<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Evidence of opposing fitness effects of parental heterozygosity and relatedness in a critically endangered marine turtle?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Phillips, Karl P.; Jorgensen, T. H.; Jolliffe, K. G.; Richardson, D. S.</td>
</tr>
<tr>
<td><strong>Publication date</strong></td>
<td>2017-08-28</td>
</tr>
<tr>
<td><strong>Type of publication</strong></td>
<td>Article (peer-reviewed)</td>
</tr>
</tbody>
</table>
| **Link to publisher's version** | [http://dx.doi.org/10.1111/jeb.13152](http://dx.doi.org/10.1111/jeb.13152)  
Access to the full text of the published version may require a subscription. |
| **Rights** | © 2017 European Society for Evolutionary Biology. This is the peer reviewed version of the following article: Phillips, K. P. et al (2017), Evidence of opposing fitness effects of parental heterozygosity and relatedness in a critically endangered marine turtle?. J. Evol. Biol., which has been published in final form at http://dx.doi.org/10.1111/jeb.13152. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving. |
| **Embargo information** | Access to this article is restricted until 12 months after publication by request of the publisher. |
| **Embargo lift date** | 2018-08-28 |
| **Item downloaded from** | [http://hdl.handle.net/10468/4690](http://hdl.handle.net/10468/4690) |

Downloaded on 2019-10-05T17:03:36Z
Evidence of opposing fitness effects of parental heterozygosity and relatedness in a critically endangered marine turtle?

Karl P. Phillips¹²³*, Tove H. Jorgensen¹⁴, Kevin G. Jolliffe⁵, & David S. Richardson¹

¹ School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich, NR4 7TJ, UK
² NERC Biomolecular Analysis Facility (NBAF), Department of Animal and Plant Sciences, University of Sheffield, Western Bank, Sheffield, S10 2TN, UK
³ Evolutionary Biology Group, Institute of Environmental Biology, Faculty of Biology, Adam Mickiewicz University, Poznań, Poland
⁴ Department of Bioscience, Aarhus University, DK-8000, Aarhus, Denmark.
⁵ Cousine Island, P.O. Box 977, Victoria, Mahé, Republic of Seychelles

* Present address: School of Biological, Earth & Environmental Sciences, University College Cork, c/o Marine Institute, Furnace, Newport, Co. Mayo, Ireland

Correspondence: Karl P. Phillips, School of Biological, Earth & Environmental Sciences, University College Cork, c/o Marine Institute, Furnace, Newport, Co. Mayo, Ireland

Phone: +353 (0) 98 42313
Fax: +353 (0) 98 42340
E-mail: karl.phillips@ucc.ie

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/jeb.13152

This article is protected by copyright. All rights reserved.
Evidence of opposing fitness effects of parental heterozygosity and relatedness in a critically endangered marine turtle?

Abstract

How individual genetic variability relates to fitness is important in understanding evolution and the processes affecting populations of conservation concern. Heterozygosity-fitness correlations (HFCs) have been widely used to study this link in wild populations, where key parameters that affect both variability and fitness, such as inbreeding, can be difficult to measure. We used estimates of parental heterozygosity and genetic similarity ('relatedness') derived from 32 microsatellite markers to explore the relationship between genetic variability and fitness in a population of the critically endangered hawksbill turtle, Eretmochelys imbricata. We found no effect of maternal MLH (multilocus heterozygosity) on clutch size or egg success rate, and no single-locus effects. However, we found effects of paternal MLH and parental relatedness on egg success rate that interacted in a way that may result in both positive and negative effects of genetic variability. Multicollinearity in these tests was within safe limits, and null simulations suggested the effect was not an artefact of using paternal genotypes reconstructed from large samples of offspring. Our results could imply a tension between inbreeding and outbreeding depression in this system, which is biologically feasible in turtles: female-biased natal philopatry may elevate inbreeding risk and local adaptation, and both processes may be disrupted by male-biased dispersal. Although this conclusion should be treated with caution due to a lack of significant identity disequilibrium, our study shows the importance of considering both positive and negative effects when assessing how variation in genetic variability affects fitness in wild systems.
Keywords: heterozygosity-fitness correlations; negative heterozygosity-fitness correlations; inbreeding depression; outbreeding depression; microsatellites; hawksbill turtle; *Eretmochelys imbricata*

**Introduction**

How genetic variability relates to individual fitness is a fundamental question in evolutionary biology (Charlesworth & Charlesworth, 1999), with potential implications for a wide range of life history parameters that affect survival and reproductive success (reviewed in Chapman et al., 2009). It is also an important concept in conservation management, with practical implications for populations of conservation concern that may be facing challenges arising from depleted genetic variation (Crnokrak & Roff, 1999; Keller & Waller, 2002; Gooley et al., 2017). A common method for studying the relationship between individual genetic variability and fitness has been to test for correlations between individual heterozygosity and fitness parameters, so called ‘heterozygosity-fitness correlations’ or HFCs (Hansson & Westerberg, 2002; Chapman et al., 2009; Miller & Coltman, 2014). HFCs have been reported in a variety of taxa (Chapman et al., 2009), but the relative roles of the hypothesised explanatory mechanisms behind the correlations remains a topic of debate (e.g. Hansson & Westerberg, 2008; Szulkin et al., 2010). With the widespread use of [assumedly neutral] microsatellite markers for HFC studies, this discussion has focussed on the relative roles of ‘general’ and ‘local’ effects (e.g. Hansson & Westerberg, 2002; Rodriguez-Quilon et al., 2015), and the associated question of the link between marker-based heterozygosity and inbreeding coefficients (e.g. Balloux et al., 2004; Berenos et al., 2016; Nietlisbach et al., 2017), though these models are not mutually exclusive. A third mechanism, ‘direct effects’, was applicable to older study techniques such as allozymes (e.g. David, 1998), but is not usually considered for microsatellites, which are typically treated as neutral and not directly subject to selection (e.g. Li et al., 2002; Li et al., 2004). Under the ‘general effects’ model, average heterozygosity across a number of independent loci is used as an estimate of genome-wide heterozygosity, which itself is treated as a proxy for an...
individual’s level of inbreeding. Individuals that are more inbred are predicted to have lower genome-wide heterozygosity and, consequently, lower fitness due to increased expression of deleterious recessive alleles and loss of heterosis (‘inbreeding depression’; Charlesworth & Charlesworth, 1987, 1999). However, this interpretation is controversial, as even when calculated from large panels of markers, multilocus heterozygosity (MLH) may be a poor correlate of pedigree inbreeding coefficients (e.g. Balloux et al., 2004). Under the ‘local effects’ model, a given marker demonstrates an HFC because it is in linkage disequilibrium with a functional locus (Hansson & Westerberg, 2002; Szulkin et al., 2010). This model is less dependent on individual inbreeding levels than general effects, but there is widespread theoretical concern about how to interpret such effects, compounded by statistical challenges in demonstrating them robustly (Szulkin et al., 2010).

