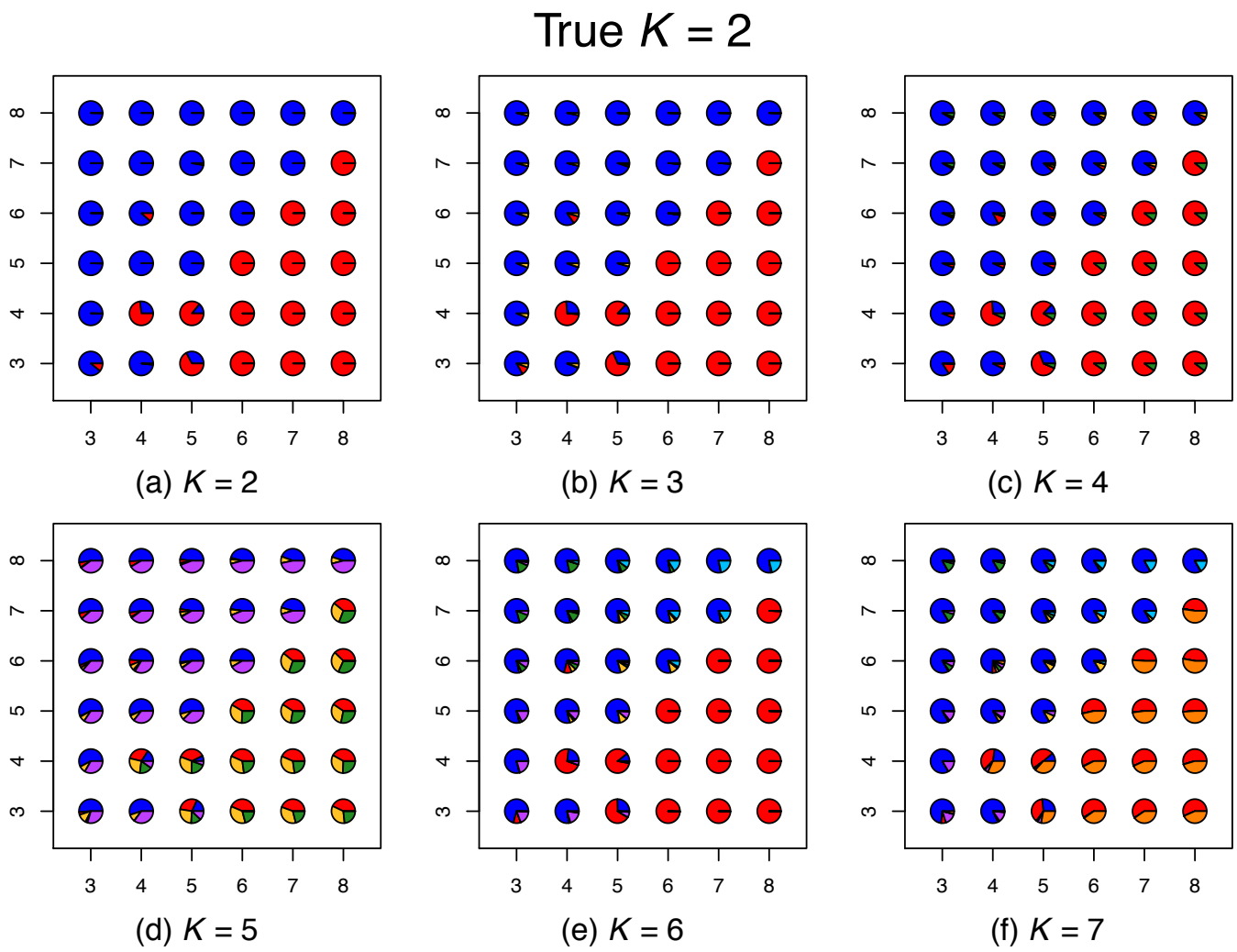

Figure S7 Map of admixture proportions estimated using a nonspatial `conStruct` model for $K = 2$ through 7. The data were simulated using two layers with nearest-neighbor symmetric migration between demes.

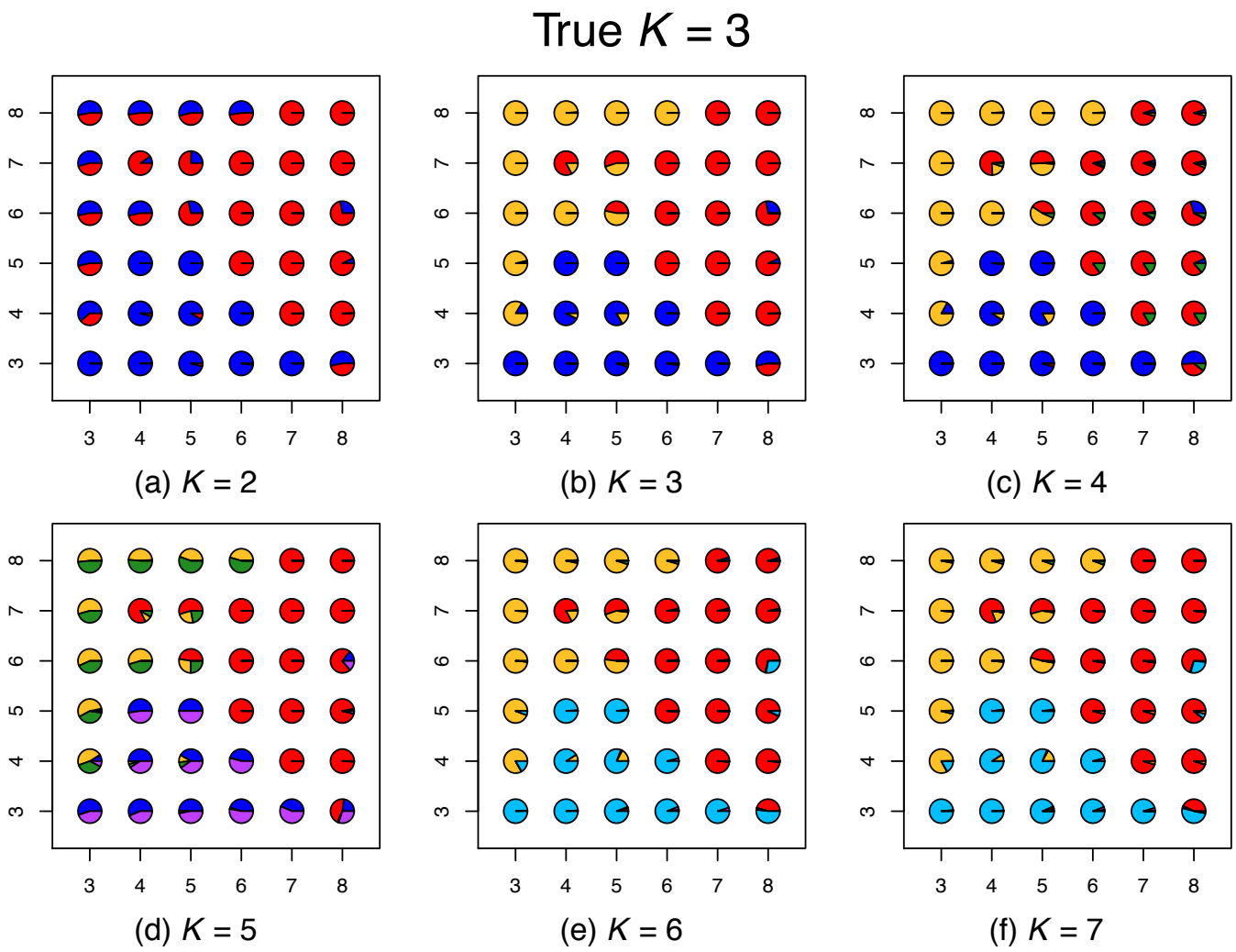

Figure S8 Map of admixture proportions estimated using a nonspatial `conStruct` model for $K = 2$ through 7. The data were simulated using three layers with nearest-neighbor symmetric migration between demes.

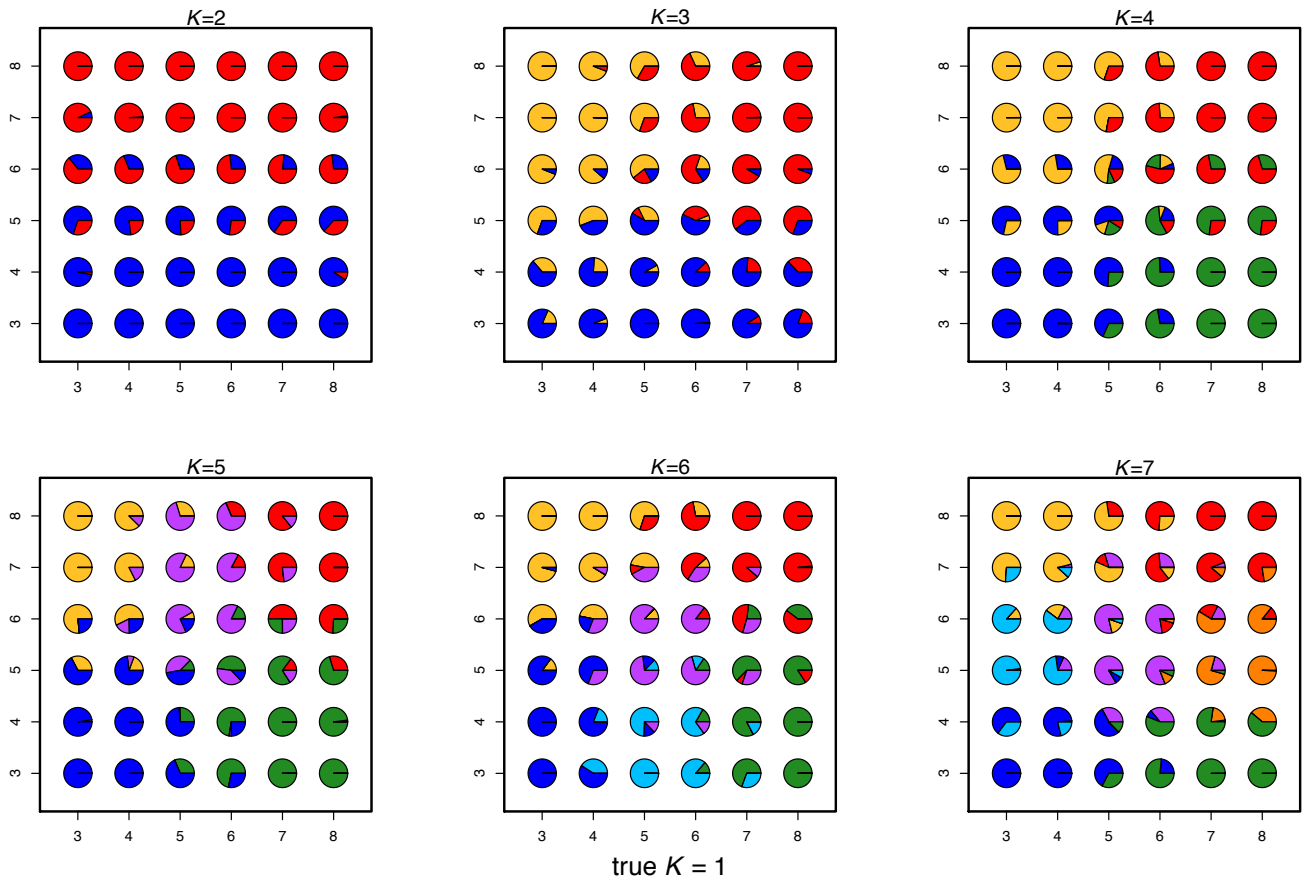

Figure S9 Map of admixture proportions estimated using ADMIXTURE Alexander *et al.* (2009) for $K = 2$ through 7. The data were simulated using one layer with nearest-neighbor symmetric migration between demes. The true value was $K = 1$, but the model with the lowest cross-validation error (i.e., the preferred model) was $K = 7$.

