G3: Genes|Genomes|Genetics Early Online, published on October 4, 2015 as doi:10.1534/g3.115.020701

Estimating trait heritability in highly fecund species

Sarah W. Davies[1,*,+], Samuel V. Scarpino[2,*,++], Thanapat Pongwarin[1], James Scott[3,4], and Mikhail V. Matz[1,+++]

[1]Department of Integrative Biology, The University of Texas at Austin, Austin, Texas, USA
[2]Santa Fe Institute, Santa Fe, New Mexico, USA
[3]Department of Statistics and Data Sciences, The University of Texas at Austin, Austin, Texas, USA
[4]Department of Information, Risk, and Operations Management, The University of Texas at Austin, Austin, Texas, USA
[*]these authors contributed equally to this work
[+]address corespondence regarding experiments to daviessw@gmail.com
[++]address corespondence regarding statistics to scarpino@santafe.edu
[+++]address general corespondence to scarpino@santafe.edu

Running Title: Estimating heritability

Keywords: Heritability; Non-model organisms; Common Garden; Binary Variable Traits; Coral Settlement

Corresponding Author Contact:

Mikhail V. Matz
Department of Integrative Biology
The University of Texas at Austin
1 University Station #C0930
Austin, Texas, 78712, USA
matz@utexas.edu
(512) 475-6424

Abstract

Increasingly, researchers are interested in estimating the heritability of traits for non-model organisms. However, estimating the heritability of these traits presents both experimental and statistical challenges, which typically arise from logistical difficulties associated with rearing large numbers of families independently in the field, a lack of known pedigree, the need to account for group or batch effects, etc. Here we develop both an empirical and computational methodology for estimating the narrow-sense heritability of traits for highly fecund species. Our experimental approach controls for undesirable culturing effects, while minimizing culture numbers, increasing feasibility in the field. Our statistical approach accounts for known issues with model-selection by using a permutation test to calculate significance values and includes both fitting and power calculation methods. We further demonstrate that even with moderately high sample-sizes, the p-values derived from asymptotic properties of the likelihood ratio test are overly conservative, thus reducing statistical power. We illustrate our methodology by estimating the narrow-sense heritability for larval settlement, a key life-history trait, in the reef-building coral *Orbicella faveolata*. The experimental, statistical and computational methods, along with all of the data from this study, are available in the R package multiDimBio.

Introduction

Organisms with high fecundity, small propagule size, and limited parental investment, also referred to as r-selected species, often exhibit higher levels of nucleotide diversity and/or standing genetic variation when compared to k-selected species (Romiguier *et al.*, 2014). Many marine species, including fish and invertebrates, exhibit these r-selected life history characteristics (Doherty & Fowler, 1994) and indeed have been shown to exhibit high levels of genetic diversity (Bay *et al.*, 2004; Davies *et al.*, 2015). However, this high genetic diversity does little to predict how a population will respond to environmental perturbations, such as those caused by climate change. Instead, the key question is not how much variation is present, but what is the heritability of the traits under selection following the perturbation.

Quantifying narrow-sense heritability, the proportion of phenotypic variance attributable to additive genetic effects (Lynch & Walsh, 1998), for non-model organisms presents both experimental and statistical challenges. Most experiments aiming to quantify narrow-sense heritability involve multi-generational breeding programs and large numbers of crosses with many culture replicates to account for "jar effects," both of which are rarely feasible in non-model species.

Here we present a quantitative genetic methodology for estimating the narrow-sense heritability of traits in highly fecund species. The method does not require the onerous sampling schemes usually required for these types of experiments. Instead, our approach leverages high fecundity by completing independent fertilizations with large quantities of eggs equally divided among sires to account for sperm competition (Figure 1). These cultures are then mixed into a single bulk culture (common garden) and divided into three replicate tanks per dam. Bulk larvae are then sorted according to the trait of interest, which in this study is a binary trait (whether or not the larvae underwent metamorphosis in response to settlement cue). Single larvae that "succeeded" and "failed" are then individually genotyped and their sire assignments are compared to the predicted distribution of sire assignments in the original design. This experimental design allows for all sires to be cultured in 'common garden' conditions, which greatly reduces the number of cultures as compared to standard approach where each family would be cultured individually, resulting in a culture number of 3x the number of sires. The narrow-sense heritability of these data can be estimated using a generalized linear mixed model with a binomial error distribution. However, as we discuss below, appropriately determining statistical significance is non-trivial. This method of quantifying heritability of binary traits is broadly applicable to many traits of interest including–but not limited to–stress tolerance, dispersal potential, and disease susceptibility. Furthermore, the framework we have developed–including the statistical methods–can be readily adapted to traits with different distributions, e.g. normally distributed phenotypes.

To demonstrate this methodology, we estimated the heritability of dispersal potential

in reef-building coral larvae. The majority of corals–like many other marine invertebrates–release gametes into the water annually. These gametes develop into planktonic larvae that are dispersed by ocean currents, representing each coral's only dispersal opportunity (Baird *et al.*, 2009). The now pelagic larvae can travel great distances before settling on a reef, but once the larva settles, it will remain there for the duration of its life. Therefore, selection for dispersal potential is limited to optimizing larval traits, which can be investigated through classical quantitative genetics, e.g. Meyer *et al.* (2009). Specifically, we determined how much variation in the early larval responsiveness to settlement cue depends on the genetic background of larvae. The experiments were performed on larvae of the hermaphroditic mountainous star coral, *Orbicella faveolata*, which is an important but endangered Caribbean reef-building coral. To analyze these data, and estimate the narrow-sense heritability of this binary trait, we developed a Monte Carlo method for performing model selection and calculating statistical power with generalized linear mixed models. The code and data are available in the R package multiDimBio (Scarpino *et al.*, 2014).

