**Supplementary Text S1: Methods used to estimate heritability**

Methods to estimate the heritability coefficient have themselves evolved in recent years. Many different methods exist, each with different drawbacks (see Visscher et al. 2008). Realized (narrow-sense) heritability is calculated using the Breeder’s equation as *h2R = R/S* (Falconer and Mackay 1996; Visscher et al. 2008), but becomes unreliable when selection acts on multiple co-evolving traits and when viability selection is operating (Falconer and Mackay 1996; Hadfield 2008).

The (mid)parent-offspring regression and ANOVA models for partitioning intra-familial variance are also common, but are strongly sensitive to rearing environments (Falconer and Mackay 1996). Genetic marker-based methods such as the Ritland multiple regression method (Ritland 1996) have also been used, but rely on large numbers of markers for accuracy (Visscher et al. 2008). More recently, a quantitative genetic mixed-effects model called the ‘animal model’ has become the standard for estimating heritability in wild populations due to their flexibility and power to estimate heritability (Kruuk 2004; Wilson et al. 2010). The animal model uses relatedness information from a known pedigree (or other methods, such as inferred pedigrees from genetic markers or genet ID) as a random effect in order to estimate the additive genetic variance, *VA*, associated with the breeding values of individuals (Kruuk 2004; Wilson et al. 2010). While this method is more flexible, heritabilities estimated while conditioning on unneeded fixed effects may result in estimates not being especially comparable among studies (Wilson 2008); thus, careful model construction is a crucial step in estimating heritability (see Wilson et al. 2010 for a step-by-step guide).

The majority of these studies in our meta-analysis (14/19 studies accounting for 53/95 estimates) used the ‘animal model’ to estimate heritability, while the remaining estimated heritability using an ANOVA-method of variance partitioning (4/19 studies and 33 estimates), while one study used the Ritland genetic marker method (accounting for six estimates). Visual inspection of residuals from our model fits suggested no additional unexplained variation that was related to heritability measurement method (Fig S11 in Supplementary Code C).

**Supplementary Text S2: Pre-processing of raw heritability estimates**

Heritability is calculated as a proportion of total phenotypic variation, and thus is constrained to fall between zero and one (Falconer and Mackay 1996). Because most classical meta-analytical statistical models assume normally-distributed uncertainty, transformation of our estimates prior to meta-analysis was necessary (Viechtbauer 2010; Lin and Xu 2020). The variance of heritability is often reported either as standard error of the mean (herein ‘SE’) and associated 95% confidence intervals or 95% Bayesian credible intervals (herein ‘CI’). Some Bayesian credible intervals are relatively asymmetric when the heritability estimate is close to zero or one, and thus are not easily converted to SEs without some information being lost. Additionally, transformations of proportional SE generally only work well for non-extreme point estimates (e.g., 0.2–0.8 for logit and arcsine-square root transformations) (Warton and Hui 2011; Wang 2018), while others have difficulty in back-calculating and interpretation (e.g. double arcsine transformation) (Schwarzer et al. 2019). Symmetric standard errors can have associated 95% confidence limits with non-sensical meanings, such as including negative values of heritability or values above one, meaning additive genetics contributing more than 100% of total phenotypic variance. On the other hand, posterior distributions and associated CIs can often be asymmetric near the boundary, violating the assumption of Gaussian-distributed SEs required in standard meta-analysis (Jackson and White 2018). To avoid and correct for these problems, we modelled *h2* and its associated CI limits on the logarithmic scale using the transformation:

|  |  |  |
| --- | --- | --- |
|  |  | (1) |

Since heritability values tended more towards the lower bound of 0 rather than the upper bound of 1, logarithmic transformation provides better estimates upon back-transformation for these low values while preserving the relative difference in standard errors compared to other transformations (see Supplementary Code Documentation A: Pre-processing). We selected a value of +0.2 to add, as this value allowed the inclusion of nearly all estimates save for three outliers with extremely large CIs (see Supplementary Code Documentation A: Pre-processing). Additionally, the logarithmic transformation somewhat normalizes the asymmetric Bayesian posterior distributions that tended to characterize heritability estimates near the lower boundary. We tested seven other transformations of proportions, such as the logit transformation on SE as well as CI and the double arc-sine square root transformation of SE, all resulted in similar model selection outcomes, suggesting that our results are robust to our choice of transformation. For estimates reporting SE (*n* = 32), we calculated the equivalent 95% CI limits on the original scale as:

