Supporting information

Appendix S2. Potential sources of bias in understanding spatial variation in selection

Many of the potential sources of bias in conducting meta-analyses of selection coefficients have been previously discussed in detail in the context of temporal variation (Siepielski *et al.* 2009), and highlighted in other reviews and syntheses (Kingsolver *et al.* 2001; Hereford *et al.* 2004; Hersch & Phillips 2004; Knapczyk & Conner 2007). These possible sources of bias include: (1) a bias against publishing studies with small sample sizes or weak estimates of selection, (2) the effects of small sample sizes, (3) the effect of few population replicates, (4) sampling error, and the (5) potential for environmental correlations between traits and fitness components to generate apparent selection. Here we perform analyses to investigate several of these potential biases in the context of spatial variation; we refer the reader to Siepielski *et al.* (2009) for additional details.

First, there is a potential publication bias against publishing studies with small sample sizes or weak selection coefficients (Kingsolver *et al.* 2001; Siepielski *et al.* 2009). Plotting the selection coefficients against sample size revealed that studies with small samples sizes and weak estimates of selection are well represented, and that sample size does not appear to have an appreciable impact on estimates of selection coefficients (Figure S1).

Second, given that there was considerable variation in the sample sizes of studies included in the database, this variation could affect our estimates of variation (the SD) and the inferred changes in sign of selection (the proportion of positive coefficients) among populations. Looking at the relationship between SD of the absolute values of selection coefficients and sample sizes reveals a weak positive trend for the SD to increase with greater sample sizes (Figure S2). However, this pattern is largely driven by a few estimates. If anything, this pattern demonstrates variation in sample size is not a source of bias because it indicates that those estimates that should have the best power to detect selection (i.e., large sample sizes should have smaller SE's) would be the ones that potentially bias our overall estimates of variation in selection.

Third, spatial replication was heavily biased towards a relatively low number of replicates and this variation in the extent of replication could influence our results. When looking at the number of spatial replicates within a dataset, as the number of spatial replicates increases, the SD of the absolute values of the estimates decreases (Figure S3). This makes sense because the SD is bounded at zero, and so any source of error, either in the estimates themselves, or more seriously, due to low levels of spatial replication, will upwardly bias the estimated SD of selection coefficients. We do note thought that limited replication and it's effects on the SD affect only consideration of the distributions of estimated selection coefficients (as in Figure 2 of the main text), but the meta-analytic approaches integrate over uncertainty generated both by finite sampling within populations and finite sampling among populations (as shown in Figure 3 of the main text). Regardless, given that the bulk of the studies in the database include few estimates, caution in interpreting the magnitude of variation in the SD among studies is warranted. It could also be the case that studies with less replication are those where the researcher specifically chose populations in contrasting environments to explicitly examine how selection varies over some environmental factor. Our analysis of these subsets (main text) shows that this is likely the case to some degree. The large scale studies, with multiple populations, are likely to be less biased in this way.

Fourth, as noted in our presentation of the formal meta-analytical model, we recognize that sampling error could account for much of the variation that we observe. While our meta-analysis explicitly controls for sampling error, we also present here a second analysis to estimate the 'real' variation in selection using a variance components analysis (Cooper & Hedges 1994; see Siepielski et al. [2009] for details). In brief, we estimated the variance in selection within a study and then remove the fraction of variation that can be attributed to sampling error (as quantified with the SE). We find that the mean percent of variation that is real (i.e., after explicitly accounting for sampling error) varies considerably, but is generally low (β mean =

22.96%, range: 0 – 97.54%; *s* mean = 34.54%, range: 0-98.9). These results are consistent with the results of the meta-analysis and likewise indicate that despite sampling error, there is still true variation in selection among populations. Nevertheless, and as before, we caution the reader that this study is subject to the above sources of bias and our database contains both real variation as well as sampling error.