Materials and methods

Experimental Framework

Our experimental framework, which is summarized in Figure 1, proceeds in four steps. First, we perform crosses between the desired number of parents. Second, all offspring from a single dam are reared in the same environment ('common garden'). Third, offspring are phenotyped for the trait of interest and genotyped to determine paternity. Fourth, these data are analyzed using random-effects models and a permutation test to determine statistical significance. What follows is a detailed description of how to estimate the narrow-sense heritability of coral settlement using this framework.

Application of the experimental framework to coral settlement

Crossing design and larval rearing

One day prior to the annual coral spawn on August 7, 2012, ten independent *O. faveolata* colony fragments (10cm x 10cm) were collected from the East Flower Garden Banks National Marine Sanctuary, Gulf of Mexico. Colonies were maintained in flow through conditions aboard the vessel and were shaded from direct sunlight. Colonies were at least 10m apart to avoid sampling clones, as clones within reefs have been detected in this genus (Severance and Karl, 2006; Baums et al., 2010). However, intracolony variation (chimerism) in scleractinian corals is very rare (Puill-Stephan et al., 2009), so each sire was assumed to only produce sperm of a single genotype. Prior to spawning, at 20:00CDT on August 8, 2012, colonies were isolated in individual bins filled with 1μm filtered seawater and were shaded completely. Nine colonies spawned at approximately 23:30CDT. From these spawning colonies, we collected gamete bundles and separated eggs and sperm with nylon mesh. Each colony was used as an independent sire, with no additional sperm/sires included in this study. Samples from each sire were preserved in ethanol for genotyping.

Divers collected gamete bundles directly from three colonies during spawning and eggs were separated to serve as maternal material (N=3 dams). Eggs were divided equally among fertilization bins (N=9 per dam) and sperm from each sire was added at 0200CDT on August 9, 2012 for a total of 27 fertilization bins. Control self-cross trials verified that self-fertilization was not detectable in our samples. After fertilization, at 0800CDT, excess sperm was removed by rinsing with nylon mesh, and embryos for each dam across all sires were pooled in one common culture. Densities were determined and larvae were stocked into three replicate culture vessels at 1 larva per 2ml for a total of nine culture containers (N=3 per dam). Larvae were transferred to the University of Texas at Austin on August 10, 2012. Following spawning, colony fragments were returned to the reef. All work was completed under the Flower garden Banks National Marine Sanctuary permit #FGBNMS-2012-002.

Common Garden Settlement Assay

On August 14, 2012, 6 day old, pre-competent larvae from the three replicate bins for a single dam were divided across three settlement assays. Four hundred healthy larvae per culture replicate were added to a sterile 800ml container with five conditioned glass slides and finely ground, locally collected crustose coralline algae (CCA), a known settlement inducer for this coral genus (Davies *et al.*, 2014). Cultures were maintained for three days after which each slide was removed and recruits were individually preserved in 96% ethanol, representing larvae exhibiting "early" responsiveness to settlement cue. Culture water was changed, new slides were added with additional CCA and larvae were maintained until they reached 14 days old. All settled larvae on slides were discarded and 50 larvae per culture were individually preserved in 96% ethanol. Larvae from the other two dams were not used in these assays due to high culture mortality.

Larval DNA Extraction

Larval DNA extraction followed a standard phenol-chloroform iso-amyl alcohol extraction protocol, see Davies *et al.* (2013), with modifications to accommodate for the single larva instead of bulk adult tissue.

Parental Genotyping

Sire genotyping was completed using nine loci from Davies *et al.* (2014) and four loci from Severance *et al.* (2004) following published protocols. Amplicons were resolved on agarose gel to verify amplification and molecular weights were analyzed using the ABI 3130XL capillary sequencer. GeneMarker V2.4.0 (Soft Genetics) assessed genotypes and loci representing the highest allelic diversities amongst the sires were chosen as larval parentage markers. To ensure that each sire was a unique multilocus genotype (MLG) and that the relatedness between sires was negligible, we compared the allelic composition of each sire across six microsatellite loci (MLG) and calculated the Probability of Identity at each locus in GENALEX v6.5 Peakall & Smouse (2006).

Table 1. Summary of the six microsatellite loci from Davies *et al.* (2013) used in paternity assignment.

Locus	Observed (bp)	N_a	Fluorescence
M_fav4	375-391	5	FAM
maMS2-5	280-328	20	FAM
maMS8	197-203	3	FAM
M_fav6	387-429	11	HEX
M_fav7	453-498	9	HEX
maMS2-8	187-205	10	NED

Larval Parentage

To compensate for the low larval DNA concentrations, 3μl of each single extracted larva (unknown concentration) was amplified in a multiplex reaction with six loci from Davies *et al.* (2013) with the following modifications: 1μM of each fluorescent primer pair (N=6) and 20μL reaction volumes (Table 1). Alleles were called in GeneMarker V2.4.0 and offspring parentage was assigned based on presence/absence of sire alleles. Data were formatted into a dataframe consisting of the number of early settlers and swimming larvae that were assigned to each sire (A-J) from each of three replicate bins (1-3).

Statistical Methods

Estimating narrow-sense heritability from binary data

In principle, estimating narrow-sense heritability for a binomially distributed trait, such as coral settlement, is straightforward, see Gilmour *et al.* (1985); Foulley *et al.* (1987); Vazquez *et al.* (2009); Biscarini *et al.* (2014, 2015). The desired quantity is the among-sire variance, denoted as τ^2, which can be estimated using a generalized linear mixed model with a binomial error distribution. Although this a departure from the standard threshold approach for estimating the heritability of binomial traits, it is now fairly common in the quantitative genetics literature, see Foulley *et al.* (1987) and Vazquez *et al.* (2009).