Many HFC studies have reported positive linear effects (Chapman et al., 2009; Küpper et al., 2010). Such positive effects may have been overrepresented due to publication bias (Coltman & Slate, 2003; bias may be diminishing – see Chapman et al., 2009), biases arising from the properties of the genetic markers used (Küpper et al., 2010), and because of a tendency for HFC studies to be conducted on small populations that have high inbreeding variance (see Coltman & Slate, 2003; Chapman et al., 2009; Küpper et al., 2010). The other end of the heterozygosity spectrum – negative HFCs – has received less attention (but see Szulkin & David, 2011). Negative multilocus HFCs may represent outbreeding depression, where population admixture breaks up coadapted gene complexes or disrupts local adaptation (Templeton et al., 1986; Waser, 1993). This phenomenon has been observed in captive populations (Lacy et al., 1993), but its importance among wild populations is less clear (Marshall & Spalton, 2000; Szulkin & David, 2011). Somewhere between deleterious inbreeding and deleterious outbreeding there should be an optimal level of outcrossing that maximises fitness. This has been well-demonstrated in plants (Waser & Price, 1989, 1994; Willi & van Buskirk, 2005). Studies that have demonstrated both inbreeding and outbreeding depression on animals are few (e.g. Marshall & Spalton, 2000; Neff, 2004; Escobar et al., 2008; see also Edmands, 2007), but have highlighted important processes. For example, inbreeding and outbreeding may act
on the same trait, or simultaneously on different traits, and directions of HFC may differ between age classes and sexes (e.g. Marshall & Spalton, 2000; Escobar et al., 2008; Olano-Marin et al., 2011b).

An individual’s genetic variability may affect the fitness of its offspring, e.g. through differential fertilisation success (e.g. Bretman et al., 2009; Fitzpatrick & Evans, 2009), egg hatchability (e.g. Keller, 1998; Cordero et al., 2004), or because variability correlates with the quality of parental care (e.g. Richardson et al., 2004; Brouwer et al., 2007). The genetic similarity/relatedness (henceforth ‘relatedness’) of the parents – which is directly related to offspring genetic variability – may also be important (e.g. Bensch et al., 1994; van de Casteele et al., 2003). Thus, the success of a given breeding event could be determined by the heterozygosity of each parent, but also by the average heterozygosity of the offspring that they produce as a result of their relatedness (Cordero et al., 2004).

HFC studies are biased towards species that mature quickly and have short lifespans, towards mammals and birds, and against species of high fecundity (Chapman et al., 2009). However, it is important for the development of any theory in evolutionary biology that it be tested against a range of life history backgrounds. Here, we test for correlations between individual genetic variability and fitness in a population of the critically endangered hawksbill turtle (Eretmochelys imbricata; family Cheloniidae), a long-lived, slow-to-mature, migratory reptile. We focus on two parameters that relate directly to marine turtle fitness: the number of eggs in a clutch and the proportional success of those eggs. For clutch size, we test for correlations with maternal heterozygosity, whereas for egg success we also test for correlations with paternal heterozygosity and parental relatedness. We also compare multilocus (general effect) and single-locus (local effect) models. Our study population in the Republic of Seychelles is one of the world’s most important populations of hawksbill turtles, but has declined substantially since the islands were colonised by humans 200+ years ago (Mortimer, 1984, 2004). Assessing the link between genetic variability and
fitness in this population may help conservation managers better understand the processes driving variation in reproductive output in this species.

**Methods**

*Field methods*

Sampling of adult female and hatchling turtles was conducted on Cousine Island, Republic of Seychelles (04° 21' S, 55° 38' E), over two nesting seasons (Sep-Apr) spanning 2007-2009. Laying females and nest locations were identified by patrolling Cousine’s 1 km beach hourly between 6 am and 6 pm (hawksbills in Seychelles nest almost exclusively by day; Diamond, 1976). Observed females were measured (curved carapace length), tagged (unique titanium tags issued by Seychelles Island Foundation), and sampled (6 mm sterile biopsy from trailing edge of foreflipper). Cousine’s conservation managers relocate the majority (> 95%) of clutches to ‘erosion safe’ beach zones and line nest chambers with nets to limit predation from Oocypode crabs (Hitchins et al., 2004). Eggs were counted during this process, which was undertaken within 12 hrs of laying (later disturbance can increase egg mortality; Parmenter, 1980). Because protective nets prevent hatchlings escaping, nests were checked daily for signs of activity from 10 days prior to the projected hatching date. Upon release, live hatchlings were counted, and tissue samples (2 mm biopsy) were taken from the marginal scute of 20 randomly-chosen hatchlings per clutch (all hatchlings if fewer than 20 emerged). The number of unhatched eggs and dead hatchlings remaining in the nest were counted, with unhatched eggs assigned as either ‘developed’ (embryo in evidence) or ‘undeveloped’ (no evident embryo).