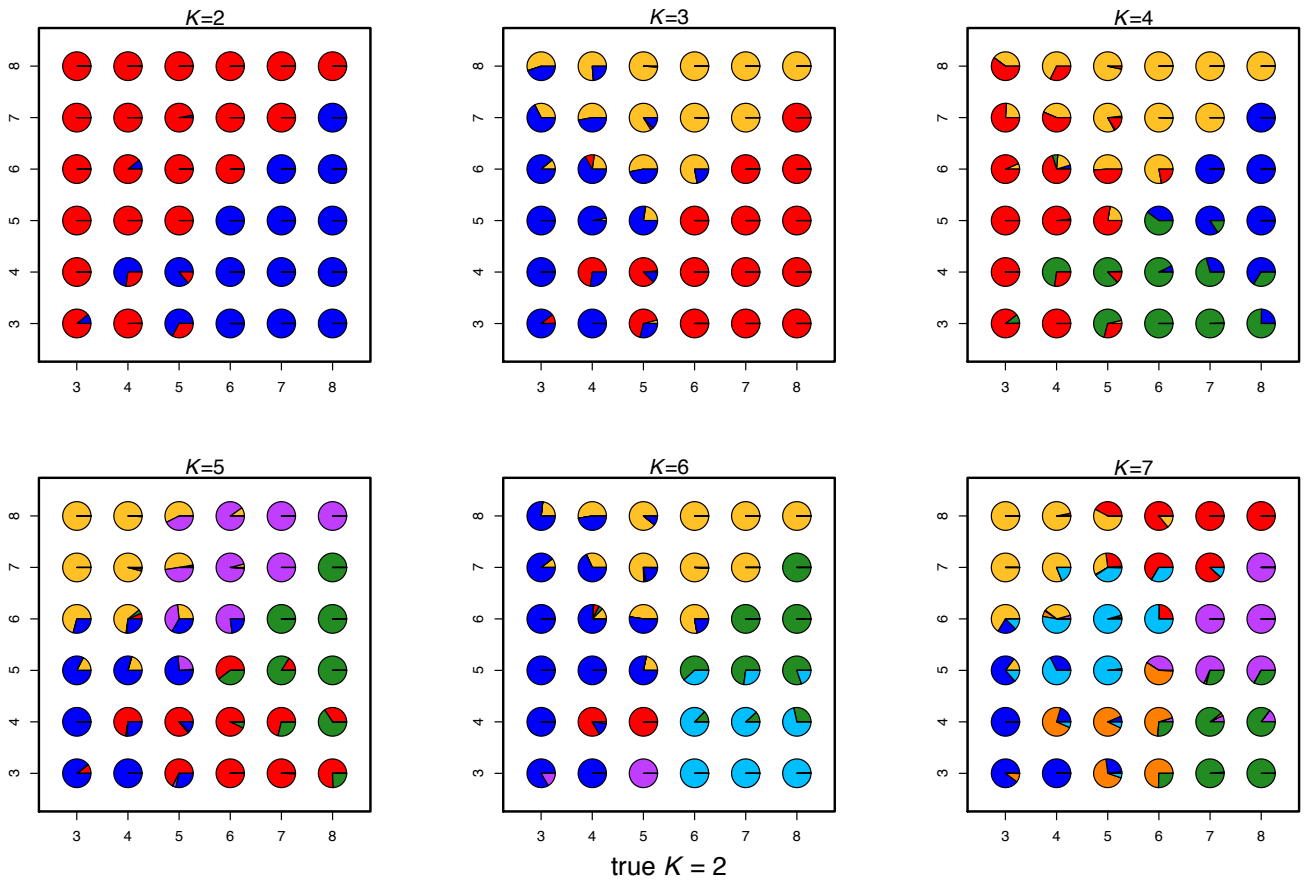

Figure S10 Map of admixture proportions estimated using ADMIXTURE Alexander *et al.* (2009) for $K = 2$ through 7. The data were simulated using two layers with nearest-neighbor symmetric migration between demes. The true value was $K = 2$, but the model with the lowest cross-validation error (i.e., the preferred model) was $K = 7$.

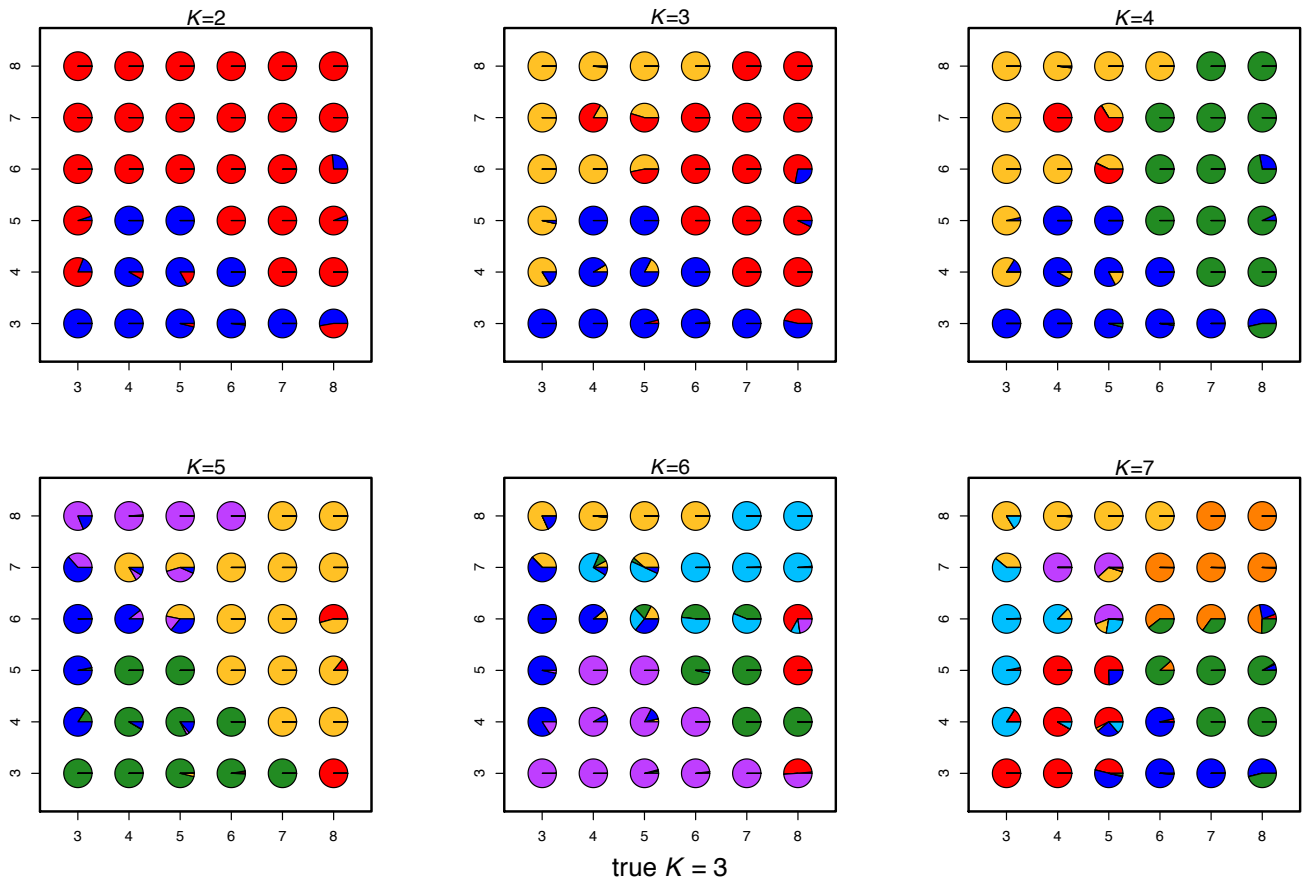

Figure S11 Map of admixture proportions estimated using ADMIXTURE Alexander *et al.* (2009) for $K = 2$ through 7. The data were simulated using three layers with nearest-neighbor symmetric migration between demes. The true value was $K = 3$, but the model with the lowest cross-validation error (i.e., the preferred model) was $K = 7$.

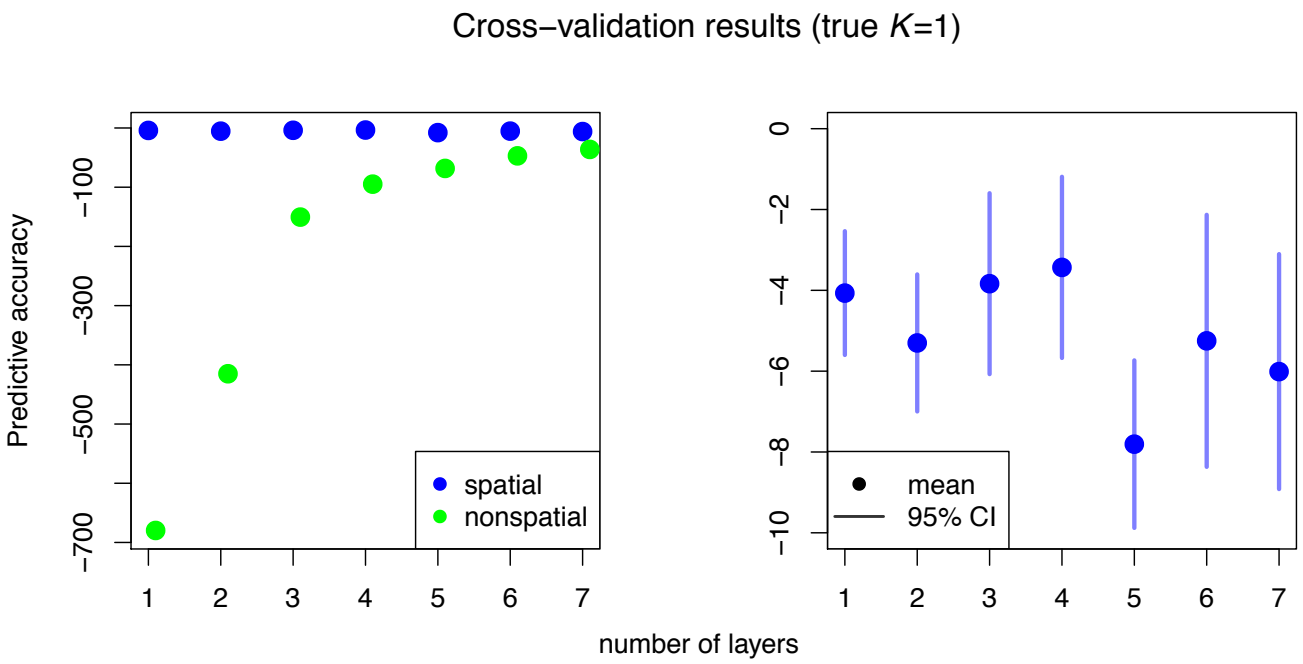

Figure S12 Cross-validation results for data simulated under $K = 1$, comparing the spatial and nonspatial `conStruct` models run with $K = 1$ through 7. The right panel zooms in on just the spatial cross-validation results.