Suppose we have binary observations $y_{ij} \in \{0, 1\}$ where i index units (sires) and j indexes observations within units. The model is simple Bernoulli sampling, parameterized by log

odds:

$$P(y_{ij} = 1) = \frac{1}{1 + \exp(-\psi_{ij})} . \tag{1}$$

We will assume that the log odds have a sire-level random effect:

$$\psi_{ij} = \alpha + \beta_i , \quad \beta_i \sim N(0, \tau^2) .$$

Thus we have a simple binary logit model with a single random effect. A standard result on logit models is that we can represent the outcomes y_{ij} as thresholded versions of an latent continuous quantity z_{ij} (Holmes *et al.*, 2006):

$$y_{ij} = \begin{cases} 1 & \text{if } z_{ij} \geq 0, \\ 0 & \text{if } z_{ij} < 0. \end{cases}$$
$$z_{ij} = \alpha + \beta_i + \varepsilon_{ij} ,$$

where ε_{ij} follows a standard logistic distribution. Note this non-standard form of latent-threshold model, wherein the errors ε_{ij} are logistic rather than normally distributed. Upon integrating out the z_{ij}'s (which are often referred to as latent or data-augmentation variables), we recover exactly the logistic regression model of Equation (1) with a sire-level random effect.

In light of this, we can interpret narrow-sense heritability in terms of the ratio of predictable to total variation in our logistic random-effects model. This is often referred to as Bayesian R^2, by analogy with the classical coefficient of determination in a regression model:

$$R^2 = \frac{\mathrm{var}(\beta_i)}{\mathrm{var}(z_{ij})} = \frac{\mathrm{var}(\beta_i)}{\mathrm{var}(\beta_i) + \mathrm{var}(\varepsilon_{ij})} = \frac{\tau^2}{\tau^2 + \pi^2/3} ,$$

exploiting the facts that the β_i and ε_{ij} are independent and that the variance of the standard logistic distribution is $\pi^2/3$. The above equation for the Bayesian R^2 is the narrow-sense

heritability for the animal model. Therefore, the among-sire variance can be transformed into an approximation of narrow-sense heritability under the sire model by multiplying the Bayesian R^2 by four, see Foulley *et al.* (1987) and Vazquez *et al.* (2009) for a more detailed derivation and Lynch & Walsh (1998) for a discussion of the assumptions this approximation relies on.

However, under this model, determining whether statistical support exists for an among-sire variance greater than zero remains a challenge. Traditionally, an approach to the problem would be to fit two models, one where τ^2, the among-sire variance, is a free parameter and one where it is constrained to zero. These models can then be compared, and model selection performed, using a likelihood ratio test, or in this case the difference in each model's deviance, which is equivalent to a likelihood ratio test for nested models. Although, critically, this is a special kind of likelihood ratio test because the null hypothesis resides on the edge of the parameter space. The large sample reference distribution for this type of test is usually considered to be a 50% mixture of a point of mass at zero and a $\chi^2(1)$ (Self & Liang, 1987). However there is still substantial debate in the literature about what mixture should be used – e.g., Crainiceanu *et al.* (2003) – and it is not clear whether any of these mixtures are valid null distributions for finite sample sizes.

Instead, our approach is to construct a permutation-based method for calculating a p value for the likelihood ratio test and performing model selection. This test is simple to implement, as it only involves randomly shuffling the identity of each offspring's sire a large number of times (say, 500) and re-fitting the random-effects model to each shuffled data set. This avoids making assumptions about the asymptotic distribution of the test statistic that may fail to hold for finite sample sizes.

Monte Carlo simulation for the likelihood ratio test

Our simulations assume a fixed probability of settlement, p_{settle}, to be equal across all sires, in this case $p_{settle} = 0.285$ (the global mean), and simulate 1,000 data sets where the number of offspring for each sire in each of three bins is drawn from a negative binomial distribution

with $\mu = 4.63$ and size = $\mu^2/(\sqrt{12.63} - \mu)$, again these are the empirically observed values across sires. The resulting 1,000 data sets have the same structure as the observed data, but the only among sire variability comes from sampling, the true $\tau^2 = 0$. For each simulated data set, we calculated the likelihood-ratio test statistic. This provides a Monte Carlo approximation to the true sampling distribution of the test statistic under the null.

Power analysis

With any novel experimental design, it is desirable to construct a method for estimating its statistical power. Using the Monte Carlo approach designed to calculate p-values for likelihood ratio tests, we can simulate data sets with an arbitrary number of sires, number and variance in offspring, among-sires variance, and number of bins. By repeatedly simulating data sets using fixed combinations of these parameters, the statistical power is simply the fraction of times we correctly reject the null hypothesis. Similarly, the false positive rate is the fraction of times we falsely reject the null hypothesis.

Implementation

All code and data developed for this study are available in the R package multiDimBio (Scarpino *et al.*, 2014). The statistical models were fit using the R packages stats in R version 3.2.1 (R Core Team, 2015) and lme4 version 1.1-8 (Bates *et al.*, 2015).

Results

Sire Independence

Each sire was determined to be a unique multilocus genotype (MLG) across the six microsatellite loci indicating that no clones were collected (Table 2). In order to ensure that each sire could be considered independent, we calculated the Probability of Identity at each locus and found that these probabilities ranged from 3.2E-01 for a single locus down to 2.0E-06 when all six loci are considered and therefore each sire was considered independent.

Table 2. Summary of paternity assignment results. Values are the microsatellite lengths for each of six loci from Davies *et al.* (2014).

	Locus 1		Locus 2		Locus 3		Locus 4		Locus 5		Locus 6	
Sire	MaMS8	MaMS8.1	Sev5	Sev5.1	Mfav4	Mfav4.1	Mfav6	Mfav6.1	Mfav7	Mfav7.1	Sev8	Sev8.1
A	200	200	280	322	379	379	391	391	453	465	190	196
B	200	203	292	322	379	379	389	391	471	486	187	190
C	200	200	283	313	375	375	419	429	453	471	190	193
D	197	200	301	322	375	379	423	423	465	486	190	196
E	200	200	283	316	375	391	389	389	453	474	190	193
F	197	197	307	313	375	375	391	391	462	471	190	202
G	197	200	301	328	379	379	391	391	474	474	193	205
H	197	200	280	307	383	383	389	389	453	453	190	193
1J	197	200	280	313	379	379	389	389	477	498	193	193

Parentage

Larvae that amplified at > 2 loci were considered successful amplifications. A total number of 55 recruits (binary successes) were collected and of these 47 were amplified and 37 were assigned parentage. A total number of 129 swimming larvae (binary failures) were extracted and of these 112 amplified successfully and 81 were assigned parentage.

