

Pedigree metrics: What are they good for?

Joel L. Pick^{1,11*}, Barbara Class^{2,11}, Paul Acker³, Anne Charmantier^{2,11}, Niels J. Dingemanse^{4,11}, Ned A. Dochtermann^{5,11}, Julien G.A. Martin⁶, Denis Réale^{7,11}, Jane M. Reid³, Holger Schielzeth^{8,11}, Celine Teplitsky^{2,11}, Kasha Strickland¹, Yimen G. Araya-Ajoy^{9,10,11}

¹ Institute of Ecology and Evolution, University of Edinburgh, Charlotte Auerbach Road, Edinburgh, EH9 3FL, UK

² Centre d'Ecologie Fonctionnelle et Evolutive, Univ Montpellier, CNRS, EPHE, IRD, Montpellier, France

³ Department of Biology, Norwegian University of Science and Technology (NTNU), N-7491 Trondheim, Norway.

⁴ Behavioural Ecology, Faculty of Biology, LMU Munich, Planegg-Martinsried, Germany

⁵ Department of Biological Sciences, North Dakota State University, Fargo, North Dakota, USA

⁶ Department of Biology, University of Ottawa, 30 Marie Curie, Ottawa, ON, K1N 6N5, Canada.

⁷ Département des Sciences Biologiques, Université du Québec A Montréal, CP8888 succursale centre-ville, Montréal Qc, H3C 3P8, Canada.

⁸ Institute of Ecology and Evolution, Friedrich Schiller University Jena, Jena, Germany

⁹ Gjørevoll Centre, Norwegian University of Science and Technology (NTNU), Trondheim, Norway

¹⁰ Escuela de Biología, Universidad de Costa Rica, San José, Costa Rica.

¹¹ Members of the SQuID working group

* Corresponding author: joel.l.pick@gmail.com

24 **ORCID:**

25 JLP: 0000-0002-6295-3742

26 BC: 0000-0002-2361-9821

27 PA: 0000-0002-3815-772X

28 AC: 0000-0002-0691-2647

29 NJD: 0000-0003-3320-0861

30 NAD: 0000-0002-8370-4614

31 JGAM: 0000-0001-7726-6809

32 DR: 0000-0002-0419-7125

33 JMR: 0000-0002-5007-7343

34 HS: 0000-0002-9124-2261

35 CT: 0000-0001-9458-709X

36 KS: 0009-0001-2953-7915

37 YGAA: 0000-0001-7844-0477

38

Abstract

Estimating additive genetic variance (V_A) is key to understanding evolutionary potential in wild populations. V_A is commonly estimated using relatedness data from pedigrees in so-called ‘animal models’. The accuracy of these estimates depends strongly on information in the pedigree. Typically metrics describing pedigree structure (such as the number of individuals, maternities, pedigree depth etc.) are provided alongside V_A estimates. However, what these metrics represent is not well understood. Here, we first explored the evolution of metric reporting. Then, using 53 wild pedigrees, we demonstrated that precision of heritability (h^2 ; ratio of V_A to total phenotypic variance) estimates was strongly predicted by pedigree size, but other metrics provided little additional information about precision. An seldom-used analytical approximation almost exactly reproduced simulated standard errors for h^2 estimates from pedigrees and genomic relatedness matrices, offering an alternative to simulation-based power analysis for simple models. Although the utility of commonly reported metrics is limited for precision, they provide important information about risk of bias from confounding effects, which is valuable for planning future studies and assessing published ones. We demonstrated this using the example of unmodelled maternal effects. Future work on biases in h^2 estimation should explicitly take pedigree structure into account.

Keywords: pedigree, relatedness matrix, precision, risk of bias, wild population, additive genetic variance, heritability, simulation

1 Introduction

Robustly estimating quantitative genetic parameters is essential to predict the evolutionary dynamics of populations and their ability to respond to environmental changes. Animal models are currently the most common method used to estimate additive genetic variance (V_A) and covariance in traits in wild populations, and so to derive estimates of heritability (h^2 ; the ratio of V_A to total phenotypic variance) and evolvability (CV_A ; mean standardised V_A), which are both used as measures of evolutionary potential. Animal models are a type of mixed model, that use information on relatedness between individuals, which historically has been estimated from pedigrees and, more recently, measured via genomic relatedness matrices (GRMs; which estimate the realised relatedness directly from genomic data) (Henderson, 1988; Kruuk, 2004; Perrier *et al.*, 2018). Animal models are commonly cited as providing two main advantages compared to more traditional methods of estimating h^2 , such as parent-offspring regression and sibling analysis; first, by using information from all relationships in a pedigree, they should increase precision. Second, the mixed model structure allows us to account for many confounding effects (Kruuk, 2004; Kruuk & Hadfield, 2007), thereby reducing bias.

As wild populations are very diverse, both in their biology and in how they are studied, the resulting structure of the pedigrees we create is hugely variable. Among other factors, pedigrees can be affected by the life-history of the population (e.g. number of offspring per mother), their mating system (e.g. resulting in varying numbers of full and half siblings), and patterns of dispersal (male or female biased) and immigration (affecting the number of founders) in the population, as well as how data has been collected. Studies of wild populations vary widely, meaning that pedigrees may be long and cover many generations, for example where a small isolated population is studied for a long time period (as in many large ungulate studies e.g. Pemberton *et al.*, 2022), or wide and shallow if a subsample of a large population is studied for a short period of time, or has high rates of dispersal (as in many studies of passerine birds, e.g. blue tits and great tits).