|  |  |  |
| --- | --- | --- |
|  |  | (2) |

where *z\** = 1.96 in the case of large sample sizes (*N* > 30). For estimates reporting CIs (*n* = 56), we make the assumption for simplicity’s sake that frequentist 95% confidence intervals are comparable to Bayesian credible intervals (Gray et al. 2015). Next, we convert all intervals to the logarithmic scale by calculating the point estimate and the confidence/credibility interval limits, then directly convert them using the same 95% CIlwr/upr,T = *ln*(95% CIlwr/upr + 0.2) transformation as above. Finally, we obtain the transformed standard errors (*SET*) based on a rearrangement of the previous formula:

|  |  |  |
| --- | --- | --- |
|  |  | (3) |

For example, a heritability estimate and 95% CI of *h2* = 0.25 [0.1,0.4] would be transformed using the *ln*[*h2*+0.2] transform to *h2T* = –0.80 [–1.2,–0.51], then SET computed as: [–0.51 – (–1.2)]/(2\*1.96) = 0.176.

**Supplementary Text S3: Model selection results of trait type** × **heritability type and trait type** × **growth form**

Model selection to examine possible trait type × heritability type interactions used a subset of data that also allowed the inclusion of a trait type × life stage interaction. We did not fit a three-way interaction of the above factors, given that there were no studies for some combinations of levels of the three factors. Model selection supported the model of trait type × life stage (Table S5). This model again had significant residual heterogeneity (*QE51* = 98, *p* < 0.0001; *I2* = 59%; *R2* = 71%), and had coefficient values similar to the previous model of trait type × life stage + heritability type (Fig. S4; Table S6). There was one highly influential point for adult bleaching (Cook’s distance = 3.7). However, when this point was removed, both juvenile bleaching and growth remained significantly low. The fail-safe number was again large, indicating robustness to any publication bias (127 > 100). Additionally, the outcome of model selection was unchanged when data were re-analyzed without this estimate.

Finally, we tested for possible trait type × growth form interactions. Again, no three-way interactions were possible given the combinations of levels of factors with adequate representation in the data. There were limited estimates for some growth forms, such as for columnar (n=3 estimates) and encrusting (n=1) corals, thus these levels (and thus the corresponding estimates) were excluded in order to examine a complete trait type × growth form interaction. Thus, we examined interactions across branching (n=8), corymbose (n=32), and massive (n=14) coral growth forms for five traits: bleaching, growth, nutrient content, survival, and symbiont community, and this also allowed interactions of trait × life stage and trait × heritability type. Similar to results of our previously reported analysis, a model of trait type × life stage was the most strongly supported by the data (Table S7-S8; Fig. S5).

**References**

Falconer DS, Mackay TF. 1996. Introduction to Quantitative Genetics. Essex, UK: Longman.

Gray K, Hampton B, Silveti-Falls T, McConnell A, Bausell C. 2015. Comparison of bayesian credible intervals to frequentist confidence intervals. Journal of Modern Applied Statistical Methods. 14(1):43–52. doi:10.22237/jmasm/1430453220.

Jackson D, White IR. 2018. When should meta-analysis avoid making hidden normality assumptions ? (June):1040–1058. doi:10.1002/bimj.201800071.

Lin L, Xu C. 2020. Arcsine-based transformations for meta-analysis of proportions: Pros, cons, and alternatives. Health Science Reports. 3(3):1–6. doi:10.1002/hsr2.178.

Schwarzer G, Chemaitelly H, Abu-Raddad LJ, Rücker G. 2019. Seriously misleading results using inverse of Freeman-Tukey double arcsine transformation in meta-analysis of single proportions. Research Synthesis Methods. 10(3):476–483. doi:10.1002/jrsm.1348.

Viechtbauer W. 2010. Conducting meta-analyses in R with the metafor package. Journal of Statistical Software. 36(3):1–48.

Wang N. 2018. Conducting Meta-Analyses of Proportions in R. John Jay College of Criminal Justice.:1–62.

Warton DI, Hui FKC. 2011. The arcsine is asinine: the analysis of proportions in ecology. Ecology. 92(1):3–10.