Monte Carlo simulation for the likelihood ratio test

To test whether the procedure proposed in this study provided any benefits over the traditional approach to performing a likelihood ratio test, we first simulated the true sampling distribution of the likelihood ratio statistic under the null hypothesis. This was accomplished by repeatedly simulating data from a model where the true among-sire variance (τ^2) was zero. The cumulative distribution function (CDF) of this random variable is shown as a black curve (actual null) in Figure 3. We then calculated two approximations to this sampling distribution; these CDFs are also plotted in Figure 3. First, the red curve (theoretical null) shows a mixture distribution of a point mass at 0 (with probability 0.5) and $\chi^2(1)$ random variable (with probability 0.5). This is the asymptotic approximation to the true null used in the traditional likelihood-ratio test of a variance component in a mixed-effects model. Second, the dotted grey curve (permutation null) shows the estimated null distribution obtained by running the permutation test on a single simulated data set. The permutation null is clearly a better approximation to the actual null than is the theoretical null, whose distribution is shifted to the right. This fact suggests that–at

least for data sets similar to ours–the asymptotic approximation is too conservative, and will therefore lead to reduced power at a specified false-positive rate.

Statistics

Using the described experimental design and statistical methods, we were unable to detect a significant random effect of sire, although there was a trend in overall variation in early settlement among sires (Figure 2). However, by bootstrapping the data, we were able to obtain an estimated τ^2 of approximately 0.176 (0.42 standard deviation), corresponding to a narrow-sense heritability of around 0.2 (95% CI 0.0 - 1.0). Considering the number of sires used and offspring sampled in our study, the true narrow-sense heritability would have to be well above 0.6 to achieve 80% power (Figure 4a). Nevertheless, this experimental set up should be sufficiently powered to correctly fail to reject the null hypothesis if in fact the true among sire variance was zero (Figure 4b).

Power analysis

Power analysis results suggest that increasing the number of sires is the most effective mechanism to increase statistical power. Unfortunately, for heritabilities less then 0.4, very large numbers of sires will be required. The intuition is that substantial amounts of variability between sires is expected just due to sampling alone, and therefore statistical support for a non-zero heritability requires large sample sizes. Despite the lack of statistical power, this approach does have the desirable property of low false positive rates. For example, even with nine sires, we expect to have a nearly 90% chance of failing to reject the null hypothesis on data sets simulated with an among-sire variance equal to zero (Figure 4b). Lastly, if sequencing additional offspring is an option, statistical power can be improved (Figure 5).

Discussion

In this paper, we present an experimental and statistical methodology for estimating the heritability of traits in non-model, highly fecund organisms. We applied this approach

to determine whether settlement is a heritable trait in the reef-building coral *O. faveolata*. Although we did not find statistical support for a non-zero, heritability in this trait, a power analysis suggests we lacked a sufficient number of individuals. Our computational method includes code for fitting model parameters, performing model selection using a permutation test, and calculating the expected statistical power for proposed or completed studies. The power calculation method is especially important for studies requiring animal care and use approval and/or those with complex or expensive collection demands.

Previous work suggests that heritable variation exists for a variety of traits across many marine organisms (Foo *et al.*, 2012; Johnson *et al.*, 2010; Kelly *et al.*, 2013; Lobon *et al.*, 2011; McKenzie *et al.*, 2011; Parsons, 1997), including corals (Kenkel *et al.*, 2011; Meyer *et al.*, 2009). These studies have found significant heritability for nearly every trait measured in corals (Kenkel *et al.*, 2011; Meyer *et al.*, 2009, 2011; Carlon *et al.*, 2011), but see Csaszar *et al.* (2010). In fact, one study specifically quantified the additive genetic variance in settlement rates of the Pacific reef-building coral *Acropora millepora* and found $h^2 = 0.49$, however no variance around this mean was estimated (Meyer *et al.*, 2009). It would not be surprising from an evolutionary standpoint if an ecologically important life-history trait such as larval settlement was heritable in other coral species, such as *O. faveoalta*. However, in this study we were unable to detect heritable variation, likely due to insufficient numbers of individuals.

There is a rich quantitative genetics literature on estimating the heritability of binomial traits dating back to Wright (1917) and Fisher (1918); however, the first use of Generalized Linear Models fit to observed presence/absence data is from Gilmour *et al.* (1985), with key future contributions from Foulley *et al.* (1987) and Vazquez *et al.* (2009). These methods were originally developed for agricultural breeders, where fewer constraints exist on the number of families used to estimate the heritability–for example the viability of poultry (Robertson & Lerner, 1949), common genetic disorders of Holstein cows (Uribe *et al.*, 1995) and root vigor in sugar beets (Biscarini *et al.*, 2014, 2015). Uribe et al. (1995) estimated sire and residual variance components using REML modeling of 7416 paternal half-sib cows and found that

heritability of common diseases in cows ranged from 0 to 0.28. These sorts of numbers are unreasonable to sample in natural populations of corals since parentage is rarely known unless controlled crosses are completed and then the costs associated with genotyping thousands of individuals are prohibitive.

A pair of recent papers by Biscarini et al. (2014 and 2015) developed a cross-validation based algorithm for selecting single nucleotide polymorphisms that maximally classified sugar beets into high and low root vigor. Therefore, our principle contribution is in terms of model selection, in the form of a permutation test to determine whether statistical support exists for a non-zero narrow-sense heritability, and the methods application to non-model organisms. In such organisms, where breeding, collection, and/or budgetary constraints may exist, such a model-selection procedure is essential.