To summarise this diversity, metrics describing the characteristics of the pedigrees are

often presented alongside published estimates of V_A and h^2 (which we will denote as $\hat{V}_A$ and $\hat{h}^2$, respectively). However, it is not always obvious what this reporting aims to provide and how it can be used to assess the reliability of these estimates. Generally speaking, these metrics should communicate ‘information’ about the pedigree more than the simple size of the pedigree (Morrissey & Wilson, 2010). Specifically, we might want these metrics to relate to the precision of $\hat{h}^2$ and the risk of bias from unmodelled parameters.

Simultaneously, the diversity of pedigree structure means that determining the sample sizes needed to estimate h^2 well is often perceived as being difficult, and so power analysis is typically conducted via simulation (Quinn *et al.*, 2006; Morrissey *et al.*, 2007; Morrissey & Wilson, 2010). With complex models, and the use of Bayesian frameworks, this can be prohibitively time-consuming. Having other, less computationally expensive approaches to assessing power would therefore be very useful for study design. Studies of wild populations also typically contain a large variety of factors that can bias $\hat{h}^2$. Often these are not accounted for, either because they are unknown or unmeasurable. In addition to providing information on precision of $\hat{h}^2$, the structure of the pedigree will likely determine the risk of bias from such unmodelled confounding effects. For example, pedigree gaps are known to affect bias in $\hat{h}^2$ (Wolak & Reid, 2017), and pedigree structure can affect the degree of bias in $\hat{h}^2$ from unmodelled common environment effects Kruuk & Hadfield (2007) and maternal genetic effects Pick *et al.* (2025b).

Historically, reported pedigree metrics have been quite limited (e.g. the total number of records, the total number of maternal and paternal links, and total number of cohorts or maximum number of generations). Morrissey & Wilson (2010) suggested presenting additional metrics, and provided an R package to calculate them (pedantics, now available in an updated form as pedtricks; Martin *et al.*, 2026), with the aim to ‘help us to better understand the implications of variation in pedigree structure’. As well as an extended set of metrics describing the different relationships within the pedigree, these metrics also included a description of the relatedness (or A) matrix. However, the precise interpretation of many of these metrics has

not been formally articulated. In our experience, many researchers present them as a matter of habit, rather than because of the precise information they contain.

Another major contribution of Morrissey & Wilson (2010) was to demonstrate the importance of reporting metrics relating to pruned pedigrees, where uninformative individuals (those that have no phenotype and are related to less than 2 other individuals in the pedigree, i.e. they don't link individuals together) are removed. However, not all individuals in a pruned pedigree are phenotyped and the proportion of unphenotyped individuals will vary across studies. Consequently, metrics relating to a pruned pedigree may still not reflect the information available to estimate h^2 , as only the phenotyped individuals are informative.

With the advent of relatively cheap genomic sequencing, an increasing number of studies use genomic relatedness matrices (GRMs) to estimate quantitative genetic parameters. Following this increase, additional metrics have been developed in human genetics for estimating precision of $\hat{h}^2$ that refer to the relatedness matrix directly. Two metrics in particular have the potential to be very useful. First, Visscher *et al.* (2014) developed an approximation of the sampling variance of heritability based on the sample size and variance in relatedness alone, which is also available in an easy-to-use web app (GCTA-GREML power calculator; <https://shiny.cnsgenomics.com/gctaPower/>). However, it should be noted that this metric was specifically developed for GRMs generated from unrelated samples, common in human genetics, and so may not work well with relatedness matrices from wild pedigrees focussed on family units. Second, Visscher & Goddard (2015) derived an approximation of the sampling variance of $\hat{h}^2$ based on the eigenvalues of the relatedness matrix. This second approximation relaxes the assumption of an unrelated population and so may be more suitable to studies of wild population. Both methods are particularly appealing, as they can be applied to both relatedness matrices from pedigrees and GRMs from genomic data. To our knowledge, these metrics have remained largely restricted to the human genetics and animal breeding literature and, outside of the use of GRMs, evolutionary ecologists have not picked up on their utility.

Here, we first conduct a literature review to assess how commonly different pedigree metrics are reported in studies estimating h^2 using animal models in wild populations, and how this reporting has evolved over time. Second, using a collection of 53 published pedigrees, we systematically evaluate how these metrics relate to both precision and minimum detectable effect size of $\hat{h}^2$. We demonstrate that considering metrics relating to the phenotyped individuals is more important than considering all individuals in the full or pruned pedigree. Third, we illustrate how the same metrics that are useful to describe precision in a pedigree may be very different in a GRM, depending on how a population has been sampled. Finally, we discuss the potential of these metrics as indicators of the ‘risk of bias’ in $\hat{h}^2$ estimates from unmodelled confounding factors, using the example of bias caused by common environmental effects.

2 Methods

All simulations and analysis were carried out in R (version 4.4.0; [R Core Team, 2024](#)). In the reporting and description of our methods, we follow the MeRIT guidelines ([Nakagawa et al., 2023](#)).

2.1 Reporting of metrics

To get a sample of studies using animal models in the wild, we used a database of heritability estimates up to April 2020 ([Young & Postma, 2023b](#)), compiled by [Young & Postma \(2023a\)](#). From this database, we took all studies that reported heritability using animal models ($n=167$), which included studies from 1992 to 2020. There are some limitations to this database. First, it focuses on birds and mammals only. This reflects the well-described sampling bias in favour of vertebrates present across evolutionary ecology ([Rosenthal et al., 2017](#)), and has potential implications for the generality of our inferences ([Dochtermann et al., 2026](#)). Second, it updates the database compiled by [Postma \(2014\)](#) using only the papers that cite [Wilson et al. \(2010\)](#),

which will be a subset of all available papers using animal models. While this database is potentially biased toward papers likely to report metrics (due to shared authors between [Wilson et al. 2010](#) and [Morrissey & Wilson 2010](#)), it potentially represents a ‘best-case’ summary of metric reporting.