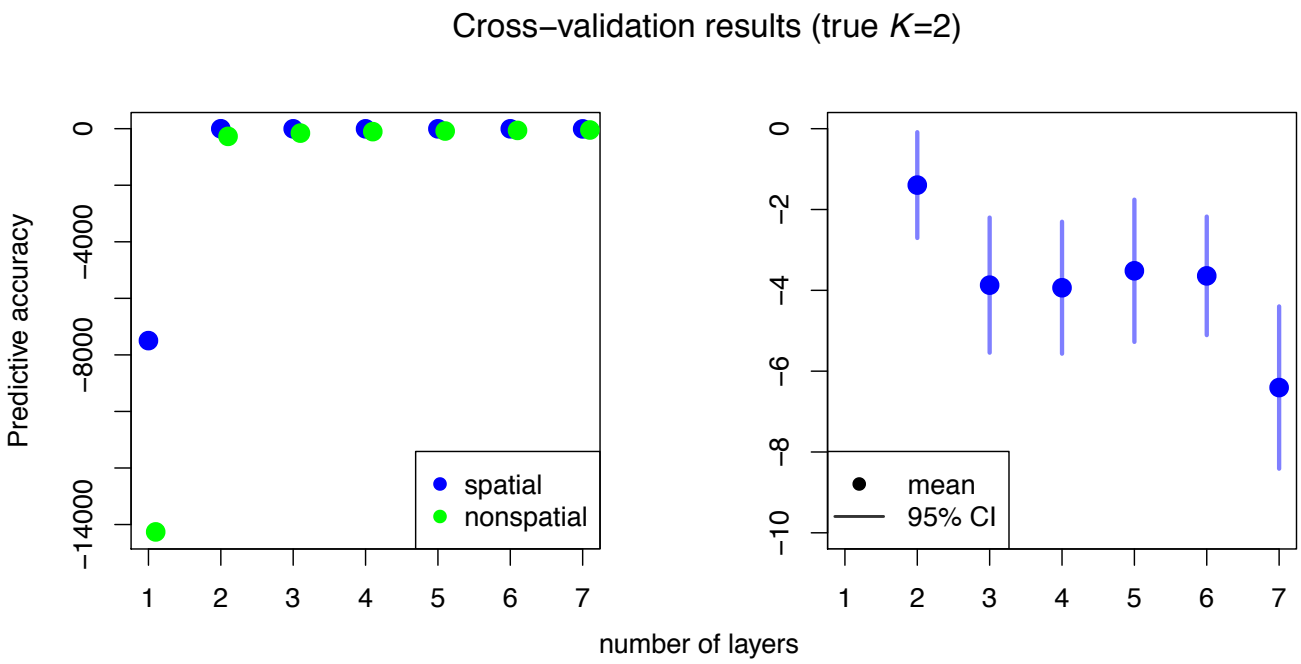

Figure S13 Cross-validation results for data simulated under $K = 2$, comparing the spatial and nonspatial `conStruct` models run with $K = 1$ through 7. The right panel zooms in on just the spatial cross-validation results.

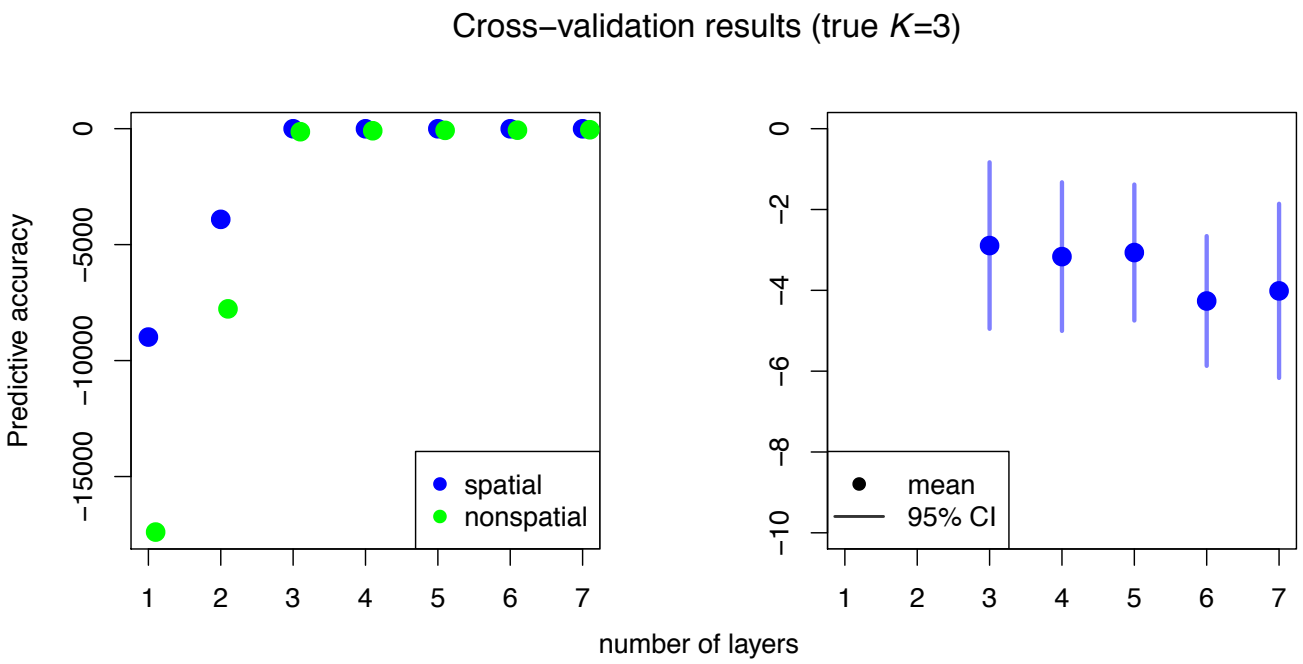

Figure S14 Cross-validation results for data simulated under $K = 3$, comparing the spatial and nonspatial `conStruct` models run with $K = 1$ through 7. The right panel zooms in on just the spatial cross-validation results.

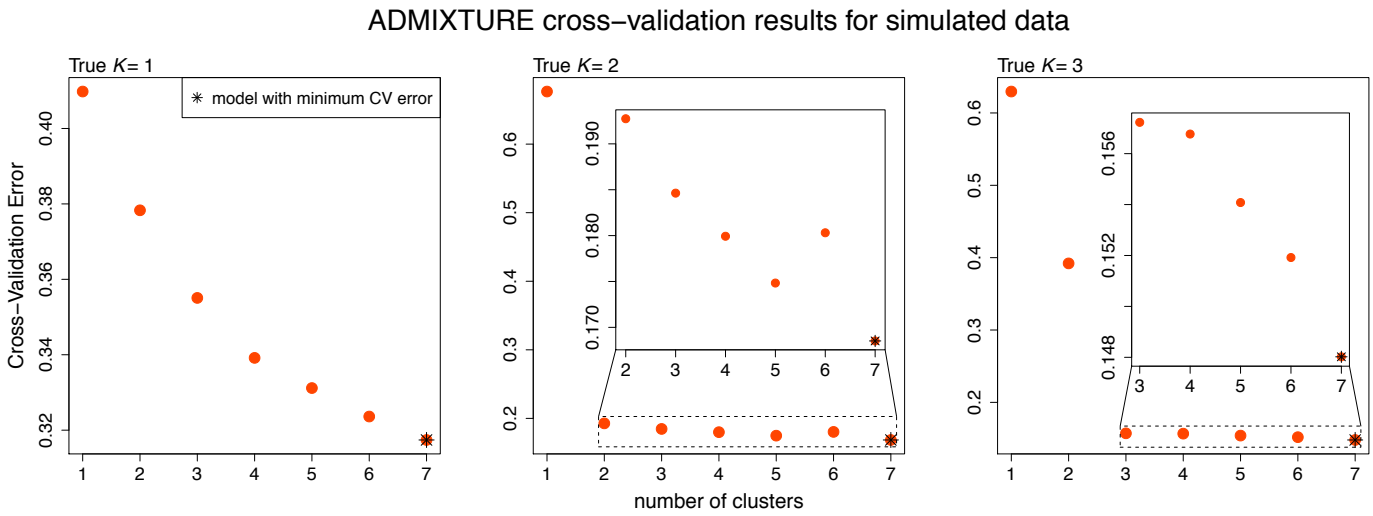

Figure S15 ADMIXTURE cross-validation results for data simulated under $K = 1$, $K = 2$, and $K = 3$, run with $K = 1$ through 7 using 50 data folds (`-cv=50`). The inset plots zoom in on cross-validation results outlined in the dotted boxes. The preferred model (with the lowest cross-validation error) is highlighted with an asterisk.

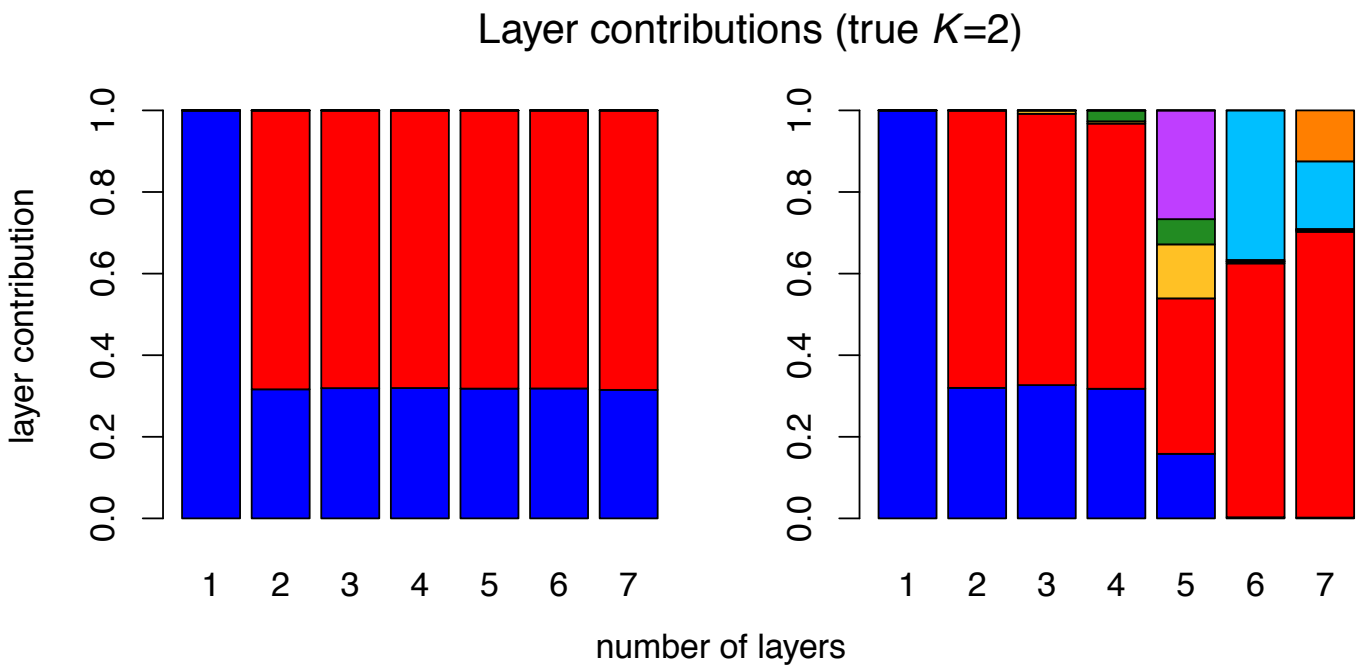

Figure S16 Layer/cluster contributions (i.e., how much total covariance is contributed by each layer/cluster), for all layers estimated in runs using $K = 1$ through 7 for the spatial model (left), and for all clusters using the nonspatial `conStruct` model (right). Data were simulated using $K = 2$. For each value of K along the x-axis, there are an equal number of contributions plotted.