Our approach has three important caveats. First, as stated in the methods section, one cannot disentangle additive variation due to sire from dam-specific sire effects under the sire model Lynch & Walsh (1998). Therefore, conservatively, heritability estimates using our approach should be considered estimates of broad-sense heritability. Second, our methods are somewhat lacking in statistical power. For heritabilities thought to be typical of studies in non-model organisms, well more than 50 individuals may need to be typed across 9 sires, see Figures 4a and 5. However, our methods perform very well with respect to minimizing the type-I error rate, see Figure 4b. Lastly, as stated in the methods, the accepted approach–based on mixtures of chi-squared distributions–has even less statistical power and was a poor approximation to our observed null distribution. Future work should focus on adapting existing methods and developing new methods to allow for smaller sample sizes. This effort is meant to be a project that will grow and develop organically; therefore, we welcome suggestions and contributions and plan regular updates to the statistical methods.

Acknowledgements

The authors acknowledge funding from the Santa Fe Institute and the Omidyar Group to SVS. Funding was also provided by the National Science Foundation grant DEB-1054766 to MVM, NSF grant DMS-1255187 to JGS, a departmental start-up grant from the Section of Integrative Biology at the University of Texas at Austin to SWD and the PADI Foundation Award to SWD. In addition the Flower Garden Banks National Marine Sanctuary is acknowledged for boat time aboard the R/V Manta.

References

Baird, A.H., Guest, J.R. & Willis, B.L. (2009) Systematic and biogeographical patterns in the reproductive biology of scleractinian corals. *Annual Review of Ecology Evolution and Systematics*, **40**, 551–571.

Bates, D., Maechler, M., Bolker, B. & Walker, S. (2015) lme4: Linear mixed-effects models using eigen and s4. *R package version 11-8*, **1**.

Bay, L.K., Choat, J.H., van Herwerden, L. & Robertson, D.R. (2004) High genetic diversities and complex genetic structure in an indo-pacific tropical reef fish (chlorurus sordidus): evidence of an unstable evolutionary past? *Marine Biology*, **144**, 757–767.

Biscarini, F., Marini, S., Stevanato, P., Broccanello, C., Bellazzi, R. & Nazzicari, N. (2015) Developing a parsimonius predictor for binary traits in sugar beet (beta vulgaris). *Molecular Breeding*, **35**, 1–12.

Biscarini, F., Stevanato, P., Broccanello, C., Stella, A. & Saccomani, M. (2014) Genome-enabled predictions for binomial traits in sugar beet populations. *BMC genetics*, **15**, 87.

Carlon, D.B., Budd, A.F., Lippe, C. & Andrew, R.L. (2011) The quantitative genetics of incipient speciation: heritability and genetic correlations of skeletal traits in populations of diverging favia fragum ecomorphs. *Evolution*, **65**, 3428–47.

Crainiceanu, C.M., Ruppert, D. & Vogelsang, T.J. (2003) Some properties of likelihood ratio tests in linear mixed models. *U RL http://legacy orie cornell edu/˜ davidr/papers, Retrieved May*, **31**, 2007.

Csaszar, N.B., Ralph, P.J., Frankham, R., Berkelmans, R. & van Oppen, M.J. (2010) Estimating the potential for adaptation of corals to climate warming. *PLoS One*, **5**, e9751.

Davies, S.W., Meyer, E., Guermond, S.M. & Matz, M.V. (2014) A cross-ocean comparison of responses to settlement cues in reef-building corals. *PeerJ*, **2**, e333.

Davies, S.W., Treml, E.A., Kenkel, C.D. & Matz, M.V. (2015) Exploring the role of micronesian islands in the maintenance of coral genetic diversity in the pacific ocean. *Mol Ecol*, **24**, 70–82.

Davies, S., Rahman, M., Meyer, E., Green, E., Buschiazzo, E., Medina, M. & Matz, M. (2013) Novel polymorphic microsatellite markers for population genetics of the endangered caribbean star coral, montastraea faveolata. *Marine Biodiversity*, **43**, 167–172.

Doherty, P. & Fowler, T. (1994) An empirical test of recruitment limitation in a coral reef fish. *Science*, **263**, 935–9.

Fisher, R.A. (1918) Xv.—the correlation between relatives on the supposition of mendelian inheritance. *Transactions of the royal society of Edinburgh*, **52**, 399–433.

Foo, S.A., Dworjanyn, S.A., Poore, A.G.B. & Byrne, M. (2012) Adaptive capacity of the habitat modifying sea urchin centrostephanus rodgersii to ocean warming and ocean acidification: Performance of early embryos. *Plos One*, **7**.

Foulley, J., Gianola, D. & Im, S. (1987) Genetic evaluation of traits distributed as poisson-binomial with reference to reproductive characters. *Theoretical and Applied Genetics*, **73**, 870–877.

Gilmour, A., Anderson, R. & Rae, A. (1985) The analysis of binomial data by a generalized linear mixed model. *Biometrika*, **72**, 593–599.

Holmes, C.C., Held, L. *et al.* (2006) Bayesian auxiliary variable models for binary and multinomial regression. *Bayesian Analysis*, **1**, 145–168.

Johnson, D.W., Christie, M.R. & Moye, J. (2010) Quantifying evolutionary potential of marine fish larvae: Heritability, selection, and evolutionary constraints. *Evolution*, **64**, 2614–2628.

Kelly, M.W., Padilla-Gamino, J.L. & Hofmann, G.E. (2013) Natural variation and the capacity to adapt to ocean acidification in the keystone sea urchin strongylocentrotus purpuratus. *Glob Chang Biol*.