These papers were reviewed by JLP, BC, PA, AC, DR, HS, CT and . We recorded whether they reported any metrics about the pedigree, and if so, which metrics. We classified metrics into groups; number of records, maternities or paternities (i.e. the number of individuals with a given parent), metrics relating to types of siblings or grandparents, number of founders, sibship size, pedigree depth (mean and max), metrics relating to inbreeding, and metrics relating to the mean and distribution of relatedness (including proportion of links > 0.25 etc.). Given that a large focus of [Morrissey & Wilson \(2010\)](#) was to promote simulations to assess power (see also [Morrissey et al., 2007](#)) and the importance of pedigree pruning, we also noted if the studies mentioned pedigree pruning or the use of simulations. We note that just because pruning and simulations are not mentioned in the manuscript, it does not mean that they were not done. Finally, we recorded whether the studies provided an archived pedigree or (inverse) relatedness (or A) matrix.

As well as generating descriptive statistics, we looked qualitatively at whether the use of these metrics changed over time - we would expect that the reporting of more diverse metrics, pruning and simulation would increase after 2010 with the publication of [Morrissey & Wilson \(2010\)](#).

2.2 Collating pedigrees

We wanted to collate a sample of pedigrees that was broadly representative of those used to estimate h^2 in the wild. To this end, JLP downloaded pedigrees from all the studies in our initial search that had a pedigree available online. This provided us with 27 pedigrees ([Bonnet et al., 2017](#); [Bourret & Garant, 2016](#); [Class et al., 2019, 2020](#); [Edme et al., 2018](#); [Fisher, 2019](#); [Fletcher et al., 2015](#); [Fletcher, 2017](#); [Kimock et al., 2019](#); [Kvalnes et al., 2016a](#);

Malenfant *et al.*, 2022; Martin *et al.*, 2017; Nicolaus *et al.*, 2016; Petelle *et al.*, 2015; Pigeon *et al.*, 2022; Poissant *et al.*, 2016; Reed *et al.*, 2016; Sakaluk *et al.*, 2014; Stedman *et al.*, 2017; Tarka *et al.*, 2013; Teplitsky *et al.*, 2011; vonHoldt *et al.*, 2020; Wheelwright *et al.*, 2014; Zub *et al.*, 2012; Zedrosser, 2019). To increase the number of pedigrees in our sample, JLP also searched the term 'pedigree' on the dryad repository on 31/10/2024. These records were screened to determine which were associated with studies estimating genetic variation in the wild, and that had a pedigree in the repository. This provided an additional 26 pedigrees (Bauch *et al.*, 2021; Becker *et al.*, 2015; Bérénos *et al.*, 2014; Bolund *et al.*, 2013; Class & Brommer, 2020a; Firth *et al.*, 2015; Foley *et al.*, 2020; Fracasso *et al.*, 2021; Gauzere *et al.*, 2020; Germain *et al.*, 2017; Godoy *et al.*, 2022; Husby *et al.*, 2012; Kim *et al.*, 2013; Kvalnes *et al.*, 2016b; Mazé-Guilmo *et al.*, 2015; McFarlane *et al.*, 2015; Messel Nafstad *et al.*, 2023; Morrissey *et al.*, 2022; Munds *et al.*, 2022; Olsson *et al.*, 2023; Ramakers *et al.*, 2019; Sauve *et al.*, 2019; Sparks *et al.*, 2022; Vedder, 2021; Westneat *et al.*, 2023; Winney *et al.*, 2019), which were also mainly composed of bird and mammal pedigrees, with one fish and one reptile pedigree. Overall, we had 53 pedigrees from 34 species (see Figure S2 for a description of the variation in pedigree sizes across taxa). For the purposes of our analysis, we assume throughout that these pedigrees are error-free. In reality, this is unlikely to be the case. Whilst this may present a source of bias in itself, this is not one we can easily assess, and is not the purpose of our study. Here, we are simply interested in having a representative sample of pedigrees that are used to estimate h^2 in wild populations.

2.3 Calculation of pedigree metrics

For each of the 53 pedigrees, JLP calculated the metrics presented in Morrissey & Wilson (2010, Table 1), using the pedtricks package (version 0.5.0 Martin *et al.*, 2026). We then compared the relationship between these pedigree metrics. As many of these metrics are counts or proportions with very low means, and all have defined ranges excluding negative values, we present the metrics on the log scale. We added 0.5 to any count-based metrics that

215 included values of 0 (maternal half-sibs, paternal half-sibs, all grandparent metrics, non-zero
216 F and $F > 0.125$) prior to log transformation (Yamamura, 1999).

217 In addition to these metrics that are often presented in wild quantitative genetics studies,
218 we calculated the two additional metrics that provide analytical approximations of the sampling
219 variance of $\hat{h}^2$ and so give a direct measure of precision. Based on Visscher *et al.* (2014), we
220 calculated

$$var(\hat{h}^2) \approx \frac{2}{N^2 var(r)} \quad (1)$$

221 where N is the number of individuals in the pedigree and r is the pairwise relatedness. We
222 calculated the variance in relatedness based on all values in the lower triangle of the A matrix.
223 As above, we note that the derivation of this metric assumes that variation in relatedness is
224 small (e.g. when individuals are unrelated, as in large human databases; discussed further in
225 Visscher *et al.* 2014), and so may not accurately predict the sampling variance of heritability
226 with wild pedigrees, where data collection is focussed on family units and so relatedness is
227 much higher. Based on Visscher & Goddard (2015), we also calculated

$$var(\hat{h}^2) \approx 2/(a - b^2/N) \quad (2)$$

228 , where

$$a = \Sigma \frac{(\lambda - 1)^2}{(1 + h^2(\lambda - 1))^2} \quad (3)$$

$$b = \Sigma \frac{(\lambda - 1)}{(1 + h^2(\lambda - 1))} \quad (4)$$

229 , where h^2 is the heritability (V_A/V_P) and λ is the eigenvalues of the A matrix (a function to
230 calculate this from a pedigree is also provided in the pedtricks package version 0.5.0; Martin
231 *et al.*, 2026). This second analytical approximation relaxes the low relatedness assumption
232 and is therefore likely more suitable for application to wild populations. We present the square
233 root of both these measures (i.e. an approximation of the expected standard error) and refer
234 to these as $SE_{\hat{h}^2}^{(1)}$ and $SE_{\hat{h}^2}^{(2)}$, respectively. In both cases, A matrices were generated using

Table 1: Table of all pedigree metrics we considered, their abbreviations (abb.) used in the results sections, what type of summary they are and a brief description.