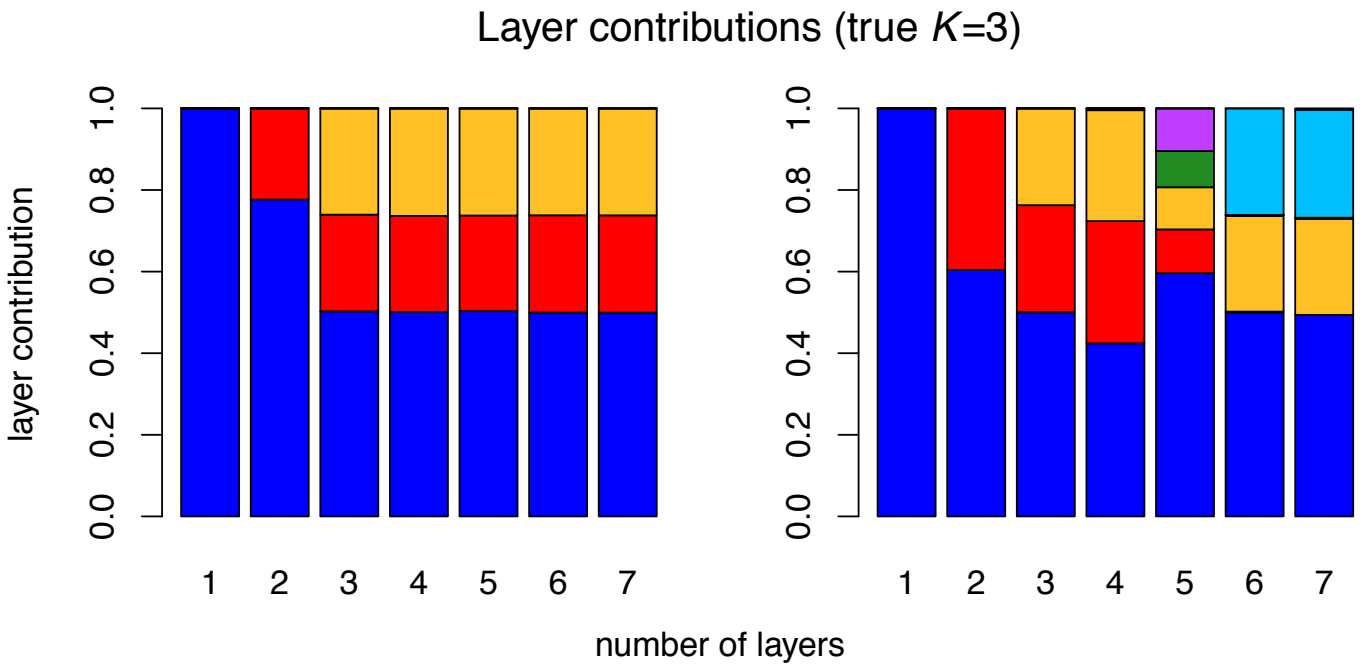

Figure S17 Layer/cluster contributions (i.e., how much total covariance is contributed by each layer/cluster), for all layers estimated in runs using $K = 1$ through 7 for the spatial model (left), and for all clusters using the nonspatial `conStruct` model (right). Data were simulated using $K = 3$. For each value of K along the x-axis, there are an equal number of contributions plotted.

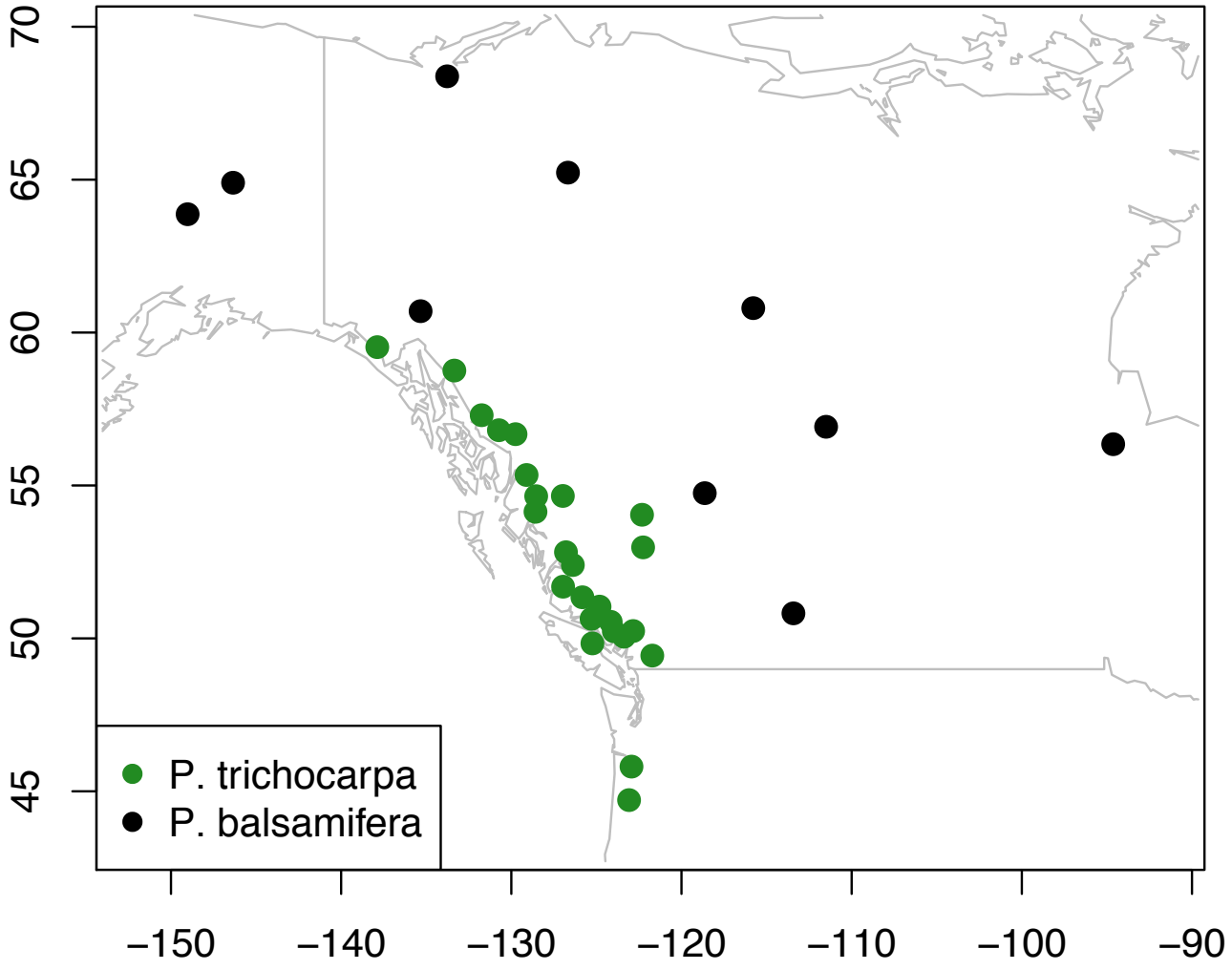

Figure S18 Map of the sampled *Populus* populations included in the analysis. The color of the sampling location denotes the putative species.

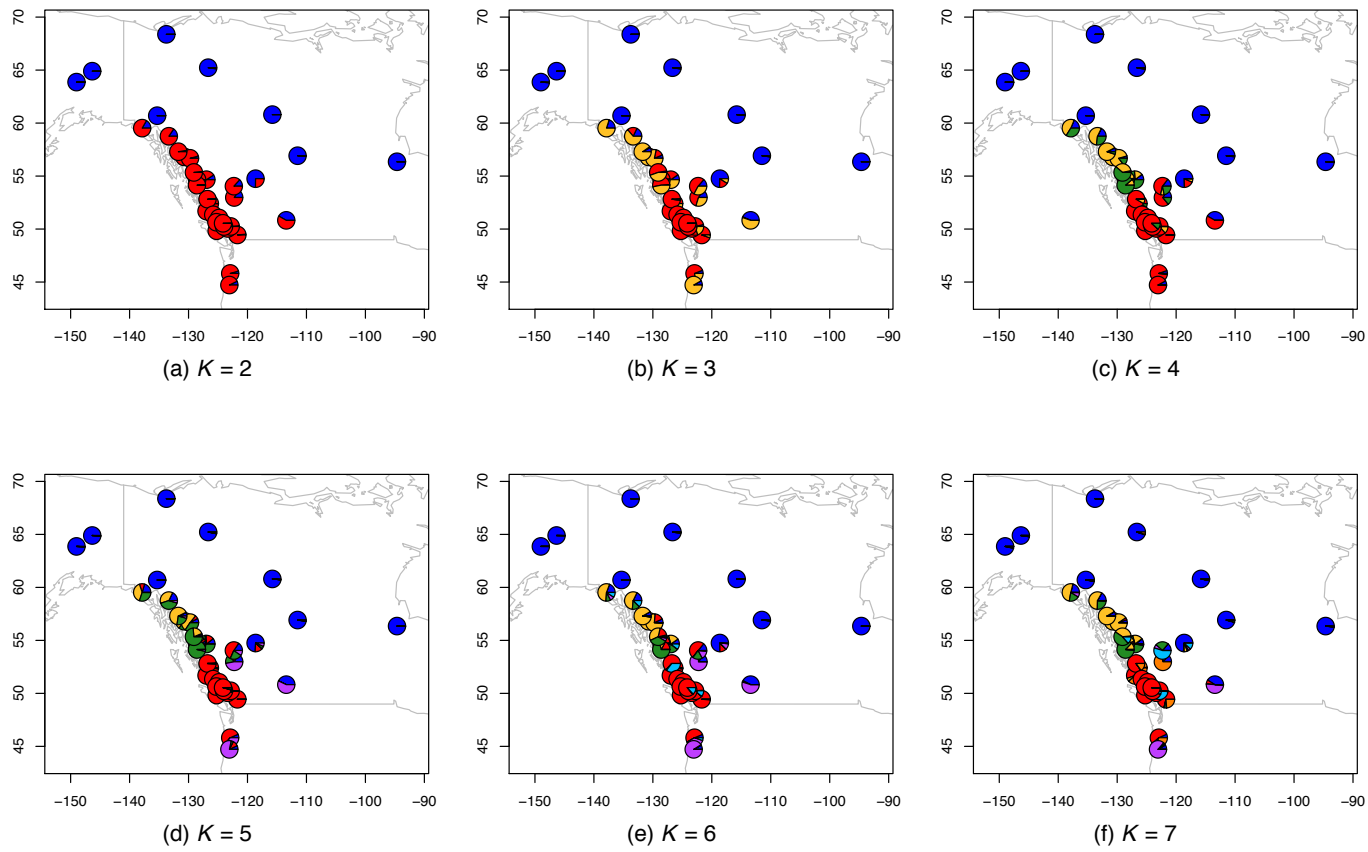

Figure S19 Maps of admixture proportions estimated for the *Populus* dataset using the spatial `conStruct` model for $K = 2$ through 7.

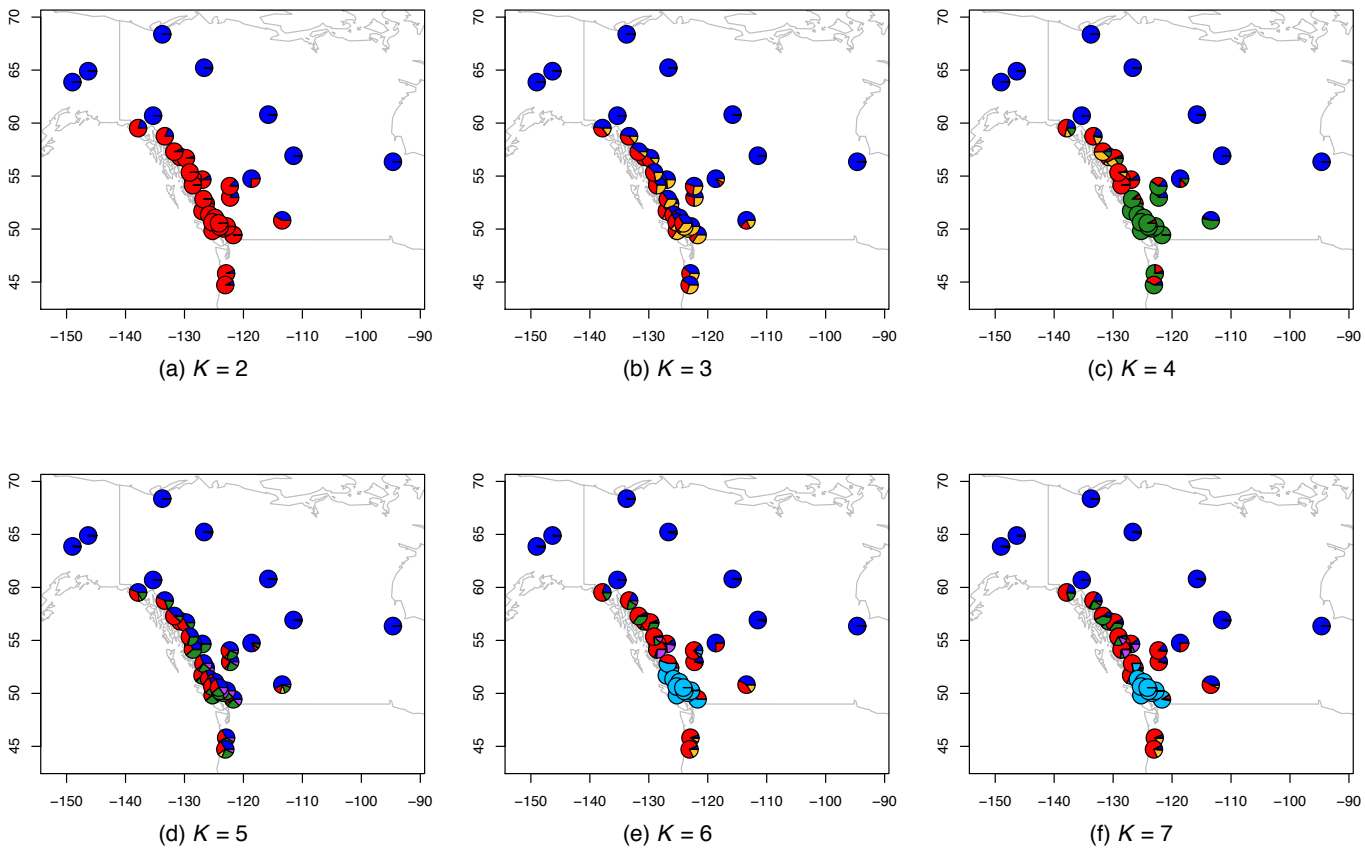

Figure S20 Maps of admixture proportions estimated for the *Populus* dataset using the nonspatial `conStruct` model for $K = 2$ through 7.