Kenkel, C.D., Traylor, M.R., Wiedenmann, J., Salih, A. & Matz, M.V. (2011) Fluorescence of coral larvae predicts their settlement response to crustose coralline algae and reflects stress. *Proceedings of the Royal Society B-Biological Sciences*, **278**, 2691–2697.

Lobon, C.M., Acuna, J.L., Lopez-Alvarez, M. & Capitanio, F.L. (2011) Heritability of morphological and life history traits in a pelagic tunicate. *Marine Ecology Progress Series*, **422**, 145–154.

Lynch, M. & Walsh, B. (1998) *Genetics and the analysis of quantitative traits*. Sinauer Associates: Sunderland.

McKenzie, L.A., Brooks, R. & Johnston, E.L. (2011) Heritable pollution tolerance in a marine invader. *Environ Res*, **111**, 926–32.

Meyer, E., Aglyamova, G.V. & Matz, M.V. (2011) Profiling gene expression responses of coral larvae (acropora millepora) to elevated temperature and settlement inducers using a novel rna-seq procedure. *Molecular Ecology*, **20**, 3599–3616.

Meyer, E., Davies, S., Wang, S., Willis, B.L., Abrego, D., Juenger, T.E. & Matz, M.V.

(2009) Genetic variation in responses to a settlement cue and elevated temperature in the reef-building coral acropora millepora. *Marine Ecology Progress Series*, **392**, 81–92.

Parsons, K.E. (1997) Contrasting patterns of heritable geographic variation in shell morphology and growth potential in the marine gastropod bembicium vittatum: Evidence from field experiments. *Evolution*, **51**, 784–796.

Peakall, R. & Smouse, P.E. (2006) Genalex 6: genetic analysis in excel. population genetic software for teaching and research. *Molecular ecology notes*, **6**, 288–295.

R Core Team (2015) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Robertson, A. & Lerner, I.M. (1949) The heritability of all-or-none traits: viability of poultry. *Genetics*, **34**, 395.

Romiguier, J., Gayral, P., Ballenghien, M., Bernard, A., Cahais, V., Chenuil, A., Chiari, Y., Dernat, R., Duret, L., Faivre, N., Loire, E., Lourenco, J.M., Nabholz, B., Roux, C., Tsagkogeorga, G., Weber, A.A., Weinert, L.A., Belkhir, K., Bierne, N., Glemin, S. & Galtier, N. (2014) Comparative population genomics in animals uncovers the determinants of genetic diversity. *Nature*, **515**, 261–3.

Scarpino, S.V., Gillette, R. & Crews, D. (2014) multidimbio: an r package for the design, analysis, and visualization of systems biology experiments. *arXiv preprint arXiv:14040594*.

Self, S.G. & Liang, K.Y. (1987) Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association*, **82**, 605–610.

Severance, E.G., Szmant, A.M. & Karl, S.A. (2004) Microsatellite loci isolated from the caribbean coral, montastraea annularis. *Molecular Ecology Notes*, **4**, 74–76.

Uribe, H., Kennedy, B., Martin, S. & Kelton, D. (1995) Genetic parameters for common health disorders of holstein cows. *Journal of Dairy Science*, **78**, 421–430.

Vazquez, A., Weigel, K., Gianola, D., Bates, D., Perez-Cabal, M., Rosa, G. & Chang, Y. (2009) Poisson versus threshold models for genetic analysis of clinical mastitis in us holsteins. *Journal of dairy science*, **92**, 5239–5247.

Wright, S. (1917) The average correlation within subgroups of a population. *Journal of the Washington Academy of Science*, **7**, 532–535.

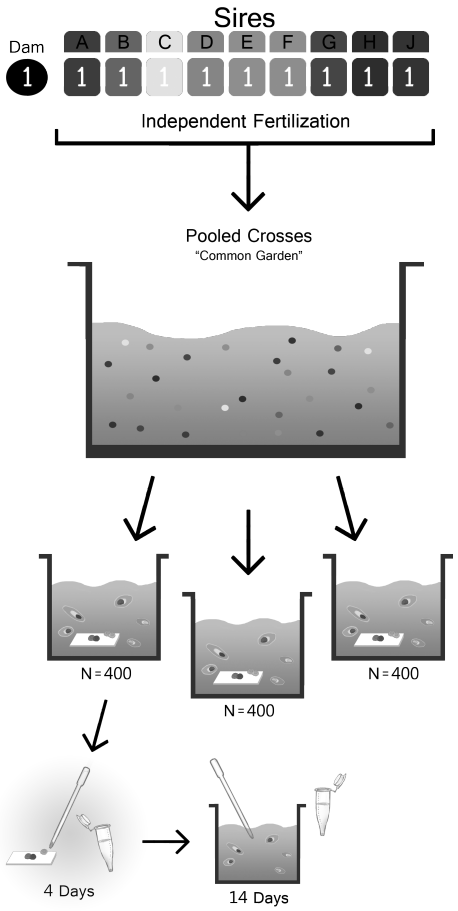

Figure 1. Diagram representing the design of the common garden experiment. First, independent fertilizations are completed for each sire and dam (in this case only one dam and nine sires are used). Second, equal quantities of fertilized embryos are pooled into one single common garden tank. This common garden is the split into three replicate tanks (N=400 larvae per tank). Settlement slides are added to each experimental tank and after 4 days the settled larvae are collected and individually preserved. Larvae were then left for an additional 10 days and settled larvae were removed every few days. N=50 larvae that remained swimming after 14 days were collected and individually preserved for genotyping, to compare their parentage to the parentage of the early-settling larvae.