Metric	Abb.	Type	Description
Records	N	Count	Number of individuals (rows) in pedigree
Maternities		Count	Number of individuals with a known mother
Paternities		Count	Number of individuals with a known father
Full sibs	FS	Count	Number of full sibling pairs
Maternal sibs	MS	Count	Number of sibling pairs sharing the same mother
Maternal half-sibs	MHS	Count	Number of sibling pairs sharing the same mother but not the same father
Paternal sibs	PS	Count	Number of sibling pairs sharing the same father
Paternal half-sibs	PHS	Count	Number of sibling pairs sharing the same father but not the same mother
Maternal grandmothers	MGM	Count	Number of individuals with a known maternal grandmother
Maternal grandfathers	MGF	Count	Number of individuals with a known maternal grandfather
Paternal grandmothers	PGM	Count	Number of individuals with a known paternal grandmother
Paternal grandfathers	PGF	Count	Number of individuals with a known paternal grandfather
Maximum pedigree depth	max. PD	Count	Maximum number of parent/offspring links between a founder and any individual
Founders		Count	Number of individuals with no known parents
Mean maternal sibship size	mat. sib. size	Average	Average number of offspring per mother
Mean paternal sibship size	pat. sib. size	Average	Average number of offspring per father
Non-zero F		Count	Number of individuals with a non-zero inbreeding coefficient
$F > 0.125$		Count	Number of individuals with an inbreeding coefficient above 0.125
Mean pairwise relatedness	mean r	Average	Average pairwise relatedness among all individuals
Pairwise relatedness ≥ 0.125	$r \geq 0.125$	Proportion	Proportion of pairwise relationships with a relatedness above 0.125
Pairwise relatedness ≥ 0.25	$r \geq 0.25$	Proportion	Proportion of pairwise relationships with a relatedness above 0.25
Pairwise relatedness ≥ 0.5	$r \geq 0.5$	Proportion	Proportion of pairwise relationships with a relatedness above 0.5
Variance in pairwise relatedness	var r	Variance	Variance in pairwise relatedness among all individuals

235 the nadiv R package (version 2.18.0 [Wolak, 2012](#)).

236 2.4 Pedigree metrics and precision in h^2

237 To assess how well the metrics predicted precision, JLP estimated the expected standard
238 error for a heritability of 0.3 for each pedigree via simulation using the squidSim R package
239 (version 0.2.6 [Pick et al., 2025a](#)). We chose 0.3 as it reflects the mean $\hat{h}^2$ from various
240 syntheses of $\hat{h}^2$ across studies and traits (e.g. [Postma, 2014](#); [Young & Postma, 2023a](#)). We
241 also tested whether our results were sensitive to this choice by repeating the simulations with
242 a h^2 of 0.1 and 0.5, which are presented in the supplements ([S3.1](#)). In the simulations,
243 all individuals in the pedigree have a phenotype. Although this is unlikely in real datasets,
244 this is essentially the underlying assumption for the reported metrics to be meaningful. We
245 note that no studies in our sample of 167 distinguished between relationships in the (pruned)
246 pedigree, and relationships between phenotyped individuals (although this has been done
247 elsewhere [Reid et al., 2014](#); [Acker et al., 2023](#)). Our simulations also assumed that all the
248 founders/immigrants comprised a single unstructured population which might not be realistic,
249 although is typically assumed in animal models.

250 For each pedigree, JLP simulated 1000 datasets and then estimated $\hat{V}_A$ for each simulated
251 dataset with an animal model using gremlin (version 1.1.0; [Wolak, 2024](#)), and from this derived
252 $\hat{h}^2$ as $\frac{\hat{V}_A}{\hat{V}_A + \hat{V}_E}$, where $\hat{V}_E$ is the estimate of residual variance. The expected standard error was
253 then estimated as the standard deviation of the sampling distribution of the $\hat{h}^2$ for each
254 pedigree. We then compared how well each of the metrics predicted the expected standard
255 error using the adjusted R^2 value derived from a linear model (using the lm function in R).
256 These comparisons were also made on the log scale and, as above, we added 0.5 to any
257 metrics that had a pedigree with a value of 0, prior to log transformation ([Yamamura, 1999](#)).

2.5 Minimum detectable effect size

We additionally estimated the minimum detectable effect size (here $\hat{h}^2$) for each pedigree. This value represents the smallest effect size that can be found to be statistically significant at a given alpha level for a given pedigree, and approximates the effect size for which the statistical power is 50%. Simulations were performed as above, this time with a simulated h^2 of 0. The sampling distribution of the $\hat{h}^2$ from these datasets represents the null distribution. The minimum detectable effect for an alpha value of 0.05 was then calculated as the upper 95% quantile of the null distribution of $\hat{h}^2$ for a pedigree. As above, we compared how well the different pedigree metrics predicted the minimum detectable heritability, on the log scale.