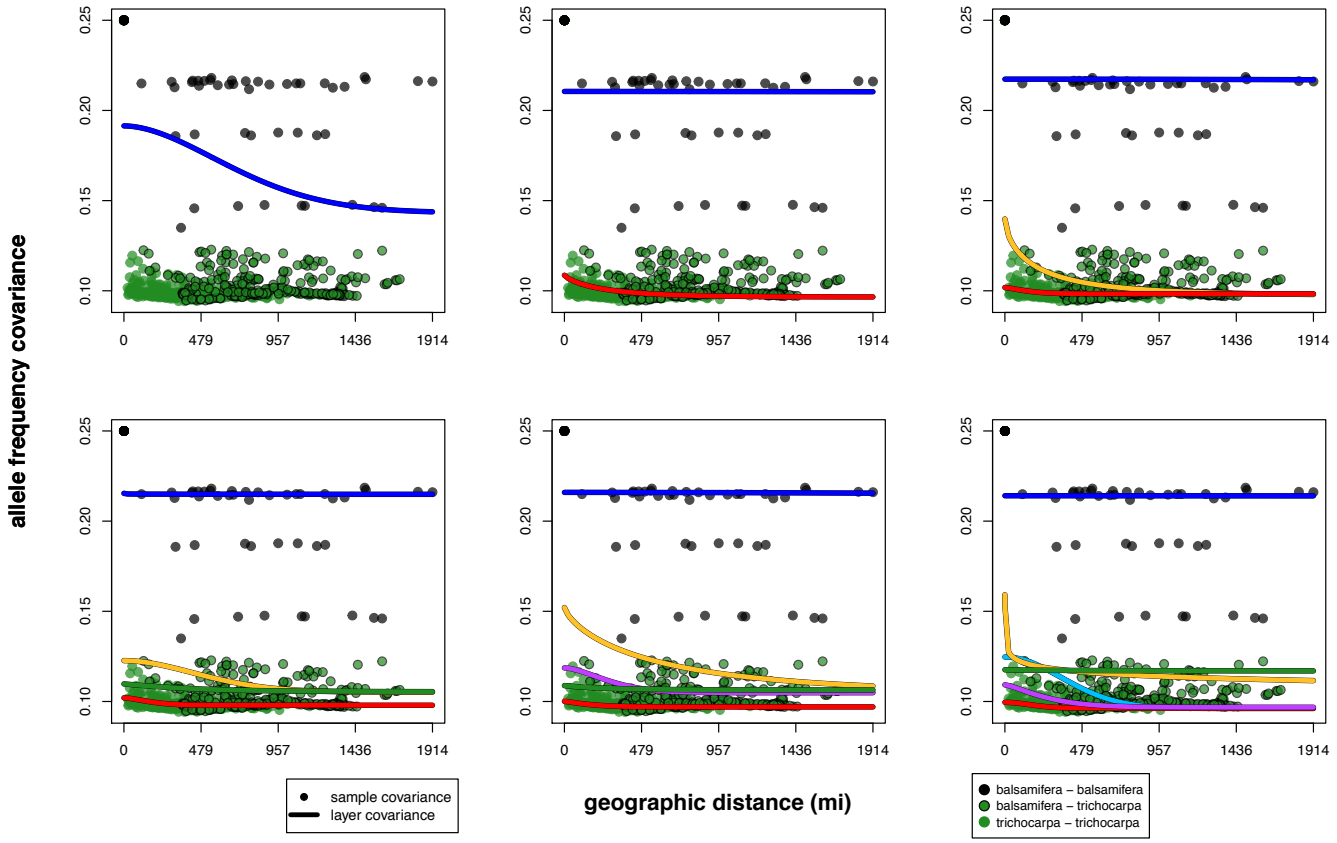

Figure S21 Plots showing the layer-specific parametric covariance curves estimated for the *Populus* data using the spatial `conStruct` model run with $K = 1$ through 6. Line colors are consistent with layer colors in Fig S19. Points are colored by the species they are a covariance between: black on black points are sample covariances between populations of *Populus balsamifera*; green on black points are sample covariances between *balsamifera* and *trichocarpa*; green on green points are sample covariances between *trichocarpa* and *trichocarpa*.

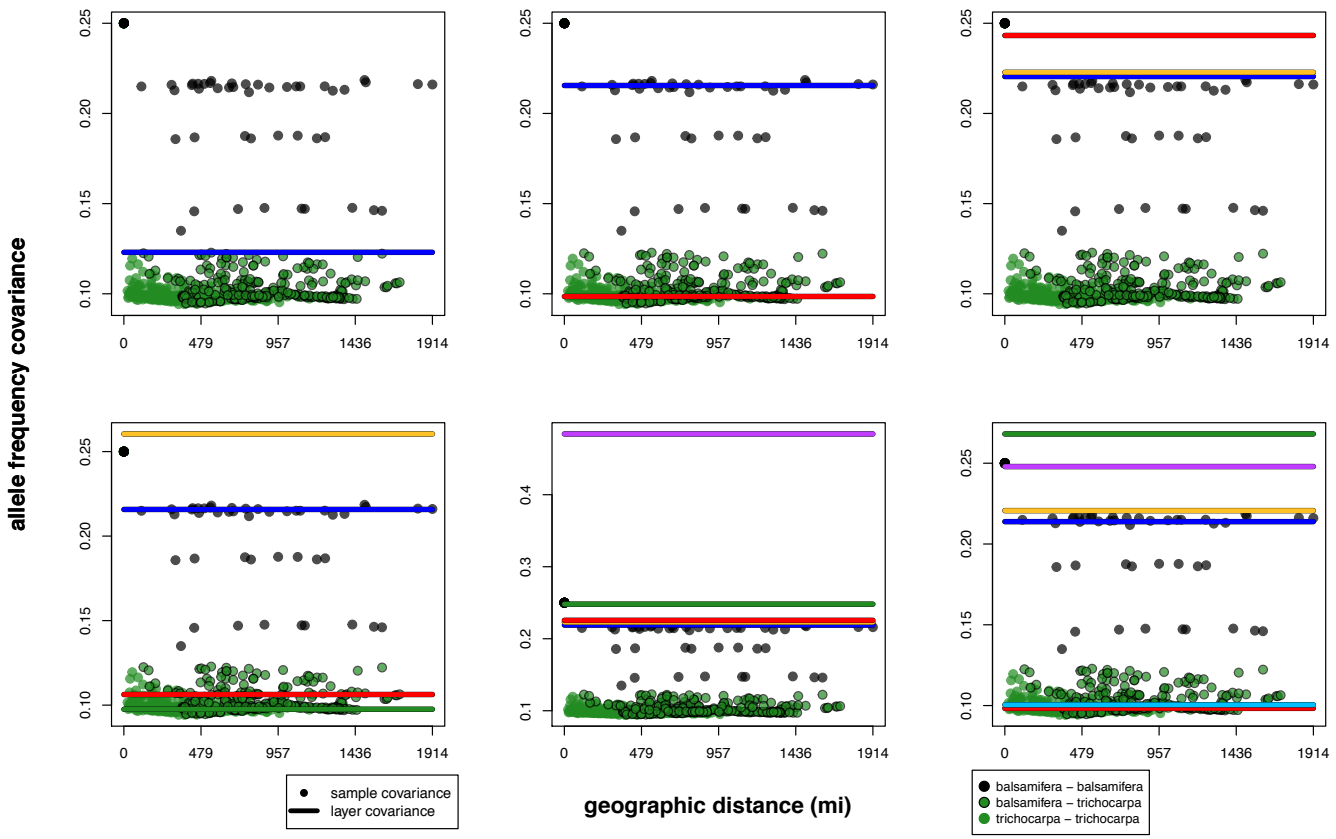

Figure S22 Plots showing the cluster-specific parametric covariances estimated for the *Populus* data using the nonspatial `conStruct` model run with $K = 1$ through 6. Line colors are consistent with cluster colors in Fig S20. Points are colored by the species they are a covariance between: black on black points are sample covariances between populations of *Populus balsamifera*; green on black points are sample covariances between *balsamifera* and *trichocarpa*; green on green points are sample covariances between *trichocarpa* and *trichocarpa*.

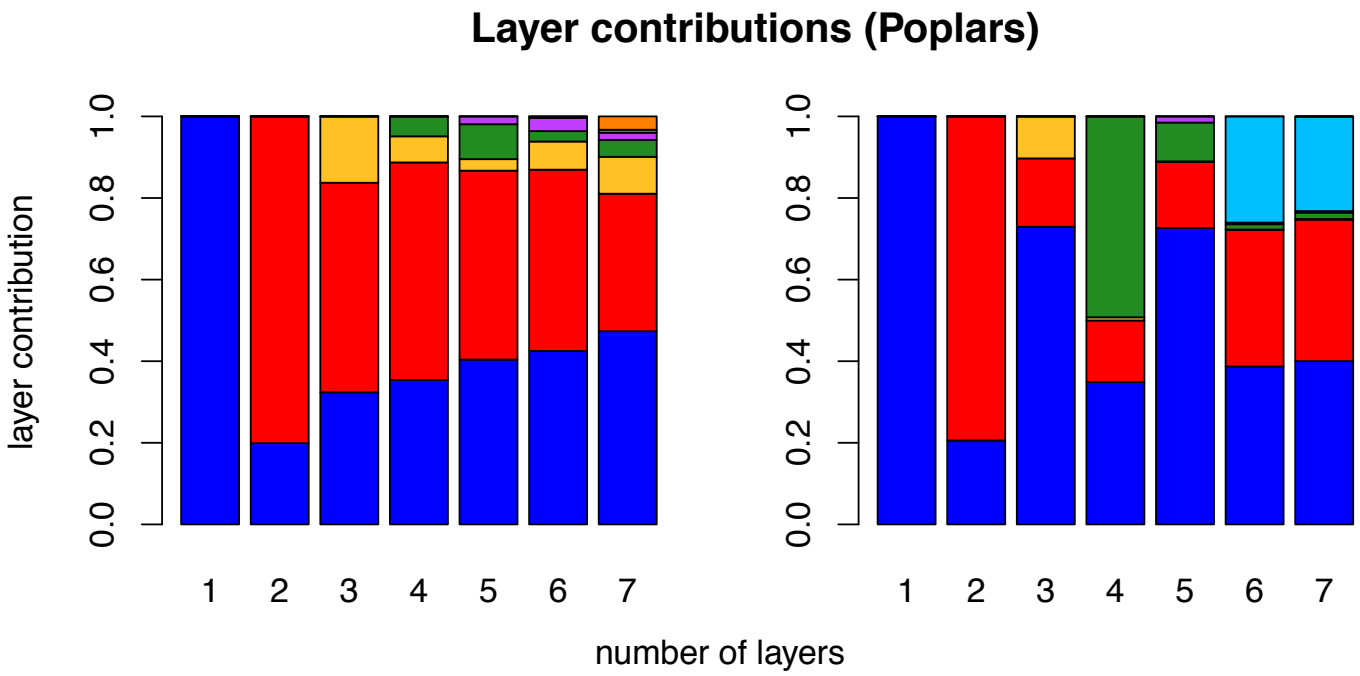

Figure S23 Layer/cluster contributions (i.e., how much total covariance is contributed by each layer/cluster), for all layers estimated in runs using $K = 1$ through 7 for the spatial model (left), and for all clusters using the nonspatial `conStruct` model (right). For each value of K along the x-axis, there are an equal number of contributions plotted. Colors are consistent with Figs S19, S21, S20, and S22.