*Molecular techniques*

Samples were genotyped at 32 microsatellite loci (details in Phillips et al., 2013). Female genotypes were only used in downstream analyses if ≥ 29/32 loci amplified. For each hatchling, its genotype for a given multiplex was not used downstream if more than four loci from a multiplex (10-11 loci) failed.
to amplify, and its whole genotype was removed if two multiplexes were discounted or if more than ten loci failed in total.

**Parentage analysis and reconstruction of paternal genotypes**

We used COLONY 2.0.4 (Wang & Santure, 2009; Jones & Wang, 2010) to identify clusters of offspring that shared a father, and to reconstruct the genotypes of these males. COLONY parameters were as in Phillips *et al.* (2013). We henceforth use ‘family’ to refer to a clutch or group of clutches produced by a single female in a given year. COLONY reconstructs genotypes on a locus-by-locus basis and provides a confidence value for each reconstruction. As in Phillips *et al.* (2013), when assembling paternal multilocus genotypes we only incorporated single-locus genotypes with confidence ≥ 0.90, and only used multilocus genotypes in downstream analyses if they contained ≥ 29/32 loci and were reconstructed from ≥ 10 offspring. All three of these thresholds bear on the fact that paternal genotypes are easier to reconstruct if the male is heterozygous or is dissimilar from the female.

Alternative thresholds might have been chosen: increasing the required minimum numbers of offspring or reconstructed loci reduces reconstruction bias, and increasing per-locus confidence increases bias (because heterozygous/dissimilar genotypes hit the threshold more easily; Appendix S1). However, we opted for consistency with the previous work, given that simulations undertaken for the present study showed that the chosen thresholds did not problematically elevate Type I error risk for the analyses we conducted (see below). Note that any conclusions arising from paternal genotype data are only applicable to families that pass the minimum offspring criterion.

All loci satisfied assumptions of Hardy-Weinberg and linkage equilibria (GENEPOP v4.1; Raymond & Rousset, 1995), and had null allele frequencies < 0.1 (CERVUS v3.0.3; Marshall *et al.*, 1998). We calculated the $g_2$ measure of identity disequilibrium in the programme RMES (David *et al.*, 2007; Oct 2009 version), as recommended by Szulkin *et al.* (2010). Significant identity disequilibrium means that locus states (heterozygous or homozygous) correlate within individuals, which suggests
that heterozygosity across the marker panel correlates with individual inbreeding (Szulkin et al., 2010).

Genetic predictors

For each genotyped adult, we calculated multilocus heterozygosity (MLH) as standardised heterozygosity (SH), which gives all loci equal weighting but corrects for missing genotypes (Coltman et al., 1999). To aid future reviews and meta-analyses we also performed all multilocus analyses with two alternative metrics: ‘internal relatedness’ (Amos et al., 2001) and ‘homozygosity by loci’ (Aparicio et al., 2006). The results did not differ substantively from those using SH (Appendix S2).

Data on single-locus heterozygosity (SLH) for each individual were coded as a series of 0’s (homozygote loci) and 1’s (heterozygote loci), and were then standardised by marker variability to reduce bias when estimating single locus partial regression slopes (see Szulkin et al., 2010). Missing single-locus genotypes were replaced with the population-level expected heterozygosity for the respective locus (Szulkin et al., 2010).

Genetic similarity between all observed pairings for which we were able to reconstruct paternal genotypes (henceforth ‘multilocus parental relatedness’) was quantified using the relatedness metric of Queller & Goodnight (1989) in COANCESTRY v1.0.1 (Wang, 2011), based on allele frequencies taken from the COLONY output. As a measure of single-locus parental similarity/relatedness, we calculated the proportion of a pair’s offspring expected to be homozygous at a given locus. Missing values were replaced with the population expected homozygosity for the respective locus.

Fitness response variables

All HFC analyses were conducted in R (v3.0 or later; R Development Core Team, 2008), using linear mixed models in the package ‘lme4’ (Bates et al., 2014) and tests of regression slope significance in the package ‘ImerTest’ (Kuznetsova et al., 2013). We tested whether maternal heterozygosity...
predicted clutch size, with female identity as a random effect (errors were near-enough normally distributed for us to favour Gaussian over Poisson for a more intuitive interpretation). We included every nest for which we knew the female’s genotype and controlled for a relationship between maternal body size and clutch size (Appendix S3). There was no significant correlation between maternal heterozygosity and body size (N = 70, P = 0.86; note that testudine growth is indeterminate).

We tested whether maternal heterozygosity, paternal heterozygosity and parental relatedness predicted the proportion of eggs in a clutch from which surviving hatchlings emerged (‘emergence success’, logit-transformed), with lay date and incubation duration included as control variables and pair identity as a random effect. Use of alternative egg success metrics (fertilisation success and hatching success) produced near-identical interpretations (Appendix S4). We excluded clutches where no egg produced a visible embryo, as either A) these lacked paternal genotype data, or B) an unexplained total clutch failure among a female’s otherwise ‘normal’ clutches implies an external factor overwhelming other processes. We also excluded multiple-paternity families, as we cannot know the contributions of each male to the proportion of eggs that fail. However, we re-ran the maternal HFC analysis of emergence success with all nests included. Multiple and single paternity families did not have significantly different emergence success (linear mixed model: difference = 0.35 ± 0.24 (SE); likelihood ratio test: df = 6,7, P = 0.14).