2.6 Reporting metrics for phenotyped individuals

As stated in the sections above, we have been assuming that all individuals are phenotyped, which is unlikely in real datasets. We further explored this limitation by comparing the expected standard error of $\hat{h}^2$ for datasets with and without missing phenotypic information. To this end, we considered a scenario in which juvenile phenotypes are only collected on individuals born into a population, meaning that only individuals with a known mother have a phenotype and so all founders/immigrants do not have a phenotype (similar to [Kruuk & Hadfield \(2007\)](#) and [Pick *et al.* \(2025b\)](#)).

JLP created pruned pedigrees for these phenotyped individuals; these were very similar to the original pedigrees, but excluded any founders without offspring. As above, JLP then simulated 1000 datasets for each of these pruned pedigrees, with a h^2 of 0.3, with phenotypes for all individuals. These datasets were analysed both with the full dataset and then again with the reduced dataset where only those with mothers have a phenotype. Depending on the pedigree there was between a 2% and 82% reduction in sample size. We then compared how well the number of records, the number of phenotyped individuals, and the second analytical approximation of SE ($SE_{\hat{h}^2}^{(2)}$) calculated both on the A matrix from all individuals and on the

A matrix of just the phenotyped individuals predicted the sampling variance, estimated from the sampling distribution of $\hat{h}^2$ (as above).

2.7 Genomic Relatedness Matrices

Whereas a relatedness matrix from a pedigree represents the expected relatedness among individuals based on pedigree information alone, a GRM represents an estimate of the realised relatedness among all individuals, and provides more fine-scale variation in relatedness for any known given relationship, such as among siblings. GRMs are therefore being used both in systems that have pedigrees (or would be amenable to creating one), and in systems where generating any pedigree would be impossible. Consequently, the structure of the relatedness matrix may vary a lot, compared to the typical family-based pedigree study. For example, randomly sampling a small proportion of individuals in a large population may result in sampling few closely related individuals, similar to studies typically used in human genomics. Here, the size of the GRM (i.e. number of individuals) may no longer be very predictive of the precision of $\hat{h}^2$.

To explore the relationship between precision of $\hat{h}^2$ and GRM size, we used six published GRMs that have been used to estimate h^2 in a variety of life-history, behavioural and fitness-related traits. As this analysis was designed to be illustrative, we did not systematically search for GRMs, rather used ones from studies known to us. These GRMs were expected to vary considerably due to how populations were sampled, and we believe this sample represents a range of data structures that are likely to exist in GRM estimation in wild populations. First, we used three GRMs from a set of studies with population sampling similar to that where we would have a pedigree, where the goal is often to try to catch everybody; Soay sheep (*Ovis aries*, [Regan et al. 2020](#), data: [Regan et al. 2019](#)), eastern water dragons (*Intellagama lesueurii*, [Strickland et al. 2021b](#), data: [Strickland et al. 2021a](#)) and Koalas (*Phascolarctos cinereus*, [Cristescu et al. 2022](#), data: [Strickland et al. 2022](#)). Second, we used two GRMs from mark-recapture systems where we expect a lower proportion

of individuals to be caught; roe deer (*Capreolus capreolus*, Gervais *et al.* 2020a, data: Gervais *et al.* 2020b) and Tasmanian devil (*Sarcophilus harrisii*, Strickland *et al.* 2024, data: https://github.com/kashastrickland/tasdevils_adaptivepotential). Finally, we used a GRM from a population of Stickleback (*Gasterosteus aculeatus*, Strickland *et al.* 2023b, data: Strickland *et al.* 2023a), which was generated through random sampling of individuals in a massive population, where we expect the relatedness among individuals to be very low.

JLP estimated the sampling variance of $\hat{h}^2$ for these six GRMs, using simulations in the same way as above, with a simulated h^2 of 0.3, as well as the second analytical approximation of SE ($SE_{\hat{h}^2}^{(2)}$; Visscher & Goddard, 2015). GRMs are covariance matrices, and so we expect them to be symmetric and positive definite. Due to the way GRMs are calculated and stored, this is not always the case. None of the GRMs were positive definite and some were not symmetric (likely due to floating point errors), and so we used functions from the Matrix package (Version 1.7-3; Bates *et al.*, 2025) to approximate symmetric, positive definite matrices. Note that $SE_{\hat{h}^2}^{(2)}$ can be calculated without a positive definite matrix, which can make quite different approximations of SE (see Section S5), and so we recommend to correct GRMs before using these methods. We then compared both $SE_{\hat{h}^2}^{(2)}$ and the sample size with the expected standard error of $\hat{h}^2$ from the simulation for both the 53 pedigrees and the 6 GRMs. We expected that 1) this analytical approximation would be a good predictor of the expected standard error for both GRMs and pedigrees, as it captures the relatedness structure of the data, and 2) the sample size would not be a good predictor of the expected standard error for the GRMs, as they are much more variable in structure than pedigrees.

2.8 Pedigree metrics and bias in $\hat{h}^2$

The structure of the pedigree will likely determine the risk of bias in $\hat{h}^2$ from unmodelled confounding effects (Kruuk & Hadfield, 2007; Wolak & Reid, 2017; Pick *et al.*, 2025b). To properly assess which metrics predict the size of bias from different unmodelled effects, we would have to conduct a much more extensive study, and systematically vary the pedigree

structure using simulations to untangle which parts of the pedigree structure affect each type of bias in isolation (for example, see [Pick *et al.*, 2025b](#)). Here, we simply wanted to highlight that these metrics may be useful for assessing risk of bias. We therefore present a demonstrative example of how some of these metrics can be used to predict the bias in $\hat{h}^2$ caused by unmodelled maternal environment effects, a bias that is well known ([Kruuk & Hadfield, 2007](#)).