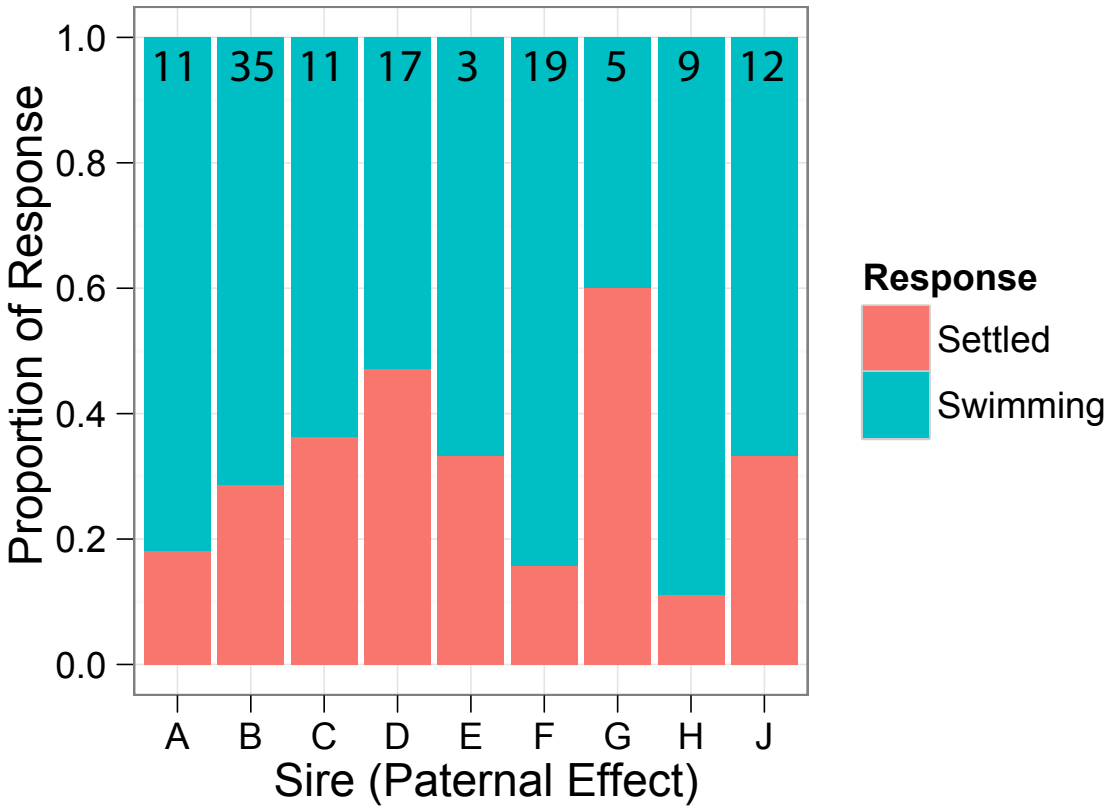

Figure 2. Proportion of settled (successes) and swimming (failures) larvae belonging to each sire. The total number of genotyped larvae assigning to each sire is indicated at the top of each bar.

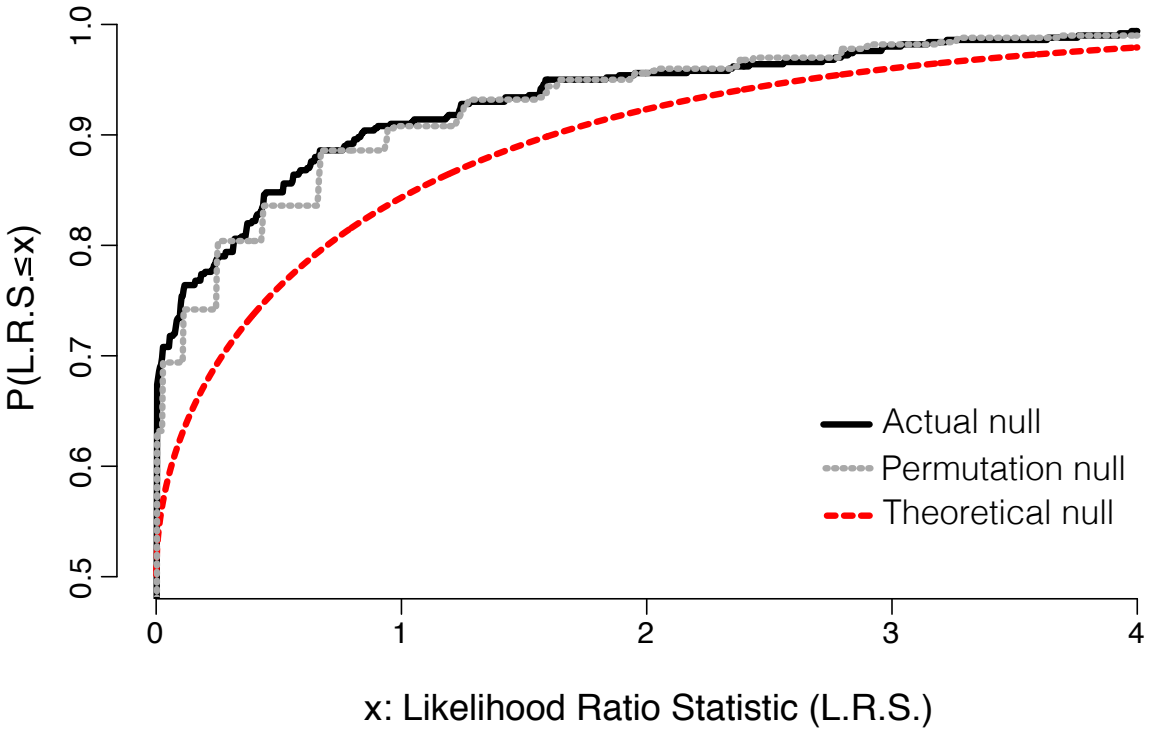

Figure 3. The cumulative distribution functions for the actual (black solid), permutation (gray dashed), and theoretical (red dashed) nulls are compared. The permutation null is a closer match to the actual null and is less conservative than the asymptotic approximation. This suggests that asymptotic approximation to the true null distribution is inappropriate for our data set.