We used a corrected Akaike Information Criterion (AICc; Akaike, 1974; Hurvich & Tsai, 1989) model ranking approach, implemented in the R package ‘MuMIn’ (Bartón, 2013), to compare thirteen multilocus models of interest for egg success. These models included main linear effects, quadratic effects, and pairwise linear interactions for the three multilocus predictors, with up to three predictors allowed into a model (see Table 2). A model was considered the nominal ‘best’ if it was ≥ 2.00 AICc units clear of the next best model. Otherwise, we considered the models comprising the top two units of AICc collectively. For each model, we calculated marginal R^2 (Nakagawa & Schielzeth, 2013) to indicate the amount of variance explained by the model’s fixed effects, and the
Akaike weight, a measure of the model's explanatory power relative to other models. Though paternal MLH and parental relatedness were correlated in our dataset (Pearson $r = -0.297$), multicollinearity was well within safe limits after all continuous predictors were zero-centred (max. variance inflation factor = 1.17; max. kappa = 2.51; concern would begin at values of 2.5 and 10.0 respectively). However, as a safeguard, any significant effects involving paternal MLH or multilocus parental relatedness were re-tested using A) bootstrapping that held each of the two parameters constant while re-sampling the other, and B) delete-one jack-knifing for each family (Appendix S5).

We initially included an interaction term between genetic variability and study season when testing multilocus predictors, but found no support either for this or for any main effect of season (data not shown).

**Single locus effects**

To test for effects of heterozygosity associated with specific loci, we used likelihood ratio tests to compare models with a multilocus metric fitted as a linear expression against respective models with all 32 single loci fitted simultaneously as covariates (Küpper et al., 2010; Szulkin et al., 2010). Only if this test is significant should a single-locus model be examined for loci with partial regression slopes significantly different from zero (Szulkin et al., 2010). AIC<sub>C</sub> is unsuitable for this test because the majority of loci are expected to be of low explanatory power and to swamp the metric. We performed these analyses for both egg number and egg success, with separate analyses for maternal heterozygosity, paternal heterozygosity, and parental relatedness on the latter. Because our sample size relative to the number of loci placed the maximum single-locus models in danger of over-fitting, we interpreted any significant likelihood ratio $P$-values relative to those derived from simulated null datasets (see below).
Simulations

To test whether inferences based on paternal genotypes were biased by the reconstruction process, we used the observed allele frequencies to generate 1000 null datasets with family sizes and locations of genotyping failures exactly matching the observed data, and with random genotyping errors (per-locus rates as estimated above) applied to females and hatchlings (details in Appendix S6). We ran each null dataset through our analysis pipeline (parentage analysis, genotype reconstruction, and statistical analyses), and compared the observed values of key summary metrics (e.g. $\Delta\text{AIC}_C$, $R^2$, and coefficient slopes) to their simulated distributions.

Results

We genotyped 95 adult females and 2455 hatchlings. After excluding 15 clutches attributed to 10 unsampled females, this was reduced to a sample of 142 clutches produced by 70 genotyped females (mean clutches per female = 2.0 ± 1.2 (SD); max. = 5). All runs of COLONY converged on the same parentage assignments. Of the 84 males inferred as contributing paternity, we were able to reconstruct 64 genotypes that met our confidence criteria. Eight of our genotyped families showed multiple paternity. Consistent with Phillips et al. (2013), no cases were observed where the paternity changed between a given female’s clutches within a season. One male sired offspring in both years of the study, but no males fertilised more than one female per year. Our final sample sizes were 140 clutches in 69 families for the clutch size HFC test, and 111 clutches in 56 families for egg success HFCs. Of the 55 male genotypes used in the egg success analyses, 49 were complete (i.e. 32/32 loci) and only one was missing >1 locus (29/32). After excluding the missing loci, all males had a harmonic mean per-locus reconstruction confidence ≥0.995, 32/55 males had confidence ≥ 0.99 for all accepted loci, only one male had >3 loci with confidence <0.99 (min = 0.98, no missing loci), and only 5/1752 accepted loci had 0.90 ≤ confidence < 0.95.
Identity disequilibrium

We found no evidence for overall identity disequilibrium among our loci (RMES, 10,000 iterations: $g^2 = 0.000$, $SD = 0.002$, $P = 0.366$), suggesting that our markers may better reflect their local genomic environment than individual inbreeding values.

Clutch size HFC

No maternal heterozygosity term (linear MLH, quadratic MLH, or SLH) was a significant predictor of clutch size (Table 1).

Egg success HFCs

For egg success (emergence), the ‘best’ model was the interaction between paternal MLH and parental relatedness (Tables 2, 3). Relative to the null model, this interaction resulted in an $\Delta$AIC$_C$ improvement of 7.67, and relative to the second-best model of 4.19. The second-best model had the same two main predictors but without the interaction (Tables 2, 3), and this was the only other model to improve significantly upon the null ($\Delta$AIC$_C = 3.48$). In terms of marginal $R^2$, the main-effects-only model improved upon the null by 7.9%, with the interaction bringing an additional 4.8% (Table 2).

