Appendix A: Literature compilation and data analysis

The most commonly reported measure of offspring size was length of embryos/newly hatched larvae or egg diameter. For a number of species (especially gastropods) the sizes of a range of different developmental stages were available. Because we were interested in variation in total maternal investment per offspring, we utilized the measure that best reflected this investment: hatchling size. We only used data for species where the eggs of two or more individuals were measured. Whilst using small numbers of females to generate estimates of among mother variance will have a tendency to overestimate variance, their use will not affect the outcome of the analyses because there were no systematic biases in the groups that contained small numbers (Quinn and Keough 2002). Data on offspring size were compiled from studies that collected females from the same population and at the same time. Our own unpublished data came from at least 5 separate broods and at least 50 offspring were measured per brood.

To examine the effects of development type on variation in offspring size at the within- and among-mother level, we used several approaches. For tests at the species level, the analysis was an unreplicated block design where 'Species' was a random, blocking factor, whereas scale of variation (within and among mothers) and development type were both fixed factors. This analysis essentially tests whether the ratio of within- to among-mother variation differed among the developmental groups. To overcome the problems associated with traditional comparative analyses (i.e. treating individual species as replicates), we also used the method of higher nodes contrasts (Harvey and Pagel 1991) and tested for an interaction between scale of variation and development type at the level of Family, Order and Class level. In the absence of comprehensive phylogenetic data, using 'Class' as the base level of

replication probably represents the most conservative approach. We included measures of offspring size that had been reported as diameters and volumes, so to test for the effects of the different measurement variables, we first ran an ANOVA to test for an effect of measurement type. Neither interaction nor main effects of Measurement type were significant (Measurement type x Development type: $F = 0.62$, $P = 0.55$, Measurement type: $F = 1.08$, $P = 0.31$) and because they were of no interest they were omitted from the final model (Quinn and Keough 2002). It should be noted that although there were significant differences in the level of offspring size variation among the different developmental modes (see Results), this will not affect the outcome of the analysis because the analysis essentially asks what is the relative level of variation in offspring within and among mothers.

To examine whether within-mother variation alone was significantly different among species with different development modes we first used a Simple Main Effects test (Quinn and Keough 2002). We found a significant difference among developmental modes (see Results) and then used a 1 sided Dunnett's test (see Quinn and Keough 2002 for details) to examine whether species with feeding, indirect development had the highest levels of within-mother variation in offspring.