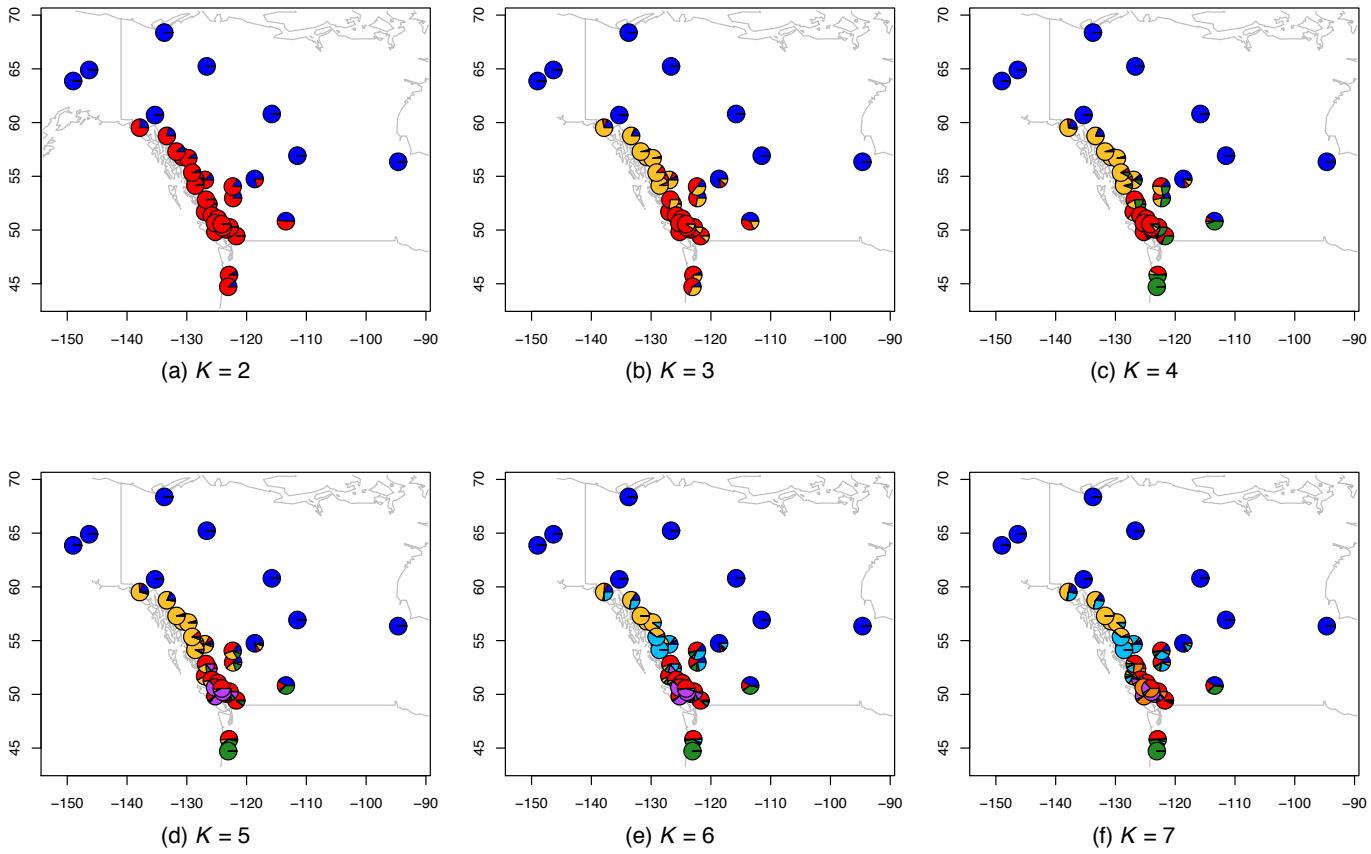

Figure S24 Maps of admixture proportions estimated for the *Populus* dataset using ADMIXTURE Alexander *et al.* (2009) for $K = 2$ through 7.

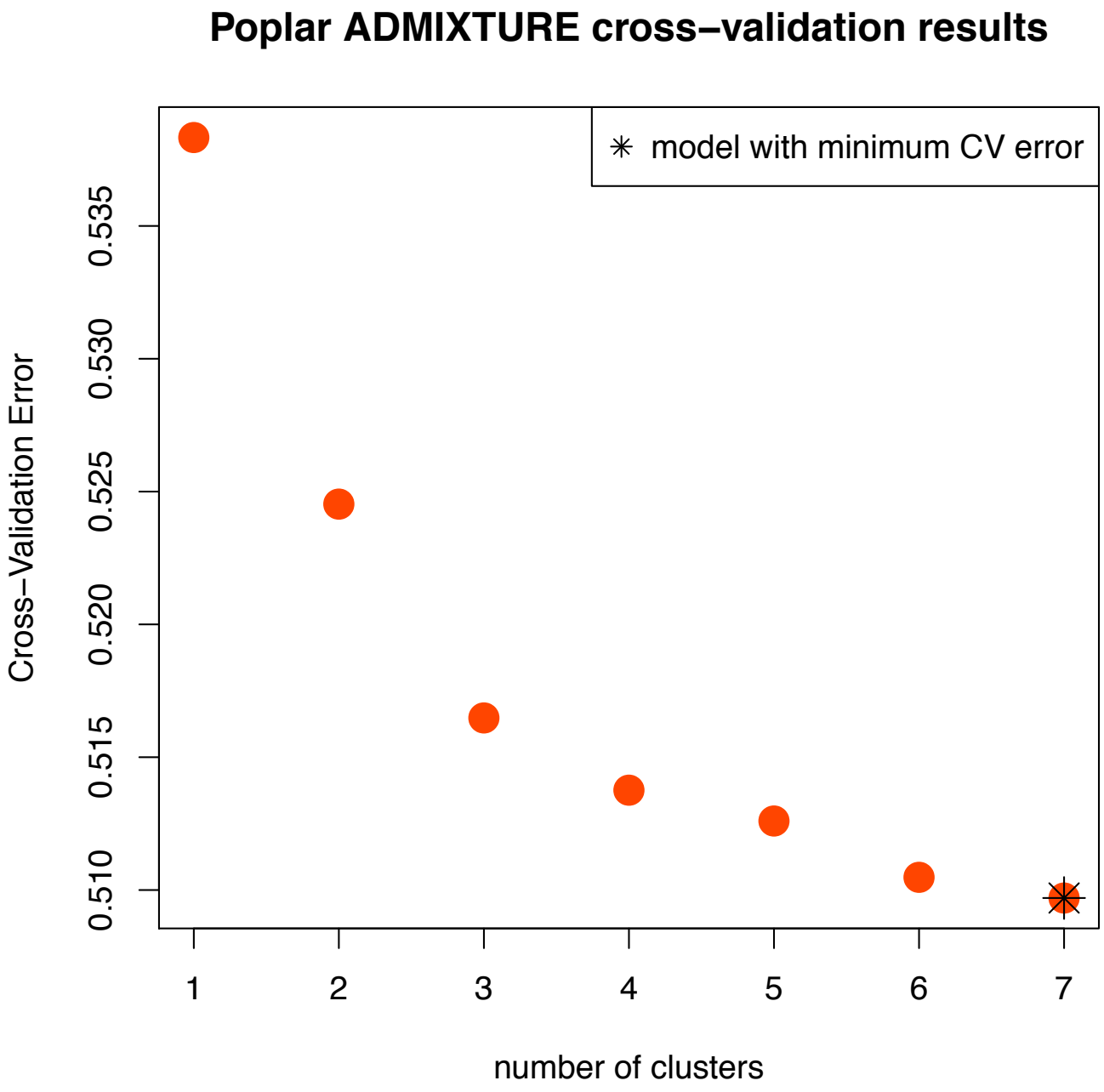

Figure S25 ADMIXTURE cross-validation results for poplar data, run with $K = 1$ through 7 using 50 data folds (`--cv=50`). The preferred model (with the lowest cross-validation error) is highlighted with an asterisk.

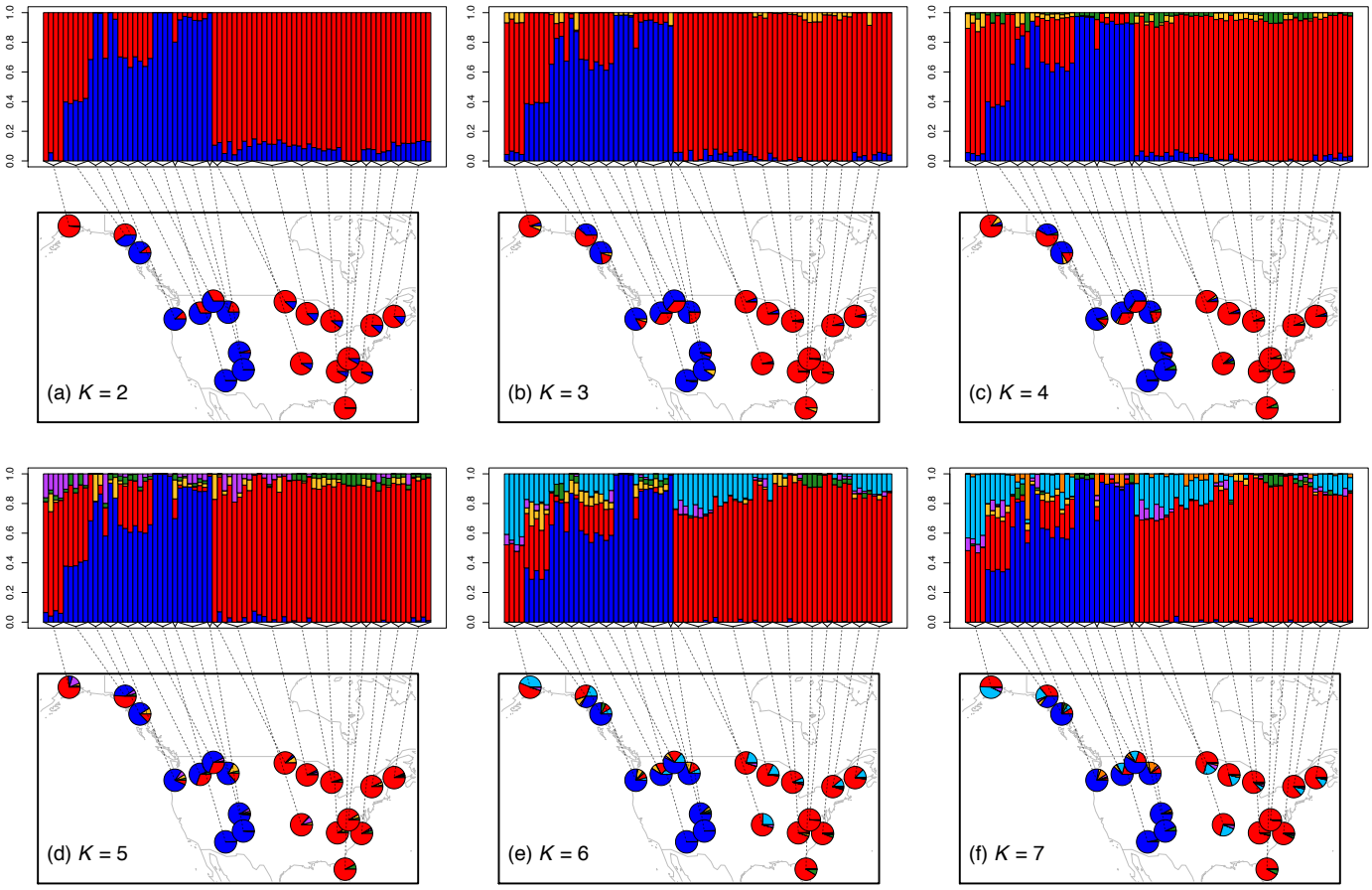

Figure S26 Map of admixture proportions estimated for the bear dataset using the spatial `conStruct` model for $K = 2$ through 7.