The interaction between paternal MLH and parental relatedness means that the effective slope of paternal MLH against fitness is negative at low values of relatedness, and becomes positive above a critical relatedness value approximately 0.75 SDs above mean relatedness ($SD = 0.10$; Fig. 1). Alternatively, the interaction can be read as parental relatedness having a negative slope when paternal MLH is low and a positive slope when paternal MLH is high (Fig. 2). (See Appendix S7 for quality control plots of model residuals). This model remained the significant best when bootstrapping for paternal MLH ($\Delta$AIC$_C$ $P = 0.003$; interaction coefficient slope $P = 0.016$; Appendix S5) or parental relatedness ($\Delta$AIC$_C$ $P = 0.006$; interaction coefficient slope $P = 0.027$; Appendix S5), and was robust to delete-one jack-knifing of families ($\Delta$AIC$_C$ improvement: mean = 7.52, $SD = 1.12$, 

This article is protected by copyright. All rights reserved.
The model was also significant when compared to our null simulations, with the observed $\Delta AIC_C$, marginal $R^2$ and coefficient slope having $P$-values of 0.018, 0.018, and 0.012 respectively (Appendix S6). These $P$-values were within 0.001 of those obtained if the null datasets were analysed with the true, complete parental genotypes, indicating a negligible change in Type I error risk from the small, non-significant biases (all within 0.1 SDs of zero – see Appendix S6) in these parameters that arise from paternal genotype reconstruction.

For maternal heterozygosity, the single-locus model of emergence success did not improve upon the linear multilocus model, even without comparison to the simulated null datasets ($df = 7, 38$, $P = 0.216$; $P = 0.893$ v. 1000 null datasets). Single locus models of paternal heterozygosity and parental relatedness at first seemed to improve significantly upon their respective linear multilocus models ($P = 0.009$ and 0.001 respectively; $df = 7, 38$), but were not significant when compared to the null datasets ($P = 0.506$ and 0.127).

No interpretive differences arose from expanding the maternal heterozygosity analyses of egg success to include six families (14 clutches) showing multiple paternity and six families (six clutches) producing too few offspring for paternal genotyping (data not shown).

**Discussion**

In this study of the relationship between genetic variability and fitness in hawksbill turtles, we found no correlation between maternal heterozygosity and either clutch size or egg success. In contrast, we found support for effects of multilocus paternal heterozygosity and multilocus parental relatedness on egg success, but interacting with each other in a way that suggests that both positive and negative HFCs may be present in our study. This result needs treating with caution given that identity disequilibrium ($g_2$) was not significantly greater than zero. However, the magnitude of improvement of fit due to the interaction ($\Delta AIC_C = -4.19$ vs second-best model, with the models only differing in presence/absence of the interaction term), the fact that the second-best model is itself a substantial improvement on the third-best model ($\Delta AIC_C = -2.13$), the ‘safe’ multicollinearity...
diagnostics, and the results from our simulations all give us confidence that this result is unlikely to be an artefact of our analysis pipeline and is thus worth further discussion. We did not find support for single locus effects, but without our simulations we would likely have committed a Type I error in this analysis due to over-fitting.

The ‘best’ model of egg success featured an interaction between paternal MLH and parental relatedness. Under this model, at low to intermediate levels of paternal MLH the effect of relatedness conforms to a traditional inbreeding interpretation, with increasing relatedness correlating with decreasing fitness. However, when paternal MLH is high, the effect of relatedness disappears and may even reverse. To think about it the other way around, low paternal MLH only has a negative effect when parental relatedness is high (i.e. when inbreeding is implied). At low and intermediate parental relatedness (outbreeding), high paternal MLH (i.e. a father that is likely outbred himself) appears to exert a negative effect on fitness. One possible explanation is that if parents are unrelated (dissimilar), and are thus already producing offspring of high heterozygosity, the addition of extra variability from a particularly heterozygous father may be deleterious (outbreeding depression). If the parents are related (similar), the fact that the father is particularly heterozygous will limit how often offspring will inherit the same allele at each locus from both parents, meaning that offspring will be more heterozygous and, in theory, suffer less from inbreeding depression. How this might apply in practice, and the long-term consequences, will be affected by the heritability of heterozygosity (Mitton et al., 1993; Nietlisbach et al., 2016) and the relationship between heterozygosity and relatedness (Roberts et al., 2006), but modelling this is beyond the scope of this study.

In our analyses of reproductive success in relation to genetic variability in the hawksbill turtle, the two top multilocus models imply the presence of positive and negative HFCs in this system, a process that could exert a stabilising influence on population genetic variability (Neff, 2004). Indeed, if the genetic predictors are interpreted in a classic inbreeding context, our result could be read (cautiously) as a tension between inbreeding and outbreeding depression, with an
optimum level of outcrossing. Studies inferring both inbreeding and outbreeding depression acting on the same fitness trait in the same system are few (e.g. Waser & Price, 1989, 1994; Marshall & Spalton, 2000; Neff, 2004), although further examples of them operating on different traits in the same population exist (Olano-Marin et al., 2011a, b). The collective implication of these studies is that some form of population structure (e.g. caused by isolation-by-distance, philopatry, or founder effects) facilitates both local adaptation and elevated inbreeding risk, and that there is a trade-off between reducing inbreeding and avoiding outbreeding depression. Could such a process operate in marine turtles? Potentially, yes. Female green turtles nesting on Ascension Island show fine-scale local adaptation despite evidence for strong male-biased gene flow between nesting beaches (Lee et al., 2007; Weber et al., 2012). More generally, studies across marine turtle species have emphasised that females show natal philopatry but that males are more dispersive (reviewed in Bowen & Karl, 2007; see also Lee, 2008; Komoroske et al., 2017), which could potentially give rise to an inbreeding-outbreeding tension analogous to that seen in some plants (Waser & Price, 1989, 1994). Fine-scale sex-biased natal philopatry has also been implicated in HFCs in blue tits (*Cyanistes caeruleus*), where positive and negative HFCs have been observed in the same population (Olano-Marin et al., 2011b; see also Szulkin & David, 2011). In our system, there is no regional structuring of nesting beaches, but pairwise genetic relatedness is higher among females nesting on Cousine than among the males that fertilise them, indicating that females are more philopatric than males (Phillips et al., 2014) and creating the potential for a dispersal-adaptation trade-off. Unfortunately we do not have comparable data from other sites in Seychelles that would allow us to test the validity, spatial scale, and strength of such an inference.