To this end, JLP simulated 500 datasets for each pedigree with $h_2=0.1$ and maternal environment variance of 0.2. A model estimating only V_A was then run on each dataset (i.e. we did not model the maternal effects). Bias in the $\hat{h}^2$ was calculated as $\frac{1}{n} \sum (\hat{\theta}_k - \theta)$ (where θ is the true value (i.e. 0.1), $\hat{\theta}_k$ is the model estimate of h^2 from the k th simulation, and n is the number of simulations). We then assessed which metrics predicted the bias, again with the metrics on the log scale.

3 Results

3.1 Reporting practices

A summary of reporting practices is shown in Figure 1. 146 of the 167 studies presented some pedigree metrics. Nearly all the studies that reported metrics presented the number of records, and most also presented maternities/paternities. A considerable number presented a pedigree depth metric (most commonly maximum pedigree depth). The reporting of other metrics was less common. As predicted, the frequency with which several metrics are presented has changed since the publication of [Morrissey & Wilson 2010](#); the diversity of metrics presented has increased, with increases in average distribution of relatedness, siblings and sibship size, grandparents, inbreeding and number of founders (Figure S1), but the proportion of publications presenting no metrics also increased (Figure S1).

Only a small proportion of papers have an accompanying pedigree or A matrix (Figure

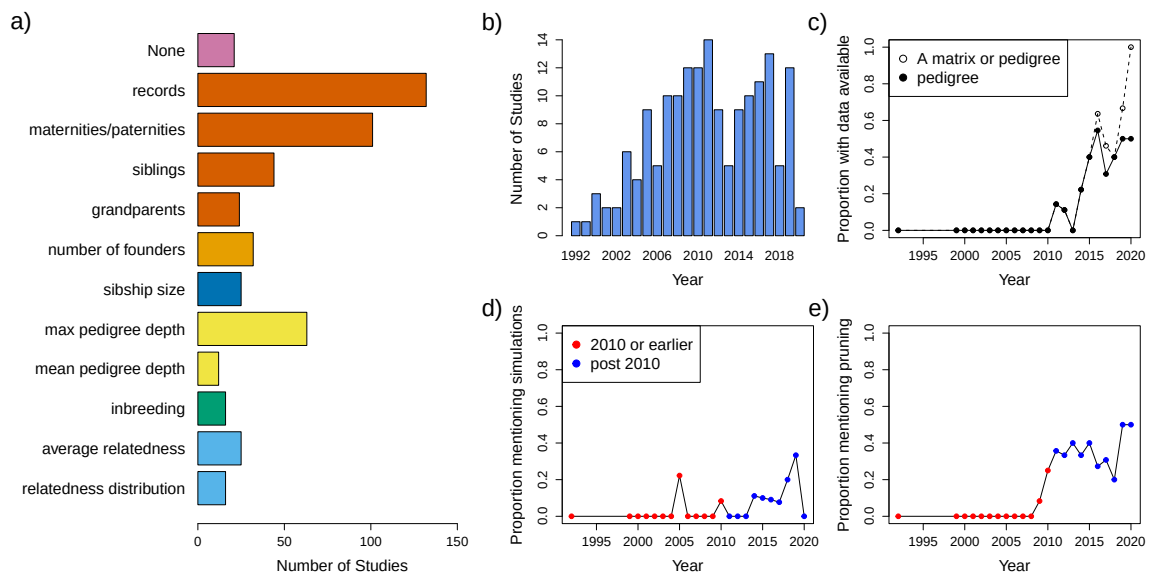


Figure 1: Pedigree metric reporting practices. a) shows the number of papers (out of a total of 167) that present each category of metric. b) The number of papers in each year. The proportion of papers providing a pedigree, or A matrix (c), reporting the use of pre-study simulations (d), and mentioning pedigree pruning (e).

1b,c). This has also changed over time, with a clear and sustained increase in the availability of pedigree data from 2010 (Figure 1c), matching the introduction of Joint Data Archiving Policy to many journals in ecology and evolution in 2010 (<https://doi.org/10.25504/FAIRsharing.z67ht2>). Few studies mention simulations, but this is clearly more prevalent after the publication of Morrissey & Wilson 2010 (Figure 1d). Discussion of pedigree pruning is more common, but still nowhere near ubiquitous. This again is much more prevalent after 2010, although the initial increase in reporting occurred prior to 2010 (Figure 1e).

3.2 Relationships between pedigree metrics

There was a more than 100-fold difference in pedigree size across our sample of 53 pedigrees, from 210 to 25,156 individuals. The count based relationships metrics (e.g. number of records, maternities, siblings etc) were all very highly positively correlated across the 53 pedigrees (Figure 2); in other words, the number of relationships of all kinds scales strongly with pedigree size (Figure 2). There was some finer scale clustering within these metrics; the close family relationship metrics (maternities, paternities, siblings) all clustered together, as

373 did the grandparents metrics.

374 There were generally lower correlations between the count based relationship metrics and
375 other metrics, indicating they provide some independent information from pedigree size. Again
376 groups of related metrics were all highly correlated within each other; the sibship size metrics,
377 the inbreeding metrics, and the A matrix based metrics (Figure 2). The relatedness metrics
378 were also negatively correlated to the pedigree size metrics. Such correlation is expected since
379 when pedigrees get bigger, for every single new close relationship, you add a disproportion-
380 ate number of 0s. The number of grandparents metrics were all highly correlated with the
381 inbreeding metrics and pedigree depth; this is also to be expected - inbreeding can only be
382 detected with grandparents (as otherwise the relatedness of parents is unknown), and similarly
383 the grandparents and the number of detected inbreeding events will increase with pedigree
384 depth.