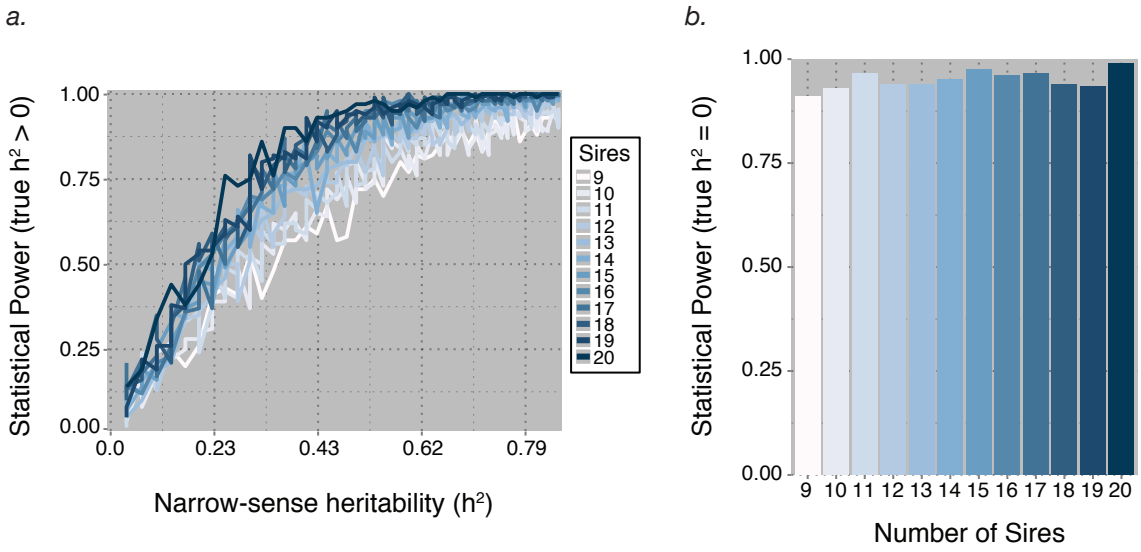

Figure 4. Power analysis for a varying number of sires. The offspring number was fixed, at $\mu = 4.63$ and size = $\mu^2/(\sqrt{12.63} - \mu)$ respectively, and the number of sires was varied between 9 and 20. In panel *a.*, the power to reject the null hypothesis of $h^2 = 0$ is plotted as a function of narrow-sense heritability (h^2), where the true value of $h^2 > 0$. In panel *b.*, the power to fail-to-reject the null hypothesis when the true value of h^2 was equal to zero is plotted for varying numbers of sires.

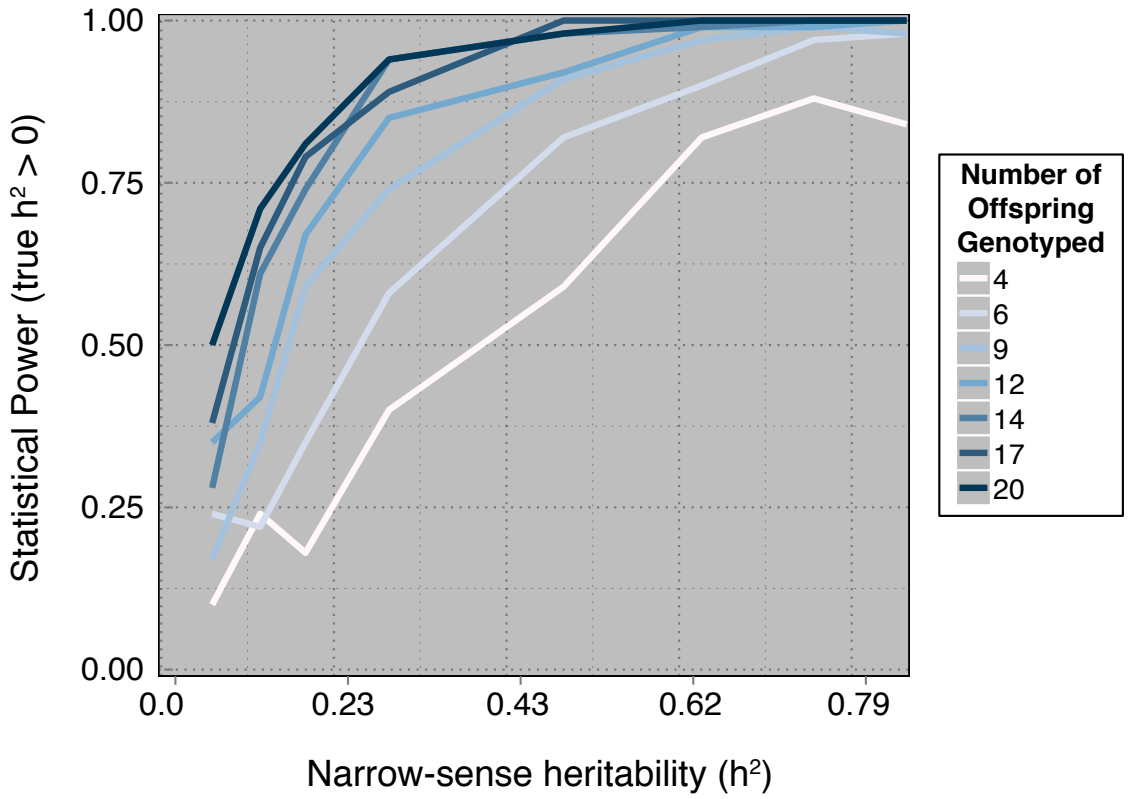

Figure 5. Power analysis for a varying number of offspring. The mean number of offspring genotyped per sire, μ, was varied between 4 and 20, while the size parameter for the negative binomial distribution was $\mu^2/(\sqrt{\mu(12.63/4.63)} - \mu)$. The number of sires was fixed at 9. The power to reject the null hypothesis of $h^2 = 0$ is plotted as a function of narrow-sense heritability (h^2), where the true value of $h^2 > 0$.