We found no HFCs associated with maternal genotypes. Examples of differences between sexes in HFCs and effects of inbreeding are common, with effects more likely to be associated with females than with males (reviewed in Olano-Marin et al., 2011b). These effects can range from early-life survival of females (Coulson et al., 1999; Olano-Marin et al., 2011b) to maternally-transmitted effects on the next generation (Brouwer et al., 2007; Bebbington et al., 2016). In our case, the

This article is protected by copyright. All rights reserved.
absence of any maternal HFC may be a consequence of female breeding strategy. Marine turtle females are capital breeders, meaning they accumulate an energy reserve (‘capital’) with which to produce and provision offspring. This is a slow and variable process (most marine turtle species nest once every 2-5 years; Miller, 1997), and is known to be affected by prevailing environmental conditions (e.g. Wood & Wood, 1980; Limpus & Nicholls, 1988). Correlations between maternal heterozygosity and reproductive metrics may therefore be difficult to detect in marine turtles without controlling for variance in rates of capital accumulation (Broderick et al., 2001). However, with sufficient long-term data it might be possible to test the hypothesis that heterozygosity affects efficiency at accumulating energy capital, and thereby remigration frequency and reproductive success, or to use biomarkers of stress such as telomeres (Plot et al., 2012; Bebbington et al., 2016).

There is considerable debate in the HFC literature as to how well MLH represents individual inbreeding status (Balloux et al., 2004; Szulkin et al., 2010; Berenos et al., 2016; Nietlisbach et al., 2017), and all HFC studies are thus urged to assess the utility of MLH as an inbreeding proxy through a test for identity disequilibrium (Szulkin et al., 2010; Miller & Coltman, 2014). We did not find significant identity disequilibrium in our study, suggesting that, despite a reasonable panel of markers, MLH in our case is not a good proxy for inbreeding. Without identity disequilibrium, the effects associated with MLH are difficult to explain (Chapman & Sheldon, 2011; Miller & Coltman, 2014). However, for several reasons we are wary of dismissing the MLH-inbreeding link as an explanation for the patterns observed. First, multilocus HFCs underpinned by inbreeding can reach significance before identity disequilibrium is significant (Szulkin et al., 2010). Second, several authors have argued that HFCs, both multi- and single-locus, are more likely to be detected when conserved markers are used, and marine turtle microsatellite loci are extremely conserved (discussed further below). Third, empirical support from meta-analyses for the premise that HFC effect sizes should be larger with greater inbreeding variance is mixed. Miller & Coltman (2014) report a significant correlation between $g_2$ and HFC effect sizes, but Chapman et al. (2009), using a coarser metric but much larger sample size, did not find a relationship between inbreeding variance and HFC effect size.
Finally, our sampled population, although large and genetically well-mixed, is not homogeneous, and shows some evidence of female natal philopatry (Phillips et al., 2014). In large populations, any kind of structure creates a greater potential for inbreeding than does full mixing (Olano-Marin et al., 2011a, b; see also Szulkin et al., 2010). We have therefore discussed MLH in its traditional inbreeding interpretation, but with caution.

Several authors have argued for the importance of marker type in HFC studies. MLH measured using microsatellites located in expressed or otherwise conserved regions may be more likely to yield HFCs than MLH using anonymous/non-conserved loci by virtue of, on average, being closer to polymorphic loci under selection (Küpper et al., 2010; Olano-Marin et al., 2011a, b; Szulkin & David, 2011; Ferrer et al., 2015, 2016). The effect is statistically still a ‘general’ one, as it is the net, cumulative effect of multiple small effects (Szulkin & David, 2011). Interestingly, the majority of marine turtle microsatellite loci characterised to date show a remarkably high degree of conservation. Indeed, almost all primers designed in the family Cheloniidae, in which the extant species started diverging approximately 63 MYA (Naro-Maciel et al., 2008), amplify across multiple family members, and some even amplify in other testudine families (e.g. Shamblin et al., 2007; Lin et al., 2008). Of the 32 loci used in our study, 18 were first characterised in species other than the hawksbill (Phillips et al., 2013). Thus our finding of significant multilocus HFCs may not be out of keeping with prevailing HFC theory, given the highly conserved nature of our markers. However, marine turtle sequence evolution is known to be particularly slow (FitzSimmons et al., 1995), potentially due to long generation times and low metabolic rates in these species (Avise et al., 1992; Scott et al., 2012). This may limit the comparability between our ‘conserved’ markers and those explicitly chosen for being conserved in taxa with faster substitution rates (e.g. Küpper et al., 2010; Olano-Marin et al., 2011a, b).

Although multiple paternity is widespread in marine turtles (reviewed in Tedeschi et al., 2015), we are aware of no published study that has demonstrated a benefit to females from multiple fertilisations (e.g. Lee & Hays, 2004; Wright et al., 2013) or that females bias paternity on genetic
grounds (Phillips et al., 2013). Our study may thus have implications for marine turtle mating systems, as it shows that, hypothetically, a female’s choice of mate can affect the success of her clutches. However, it would only help explain mate choice, rather than multiple mating sensu lato, unless postcopulatory mechanisms bias paternity to the ‘best’ male (Parker, 1970; Eberhard, 1996). Moreover, it is difficult to see how a female could assess and utilise the effect we describe here, especially if the population is widely dispersed at the time of mating, as argued by Phillips et al. (2013).