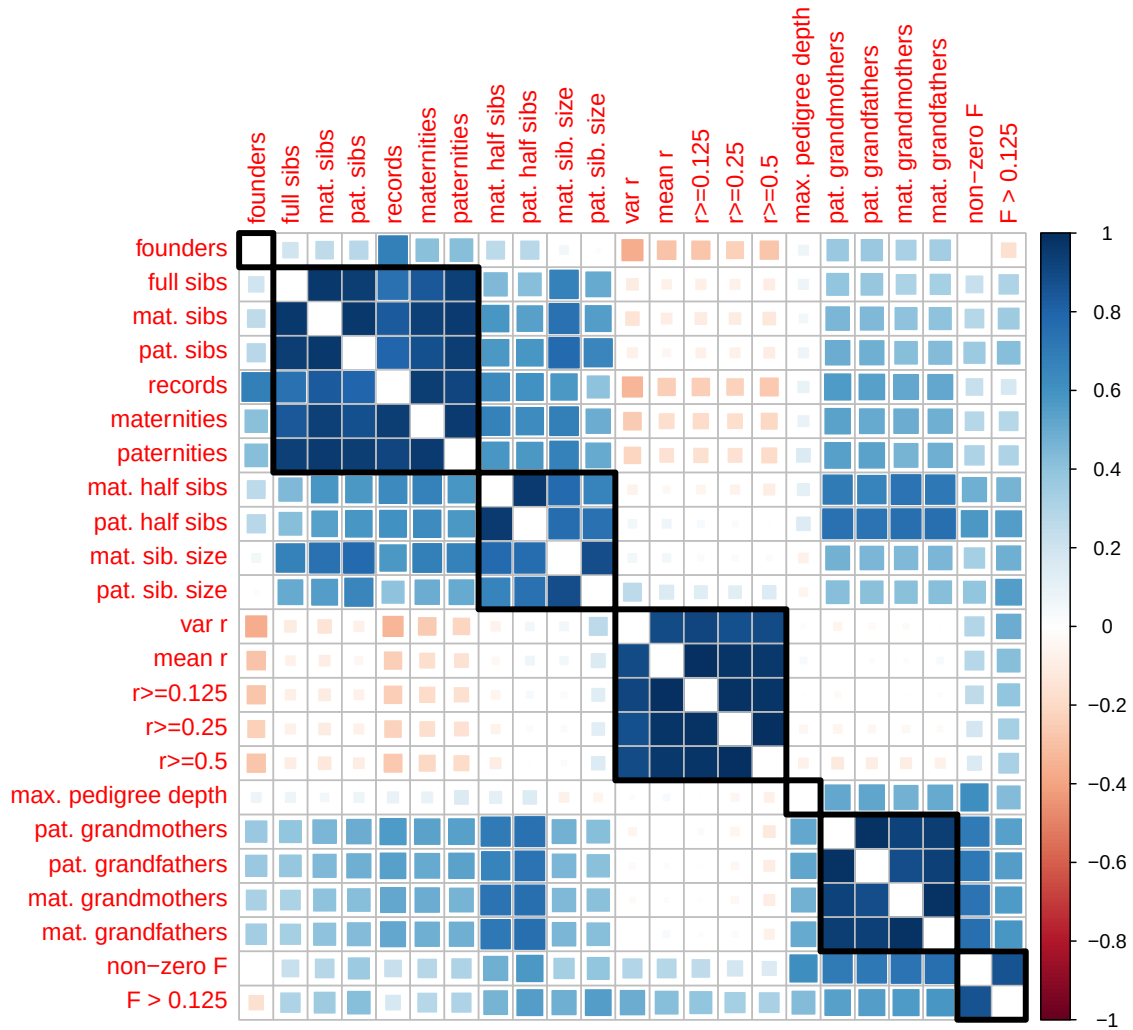


Figure 2: Correlations between pedigree metrics (See table 1 for abbreviations). Colour represents the strength and direction of the correlation, and the size of squares additionally reflects correlation strength. The boxes overlaying the correlation plot represent an illustrative clustering analysis using the R package corplot (Wei & Simko, 2021), which captures a similar grouping of metrics as shown in different colours in Figure 1a.

3.3 Predictive power for precision in $\hat{h}^2$

The sample size of the pedigree alone (e.g. total records) was a good predictor of precision in $\hat{h}^2$ (Figure 3a, S3). This might be expected to some extent, given the variation in pedigree size across pedigrees used in wild studies. The best predictor from the pedantics metrics was the number of maternities (Figures 3, S3). This metric essentially reflects the number of

individuals that are 'linked' into the pedigree. In this regard, the number of individuals with at least one parent might better capture this. We note that in this sample of pedigrees, the number of maternities is extremely highly correlated with the number of individuals with at least one parent ($r=0.998$), most likely because in many of these systems, if only one parent is known it is much more likely to be the mother, as most commonly observed systems have maternal care, meaning the mother is more likely to be identified. As there is an asymptotic relationship between maternities and expected SE of $\hat{h}^2$, when considering the number of maternities on an untransformed scale, we can see that there is only a small change in SE after around 2000 maternities (Figure 3b), compared to the rapid drop in precision for pedigree sizes below that size (although note that this will be very dependent on the underlying effect size, and the distribution of the data). These relationships are quantitatively similar when calculated with a h^2 of 0.1 and 0.5 (Figure S4).