Fifth, spatial variation in selection can also seem to be generated by spatial variation in environmental correlations between traits and fitness components. This generates what can be thought of as "apparent selection." If traits and fitness are both correlated with a shared environmental factor, this can generate a covariance between the trait and fitness that only exists because of the shared correlation with the underlying environmental factor (e.g., Price et al. 1988; Schluter et al. 1991; Rausher 1992; Stinchcombe et al. 2002; Kruuk et al. 2003). As a result, measures of spatial variation in selection can be biased by the effects of environment (Price et al. 1988, Schluter et al. 1991). This possible effect is especially relevant here because this can in turn cause apparent spatial variation in selection when the effects of the environment on fitness or phenotype varies among populations. Recent studies advocate estimating selection gradients directly from the estimated genetic variances and covariances to avoid this potential bias (Hadfield et al. 2010; Morrissey et al. 2010); however, for most studies this will be logistically difficult.

Finally, we note that few studies statistically tested whether or not selection differed among populations. While this is not a bias per se it is important to note that an ANCOVA type model would be an ideal approach for explicitly addressing this and we suggest that future studies addressing spatial variation in selection explicitly quantify the interaction between selection and population source.

1.
Cooper, H., & Hedges, L. V. (1994). *The handbook of research synthesis*. The Russell Sage Foundation, New York.

2.

Hereford, J., Hansen, T.F. & Houle, D. (2004). Comparing strengths of directional selection: how strong is strong? *Evolution*, 58, 2133-2143.

3.

Hersch, E. & Phillips, P.C. (2004). Power and potential bias in the detection of selection in natural populations. *Evolution*, 58, 479-485.

4.

Kingsolver, J.G., Hoekstra, H.E., Hoekstra, J.M., Berrigan, D., Vigniere, S.N., Hill, C.E., Hoang, A., Gilbert, P. & Beerli, P. (2001). The strength of phenotypic selection in natural populations. *Am. Nat.*, 157, 245–261.

5.

Knapczyk, F.N. & Conner, J.K. (2007). Estimates of the average strength of natural selection are not inflated by sampling error or publication bias. *Am. Nat.* 170, 501-508.

6.

Kruuk, L. E. B., Merilä, J. & Sheldon, B. C. (2003). When environmental covariance short-circuits natural selection. *Trends Ecol. Evol.*, 18,207-208.

7.

Price, T.D., Kirkpatrick, M., & Arnold, S.J. (1988). Directional selection and the evolution of breeding date in birds. *Science*, 240, 798-799

8.

Rausher, M. D. (1992). The measurement of selection on quantitative traits: biases due to environmental covariances between traits and fitness. *Evolution*, 46, 616-626.

9.

Schluter, D., Price, T. & Rowe, L. (1991). Conflicting selection pressures and life history trade-offs. *Proc. R. Soc. Lond. B,* 246, 11–17.

10.

Stinchcombe, J.R., Agrawal, A.F., Hohenlohe, P., Arnold, S.J. & Blows, M.W. (2008). Estimating nonlinear selection gradients using quadratic regression coefficients: double or nothing? *Evolution*, 62, 2435-2440.