A conceptual follow-up to our study would be to examine whether the effects we observe extend to candidate loci known to exhibit HFCs, such as immune genes (e.g. major-histocompatibility complex, Piertney & Oliver, 2006; Toll-like receptors, Grueber et al., 2012). It would also be informative to extend the study into additional years, as HFCs can vary in strength between breeding seasons, being stronger in ‘bad’ years that expose deleterious genotypes (e.g. Brouwer et al., 2007; Harrison et al., 2011; Annavi et al., 2014). More generally, how well MLH reflects individual inbreeding could be better tested by using much larger marker sets, such as next generation sequencing approaches that generate thousands of markers across the genome (e.g. restriction-site associated DNA, Hoffman et al., 2014; single-nucleotide polymorphisms, Berenos et al., 2016).

To summarise, our results emphasise the importance of looking for fitness effects of both low and high levels of genetic variability within a system, even on the same fitness trait. Such studies are relatively few, but play an important role in understanding how genetic variation is maintained in wild populations and how this might affect individual fitness. Our study is also of value to the HFC literature because of the distinct characteristics of this species – a long-lived, slow-to-mature, fecund reptile – which are all traits that are underrepresented by the species examined in HFC studies (Chapman et al., 2009). From a conservation perspective, our results suggest that both inbreeding and outbreeding may affect fitness in marine turtles, at least in the Seychelles population of hawksbill turtles. Whether these effects have been altered by the substantial population declines
caused by two centuries of overhunting is impossible to say from our study, but their mutual presence may highlight an important balance that could be disturbed by anthropogenic processes.

Acknowledgements

THIS SECTION HAS BEEN TEMPORARILY REMOVED TO AVOID COMPROMISING JEB’S RULES ON AUTHOR ANONYMITY AT REVIEW.

Turtle tissue samples were collected under a Seychelles Bureau of Standards permit (NRDC/0266) and exported in accordance with the Convention on International Trade in Endangered Species (export permit SC9117736A602; import permit 475521/01).

All data and analysis scripts will be uploaded to the Dryad repository.

List of supplementary material

**Appendix S1.** Exploration of parameters used for reconstructing male hawksbill turtle microsatellite genotypes from hatchling genotypes; Table S1.1 and Figs S1.1-S1.12.

**Appendix S2.** Tables S2.1-S2.4: re-runs of main multilocus analyses using two alternative MLH metrics (‘homozygosity by loci’ and ‘internal relatedness’).

**Appendix S3.** Fig. S3.1: relationship between clutch size (number of eggs laid) and female body size.

**Appendix S4.** Tables S4.1-S4.2 and Fig. S4.1: re-runs of analyses of multilocus effects on egg success using two alternative fitness metrics in addition to that used in the main document (emergence success; proportion of eggs producing a hatchling that leaves the nest): ‘fertilisation success’ (proportion of eggs producing a visible embryo) and ‘hatching success’ (proportion of eggs that hatch).

**Appendix S5.** Exploration of how the correlation between paternal multilocus heterozygosity and parental relatedness may affect our analyses, including bootstrap tests of both parameters; Figs S5.1-S5.4.

This article is protected by copyright. All rights reserved.
Appendix S6. Description of simulations used to generate null datasets against which to compare our observed effects on egg emergence success; Table S6.1.


References


Bates, D., Maechler, M., Bolker, B. & Walker, S. 2014. lme4: Linear mixed-effects models using Eigen

Bebbington, K., Spurgin, L.G., Fairfield, E.A., Dugdale, H.L., Komdeur, J., Burke, T. & Richardson, D.S.
2016. Telomere length reveals cumulative individual and transgenerational inbreeding


due to both individual and maternal inbreeding in a free-living mammal population. Mol.
Ecol. 25: 3152-3168.

4886-4907.

Bretman, A., Newcombe, D. & Tregenza, T. 2009. Promiscuous females avoid inbreeding by


bottleneched island species: a case study on the Seychelles warbler. Mol. Ecol. 16: 3134-
3144.

Chapman, J.R., Nakagawa, S., Coltman, D.W., Slate, J. & Sheldon, B.C. 2009. A quantitative review of

Chapman, J.R. & Sheldon, B.C. 2011. Heterozygosity is unrelated to adult fitness measures in a large,

This article is protected by copyright. All rights reserved.


This article is protected by copyright. All rights reserved.


This article is protected by copyright. All rights reserved.


This article is protected by copyright. All rights reserved.


This article is protected by copyright. All rights reserved.


**Table 1.** Model fit statistics for maternal heterozygosity as a predictor of clutch size. All models are linear mixed models fitted by maximum likelihood, with female identity as a random effect, and are compared using likelihood ratio tests. MLH = multilocus heterozygosity; SLH = single-locus heterozygosity; body size = curved carapace length. All analyses were conducted on 140 clutches laid by 69 females.

<table>
<thead>
<tr>
<th>Focal model</th>
<th>Tested against</th>
<th>$P$</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>Body size</td>
<td>Null model</td>
<td>&lt; 0.001</td>
<td>3,4</td>
</tr>
<tr>
<td>Body size + MLH</td>
<td>Body size</td>
<td>0.833</td>
<td>4,5</td>
</tr>
<tr>
<td>Body size + MLH$^2$</td>
<td>Body size + MLH</td>
<td>0.296</td>
<td>5,6</td>
</tr>
<tr>
<td>Body size + SLH</td>
<td>Body size + MLH</td>
<td>*0.081</td>
<td>5,36</td>
</tr>
</tbody>
</table>

* Not significant, and thus not corrected for over-fitting
Table 2. Multilocus heterozygosity (MLH) and parental relatedness models of proportional egg success (emergence), assessed by corrected Akaike Information Criterion (AICc). We assess thirteen models relative to a ‘null’ model that controls for clutch lay date and clutch incubation duration. All models are linear mixed models with parent pair identity as a random effect. AICc values in bold improve upon the null model by ≥ two units; underlined comprise the top two units of the AICc ranking.