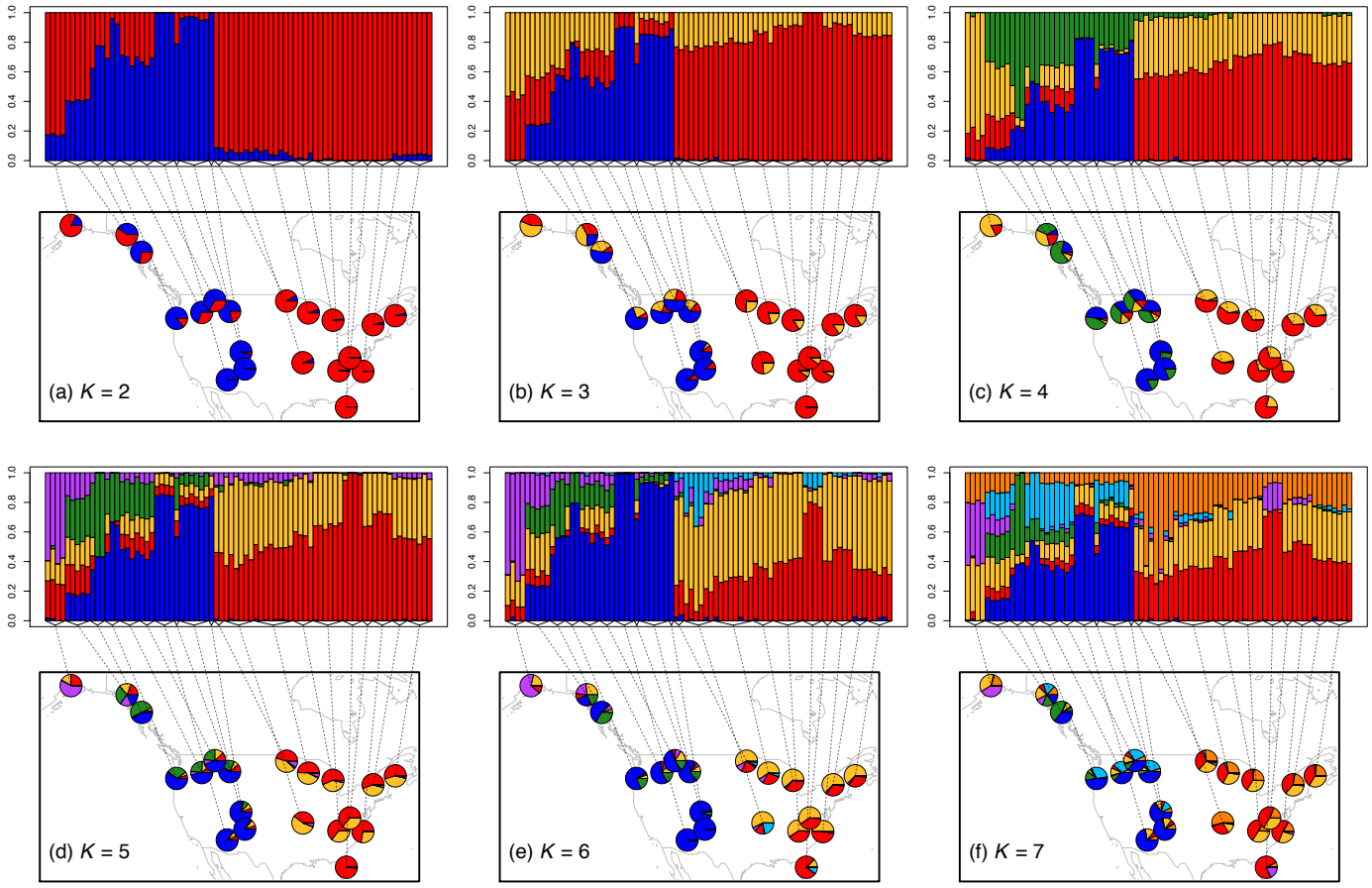

Figure S27 Map of admixture proportions estimated for the bear dataset using the nonspatial `conStruct` model for $K = 2$ through 7.

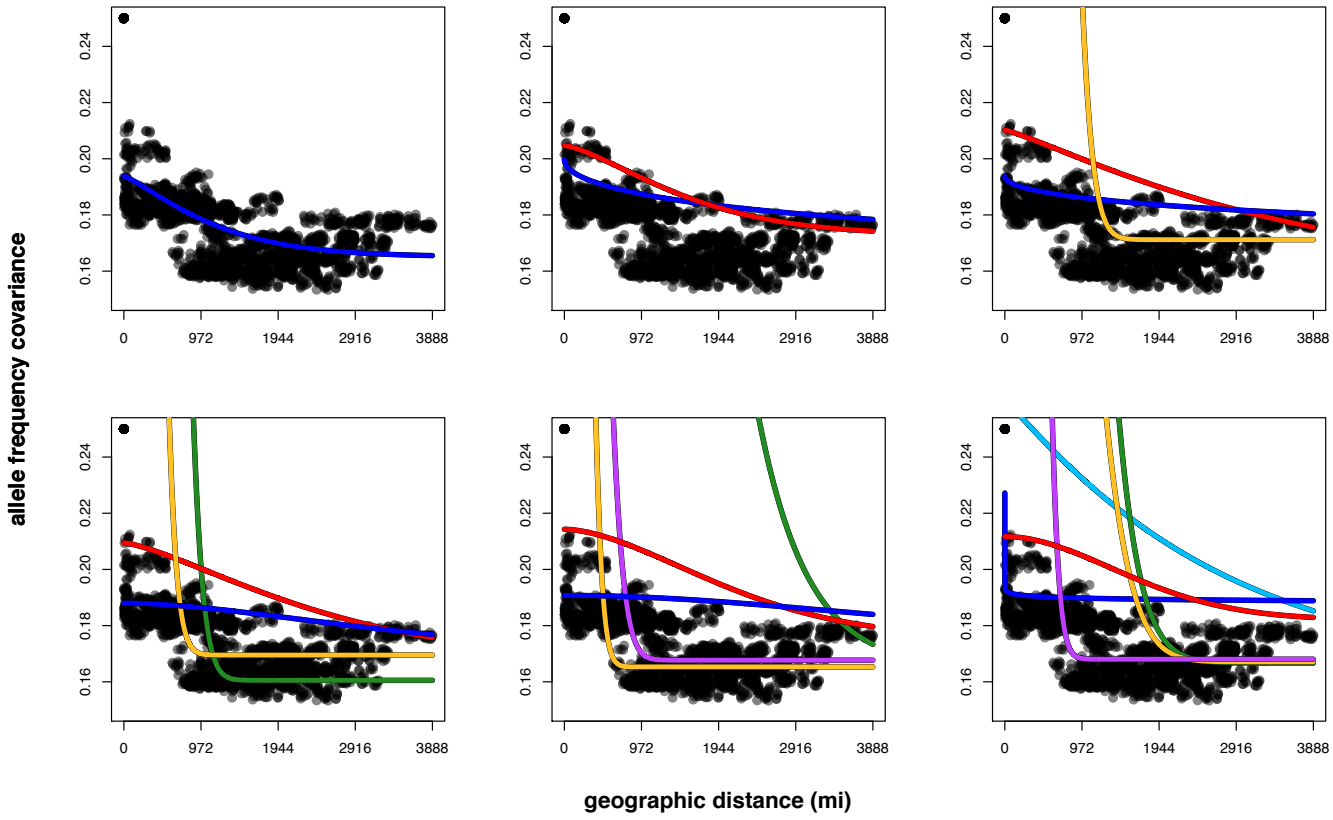

Figure S28 Plots showing the layer-specific parametric covariance curves estimated for the black bear data using the spatial `conStruct` model run with $K = 1$ through 6.

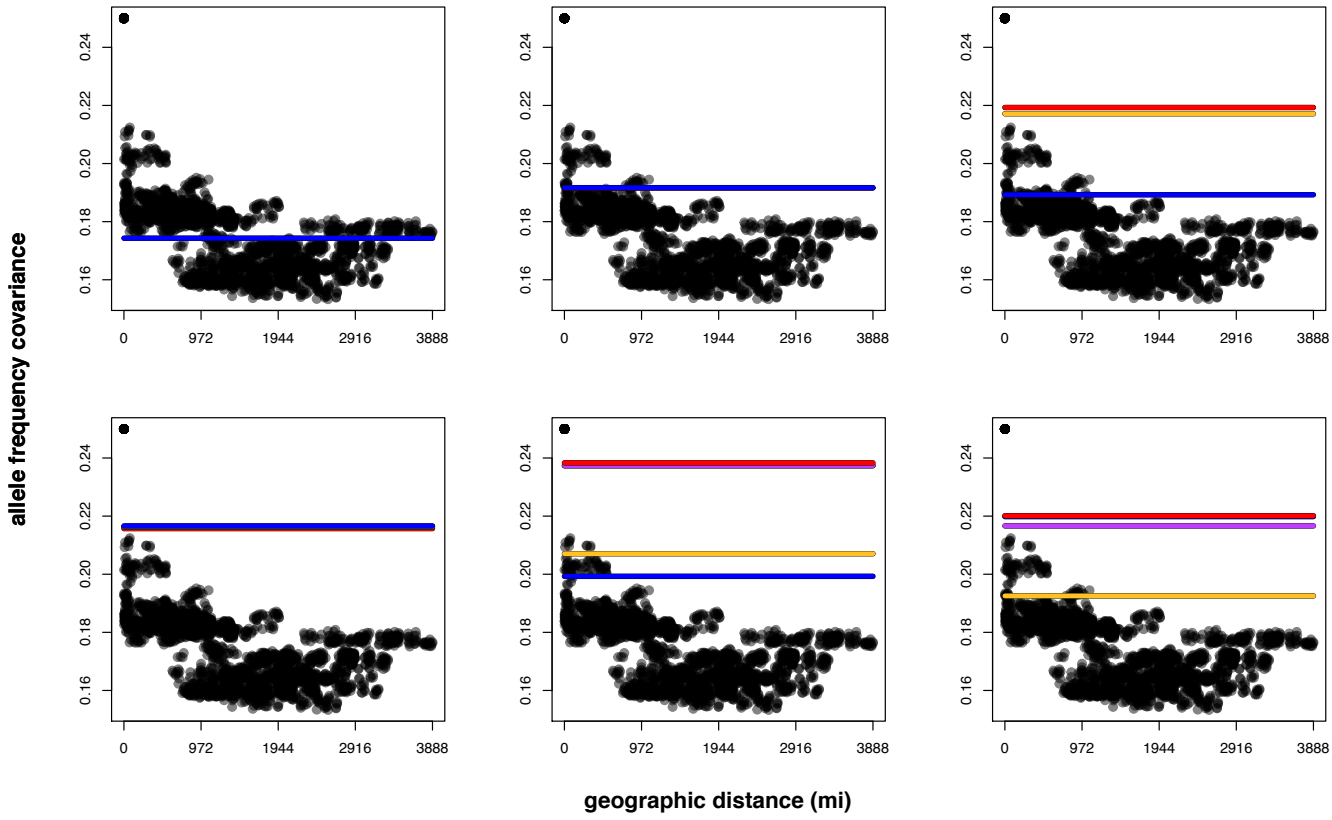

Figure S29 Plots showing the cluster-specific parametric covariances estimated for the black bear data using the nonspatial `conStruct` model run with $K = 1$ through 6.