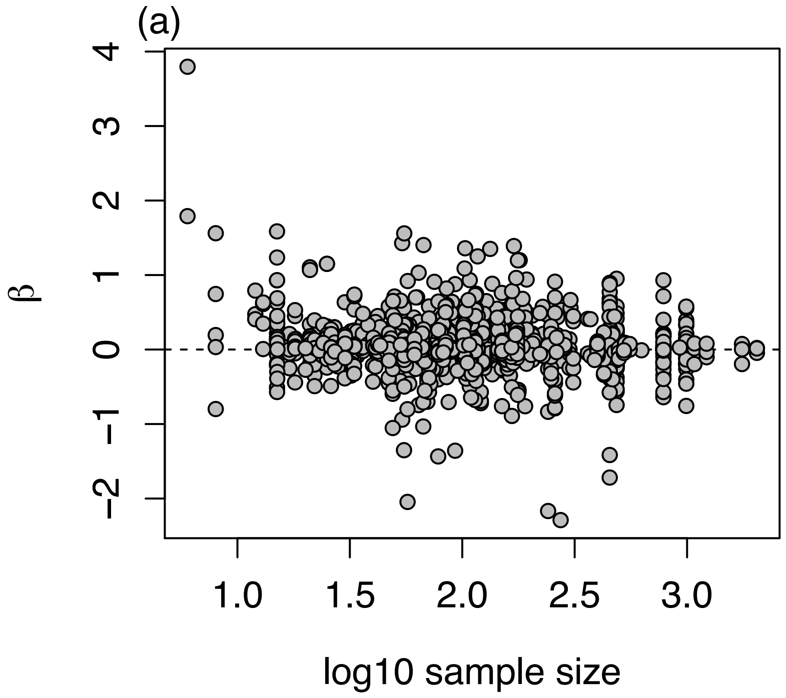
Figure Legends:

Figure S1. Selection estimates as a function of the sample size (log 10 scale). Each individual data point represents individual selection coefficients. The top panel corresponds to linear gradients (a) and the bottom panel to linear differentials (b).

Figure S2. The standard deviation (SD) of the absolute value of selection coefficients as a function of the mean sample size (log 10 scale) for a given dataset. The top panel corresponds to linear gradients (a) and the bottom panel to linear differentials (b).

Figure S3. The standard deviation (SD) of the absolute value of selection coefficients as a function of the number of populations for a given dataset. The top panel corresponds to linear gradients (a) and the bottom panel to linear differentials (b).