<table>
<thead>
<tr>
<th>Model</th>
<th>AICc</th>
<th>Akaike weight</th>
<th>Marginal R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>pat.MLH × par.rel</td>
<td>-7.67</td>
<td>75.2</td>
<td>+12.7</td>
</tr>
<tr>
<td>pat.MLH + par.rel</td>
<td>-3.48</td>
<td>9.3</td>
<td>+7.9</td>
</tr>
<tr>
<td>par.rel</td>
<td>-1.35</td>
<td>3.2</td>
<td>+3.9</td>
</tr>
<tr>
<td>par.rel²</td>
<td>-1.23</td>
<td>3.0</td>
<td>+5.7</td>
</tr>
<tr>
<td>pat.MLH + mat.MLH + par.rel</td>
<td>-1.12</td>
<td>2.9</td>
<td>+7.8</td>
</tr>
<tr>
<td>Null</td>
<td>223.49</td>
<td>1.6</td>
<td>20.4</td>
</tr>
<tr>
<td>pat.MLH</td>
<td>+0.37</td>
<td>1.4</td>
<td>+1.8</td>
</tr>
<tr>
<td>mat.MLH + par.rel</td>
<td>+0.94</td>
<td>1.0</td>
<td>+3.9</td>
</tr>
<tr>
<td>pat.MLH²</td>
<td>+1.72</td>
<td>0.7</td>
<td>+2.9</td>
</tr>
<tr>
<td>mat.MLH</td>
<td>+2.28</td>
<td>0.5</td>
<td>+0.0</td>
</tr>
<tr>
<td>pat.MLH + mat.MLH</td>
<td>+2.69</td>
<td>0.4</td>
<td>+1.8</td>
</tr>
<tr>
<td>mat.MLH × par.rel</td>
<td>+3.30</td>
<td>0.3</td>
<td>+3.9</td>
</tr>
<tr>
<td>mat.MLH²</td>
<td>+3.42</td>
<td>0.3</td>
<td>+0.7</td>
</tr>
<tr>
<td>pat.MLH × mat.MLH</td>
<td>+4.29</td>
<td>0.2</td>
<td>+2.6</td>
</tr>
</tbody>
</table>

pat.MLH = paternal MLH; mat.MLH = maternal MLH; par.rel = multilocus parental relatedness

This article is protected by copyright. All rights reserved.
‘+’ indicates multi-predictor models with main effects only

‘×’ indicates two-predictor models with main effects and interaction term

Table 3. Regression slopes for multilocus predictor models of emergence success that improved significantly upon the null model (Table 2). Tests of regression slope significance were performed in the ‘lmerTest’ R package.

<table>
<thead>
<tr>
<th>Model term</th>
<th>pat.MLH × par.rel</th>
<th>pat.MLH + par.rel</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Slope</td>
<td>SE</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.54</td>
<td>0.07</td>
</tr>
<tr>
<td>Lay date</td>
<td>-5.95 × 10⁻³</td>
<td>1.76 × 10⁻³</td>
</tr>
<tr>
<td>Lay date²</td>
<td>-1.21 × 10⁻⁴</td>
<td>4.29 × 10⁻⁵</td>
</tr>
<tr>
<td>Incubation duration</td>
<td>-5.88 × 10⁻²</td>
<td>1.54 × 10⁻²</td>
</tr>
<tr>
<td>Paternal MLH</td>
<td>-1.29</td>
<td>0.58</td>
</tr>
<tr>
<td>Parental relatedness</td>
<td>-1.45</td>
<td>0.60</td>
</tr>
<tr>
<td>pat.MLH × par.rel</td>
<td>17.11</td>
<td>6.49</td>
</tr>
</tbody>
</table>

*P < 0.05; **P < 0.01; ***P < 0.001
**Figure 1.** Effect of interaction between paternal multilocus heterozygosity (MLH) and multilocus parental relatedness on proportional egg success (emergence; logit-transformed), focusing on paternal MLH. See Tables 2-3 for model fit assessments, and Appendix S7 for quality control plots of residuals. All plots show data after controlling for significant effects of clutch lay date and incubation duration. Top-left: all data points. Top-right: families with ‘low’ parental relatedness (lower third of relatedness data; relatedness ≤ -0.03). Bottom-left: families with ‘medium’ parental relatedness (mid third of relatedness data; -0.03 < relatedness ≤ 0.04). Bottom-right: families with ‘high’ parental relatedness (upper third of relatedness data; relatedness > 0.04). All lines are fitted at mean relatedness for the focal subset (-0.10, 0.01, and 0.10 respectively).

**Figure 2.** Effect of interaction between paternal multilocus heterozygosity (MLH) and multilocus parental relatedness on proportional egg success (emergence; logit-transformed), focusing on parental relatedness. See Tables 2-3 for model fit assessments, and Appendix S7 for quality control plots of residuals. All plots show data after controlling for significant effects of clutch lay date and incubation duration. Top-left: all data points. Top-right: families with ‘low’ male paternal MLH (lower third of paternal MLH data; MLH ≤ 0.95). Bottom-left: families with ‘medium’ parental relatedness (mid third of relatedness data; 0.95 < relatedness ≤ 1.04). Bottom-right: families with ‘high’ parental relatedness (upper third of relatedness data; relatedness > 1.04). All lines are fitted at mean relatedness for the focal subset (0.91, 1.02, and 1.15 respectively).
Figure 2

[Scatter plots showing the relationship between parental relatedness and egg emergence success index for different MLH levels: All data, Low male MLH, Medium male MLH, High male MLH.]