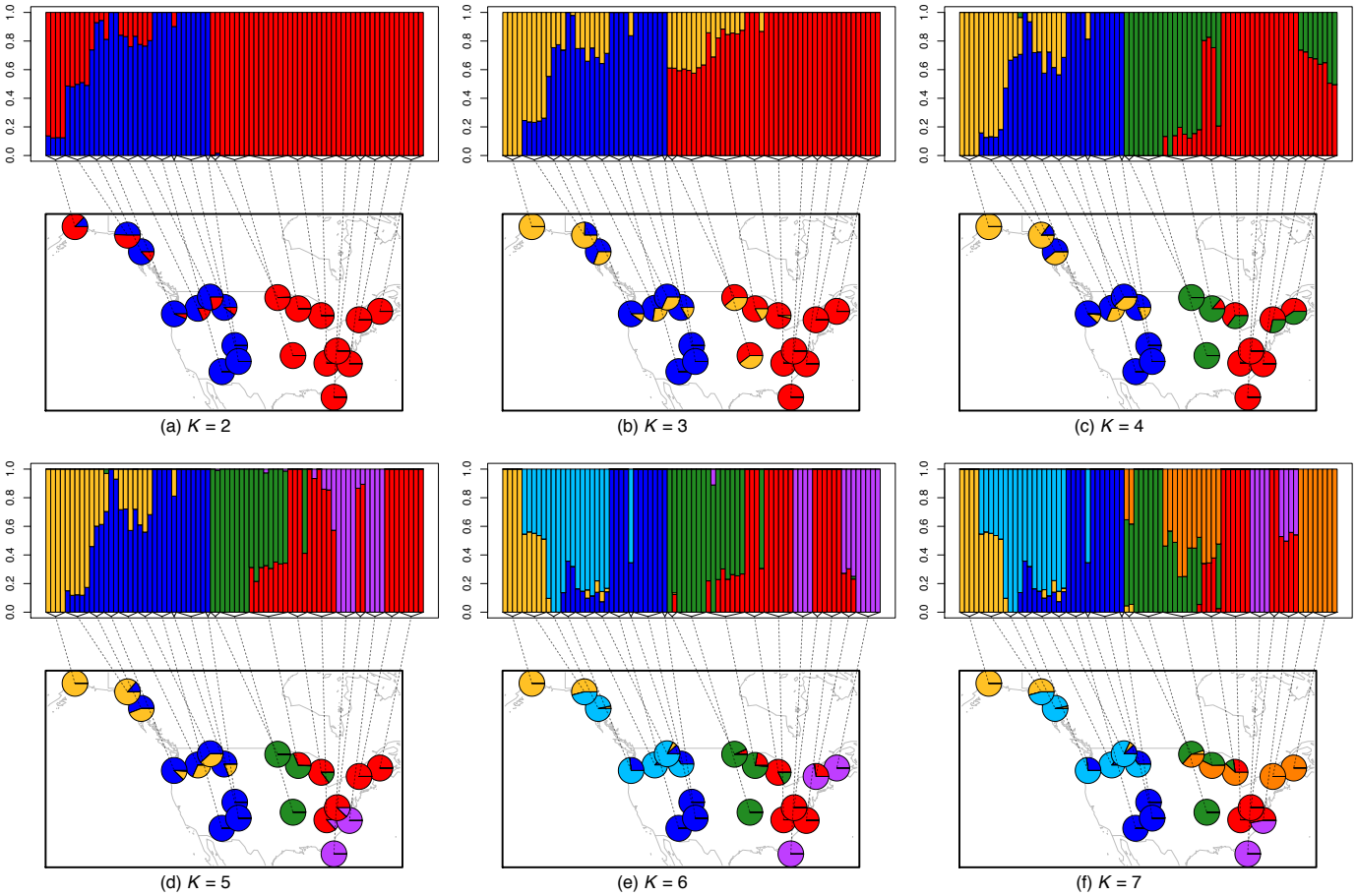

Figure S30 Maps of admixture proportions estimated for the bear dataset using ADMIXTURE Alexander *et al.* (2009) for $K = 2$ through 7.

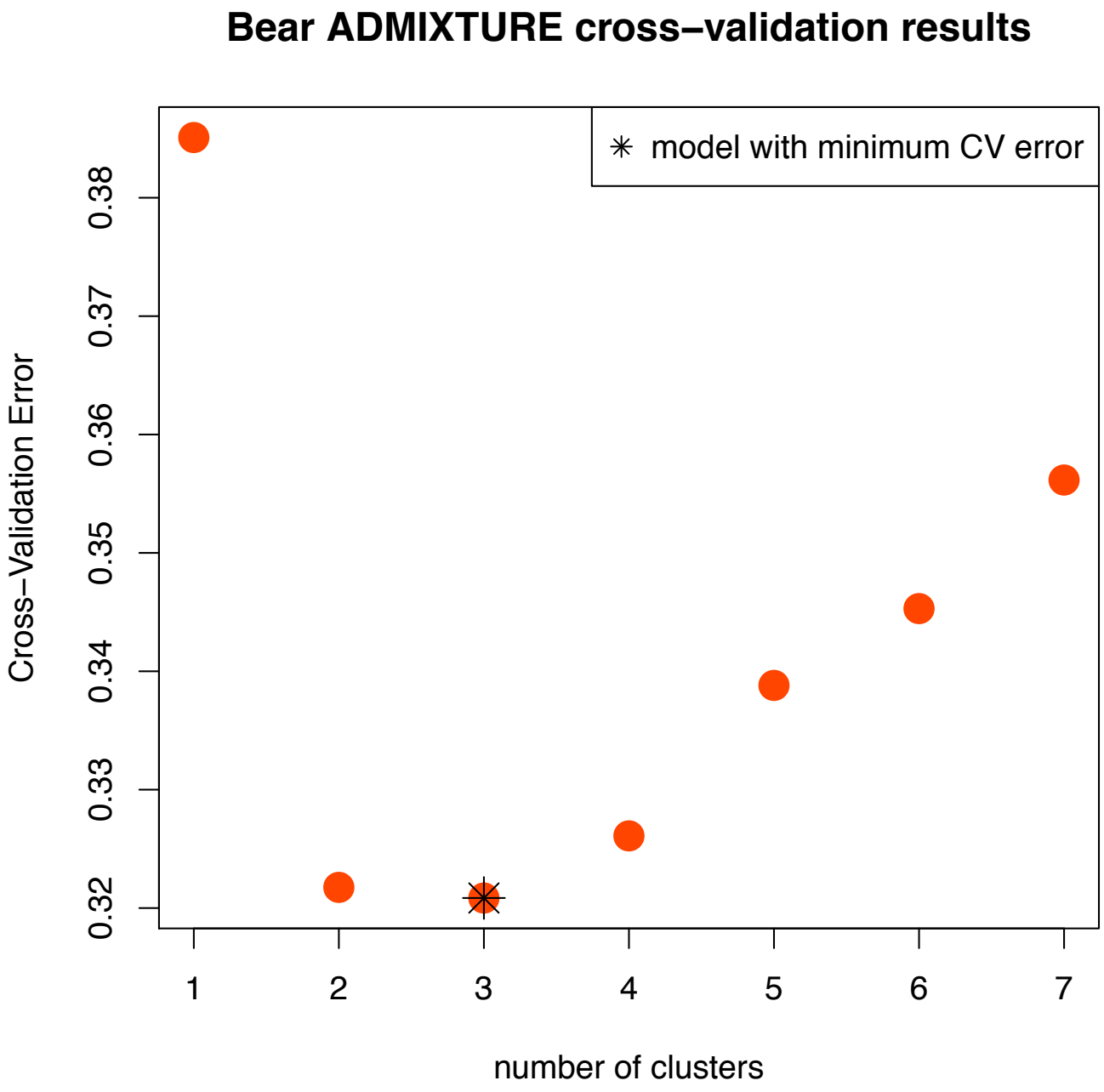

Figure S31 ADMIXTURE cross-validation results for bear data, run with $K = 1$ through 7 using 50 data folds (`--cv=50`). The preferred model (with the lowest cross-validation error) is highlighted with an asterisk.

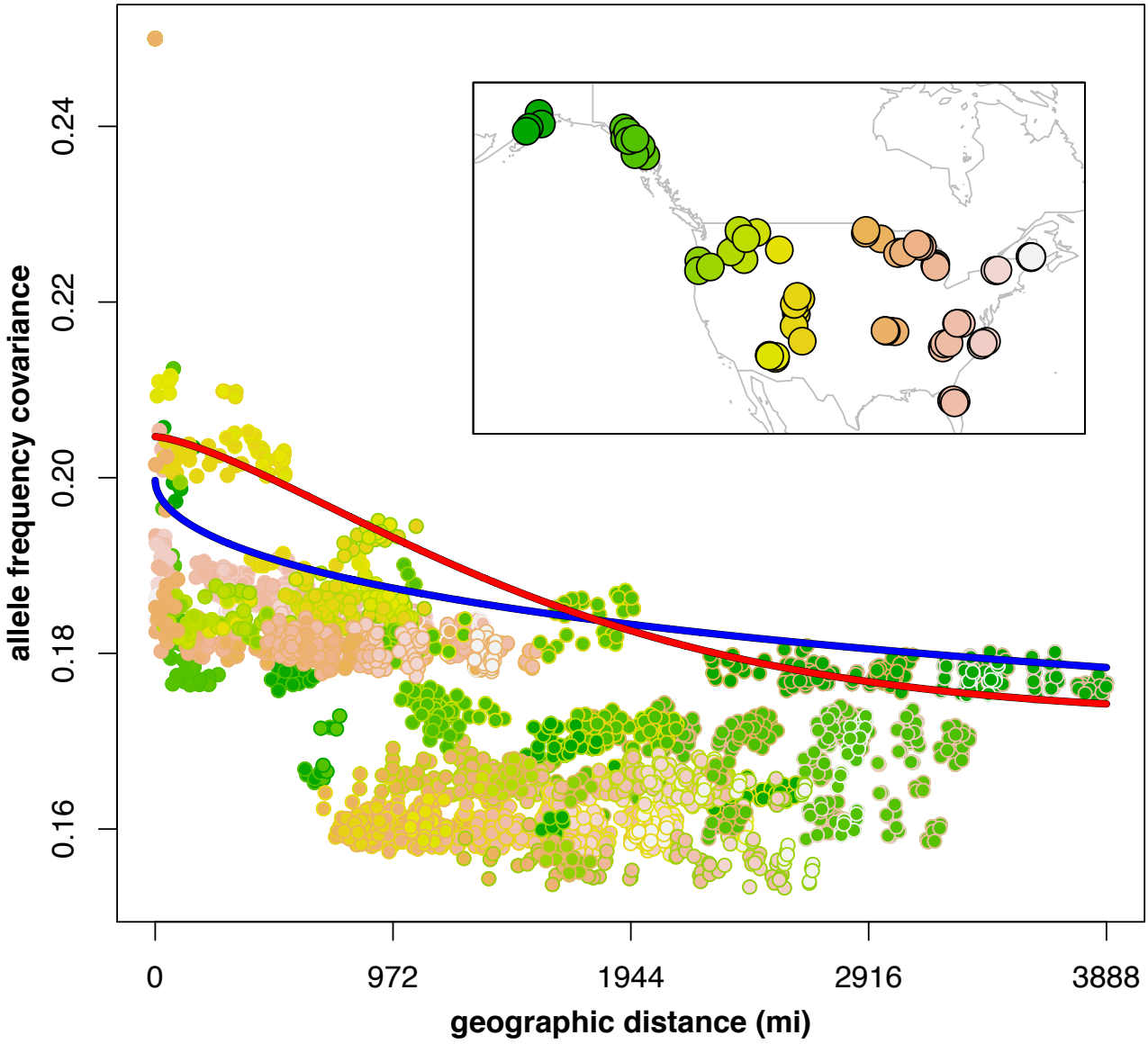

Figure S32 Allele frequency covariance between each pair of sampling locations for bear data, with each point colored to show which two locations it correponds to (see inset for colors). For instance, set of points showing covariance around 0.18 and above 2000 miles distant have dark green centers with white-to-orange borders, indicating they are comparisons between the Northwest-most (Alaskan) bears and the bears on the Eastern half of the map. The two curves show the decay of covariance within each layer for the spatial `conStruct` model with $K = 2$, as in figure S28.