Figure S1


(a)
4
3
2
1
0
−1
−2
β
1.0
1.5
2.0
2.5
3.0
log10 sample size

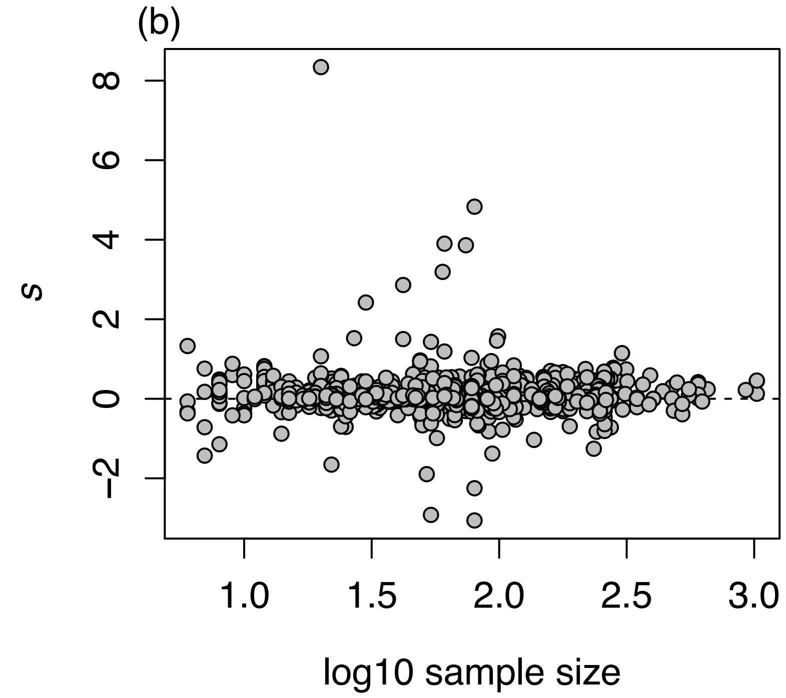

(b)
8
6
4
2
0
−2
s
1.0
1.5
2.0
2.5
3.0
log10 sample size

Figure S2

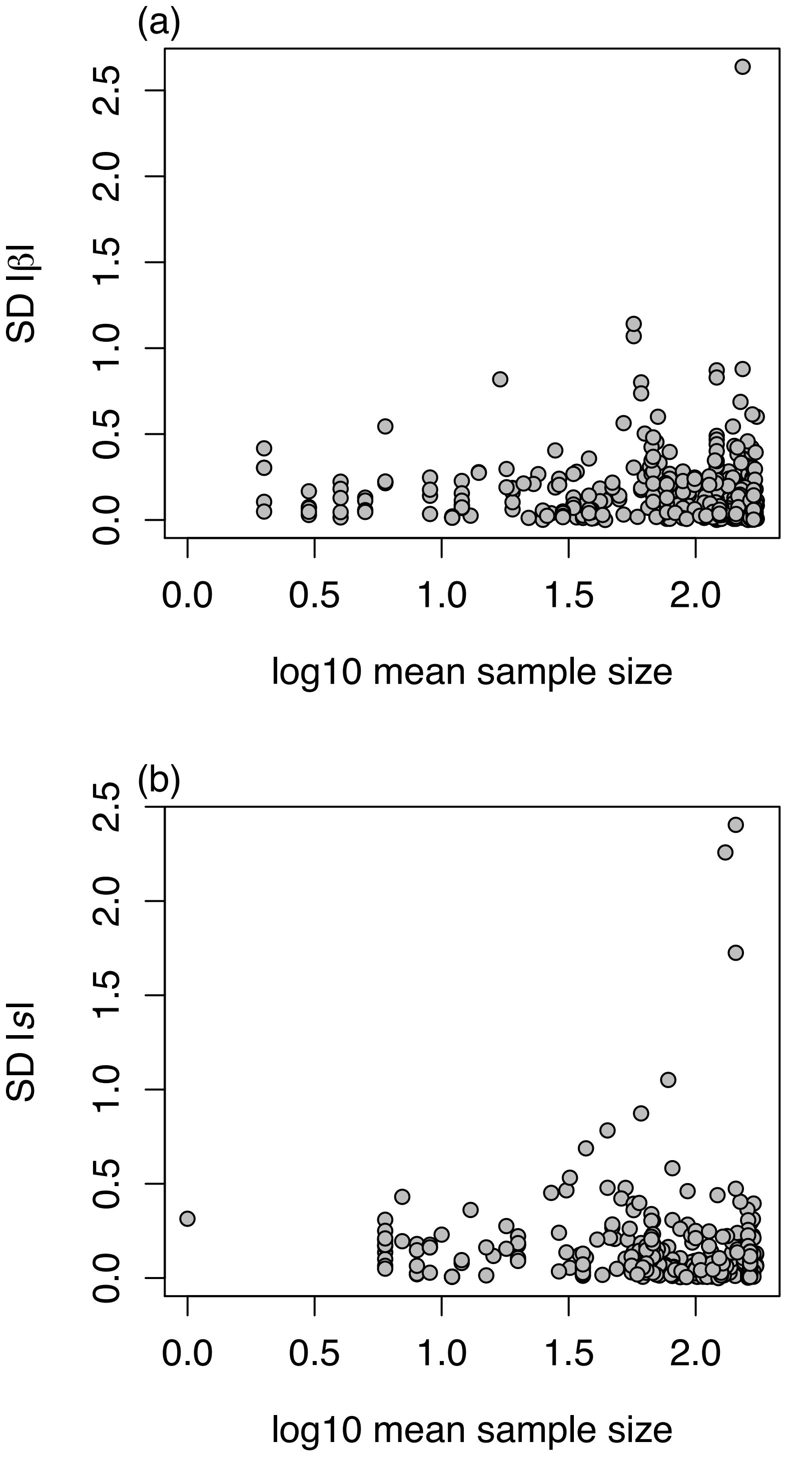
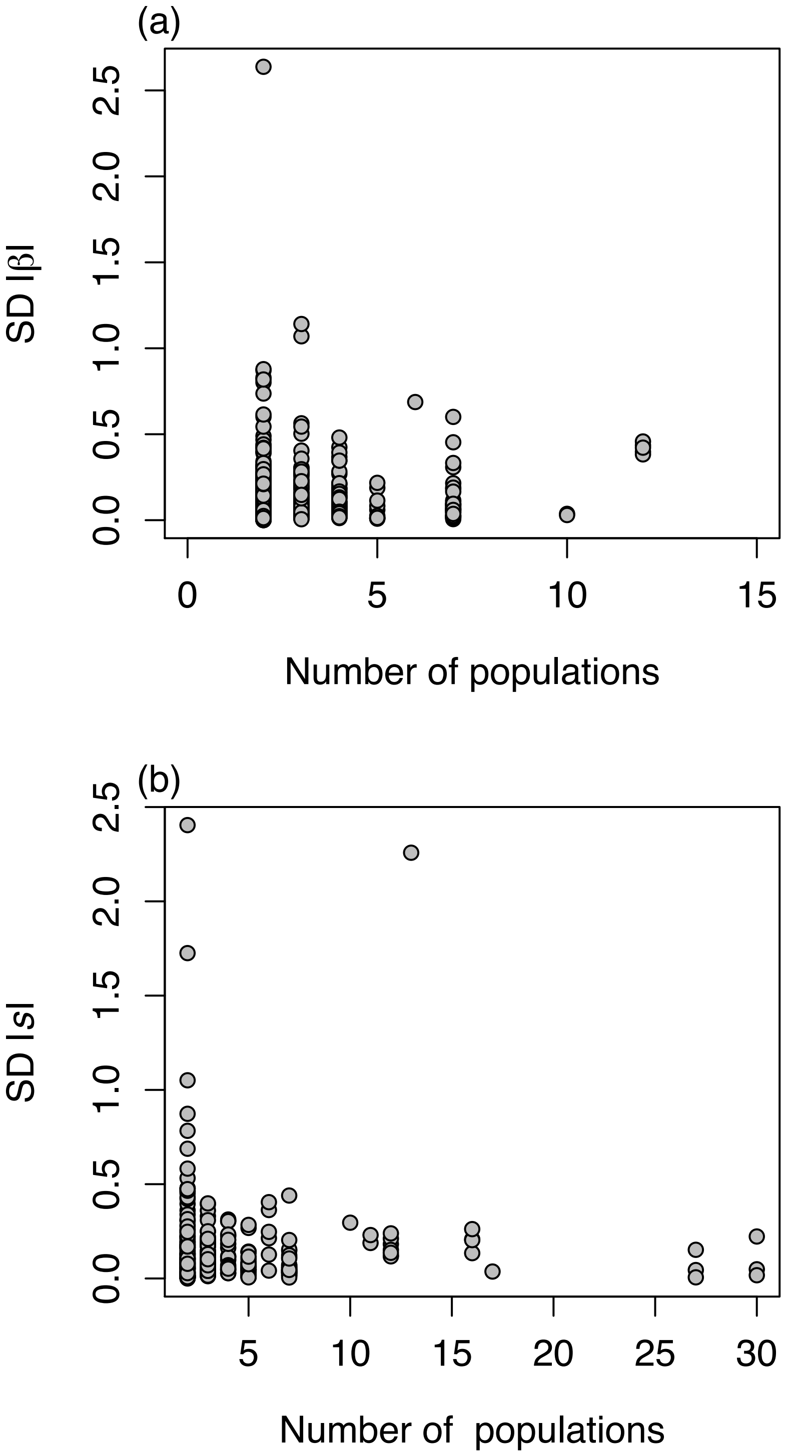

(a)
SD |β|
log10 mean sample size
0.0
0.5
1.0
1.5
2.0
2.5
(b)
SD |s|
log10 mean sample size

Figure S3



(a)
SD |β|
0.0
0.5
1.0
1.5
2.0
2.5
0
5
10
15
Number of populations
(b)
SD |s|
0.0
0.5
1.0
1.5
2.0
2.5
5
10
15
20
25
30
Number of populations