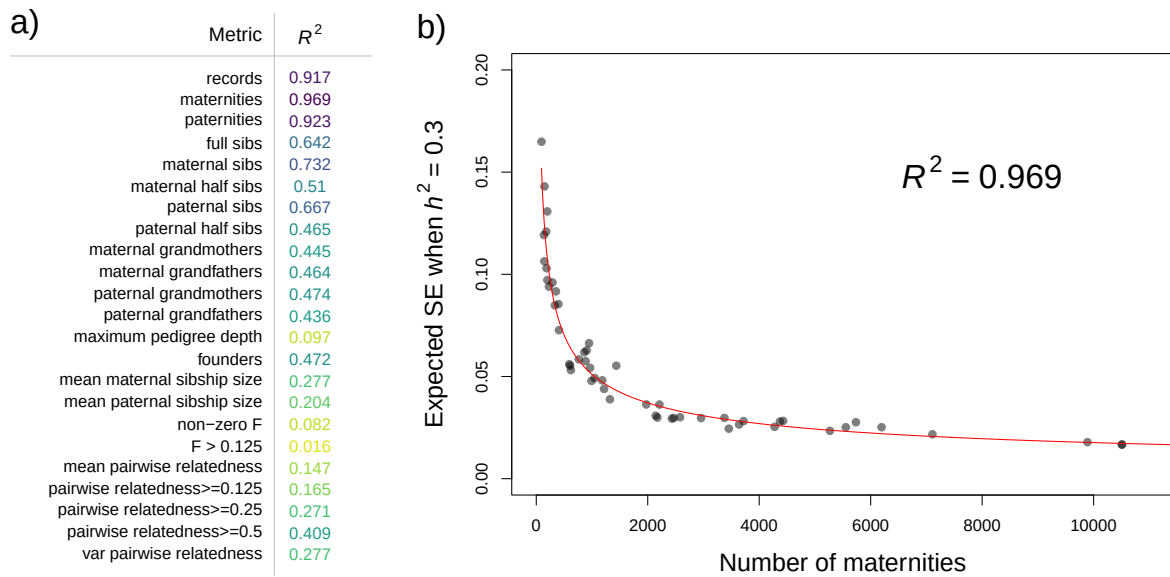


Figure 3: a) Strength of the relationship between the pedantics metrics and the expected standard error of $\hat{h}^2$ (based on simulations with an h^2 of 0.3), across 53 wild pedigrees. Adjusted R^2 values are coloured based on the strength of the relationship. b) Relationship between the expected standard error and the number of maternities. Red line shows the model prediction from a simple linear model. For clarity, the X axis is limited to 10,000, and so the figure does not show the pedigree with >20,000 maternities.

The second analytical approximation of the expected standard error of $\hat{h}^2$ ($SE_{\hat{h}^2}^{(2)}$, equation 2; Visscher & Goddard, 2015) was an almost perfect predictor of the precision in $\hat{h}^2$ in a simple

404 animal model (Figure 4b), regardless of the h^2 (Figure S5). It is therefore robust to the
 405 structure and low sample sizes of wild pedigrees. This approximation also has the advantage
 406 that it predicts the actual SE (see comparison with 1:1 line in Figure 4b). As expected, the
 407 first analytical approximation ($SE_{\hat{h}^2}^{(1)}$, equation 1; Visscher *et al.*, 2014) was a much poorer
 408 predictor of precision in $\hat{h}^2$, both in terms of R^2 but also in terms of the deviation from 1:1
 409 line (Figure 4a), consistently underestimating SE, especially at low sample sizes. Indeed the
 410 predictive power of this metric (measured as R^2) was lower than the number of maternities.
 411 This was consistent as h^2 changed (Figure S5) with predictions being slightly stronger at lower
 412 h^2 .

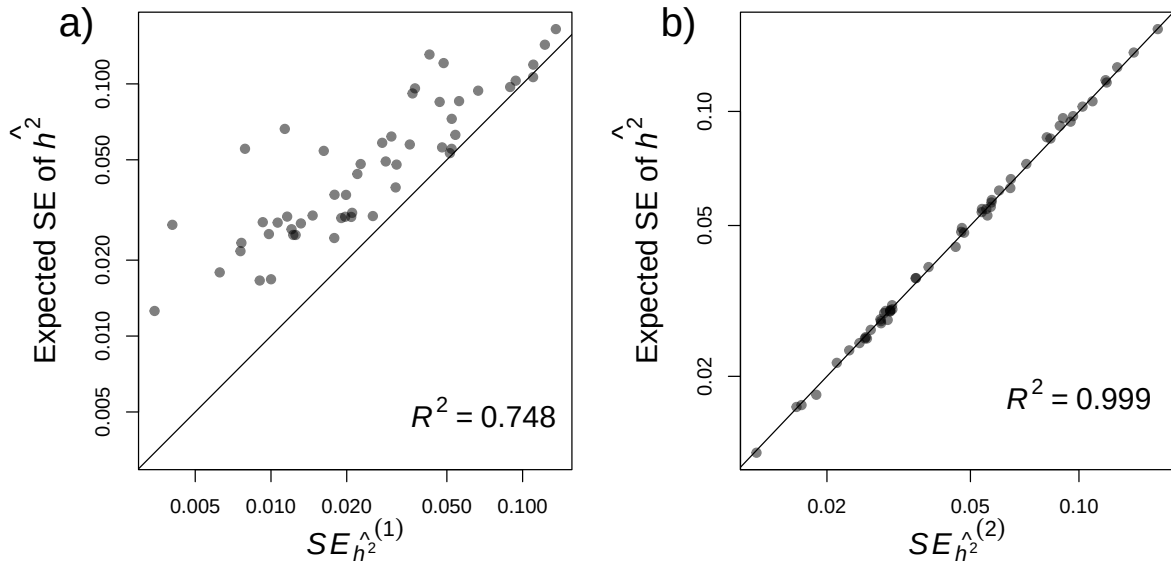


Figure 4: Relationship between the two analytical approximations of the SE of $\hat{h}^2$ and the expected standard error of $\hat{h}^2$ based on simulations, across 53 wild pedigrees. $SE_{\hat{h}^2}^{(2)}$ and the simulations were run with an h^2 of 0.3. The black line represents the 1:1 relationship.

413 3.3.1 Minimum detectable effect size

414 Similarly, the number of maternities was highly predictive of the minimum detectable $\hat{h}^2$
 415 (Figure 5). Interestingly, the number of records was far less predictive of the minimum
 416 detectable effect size than for precision.

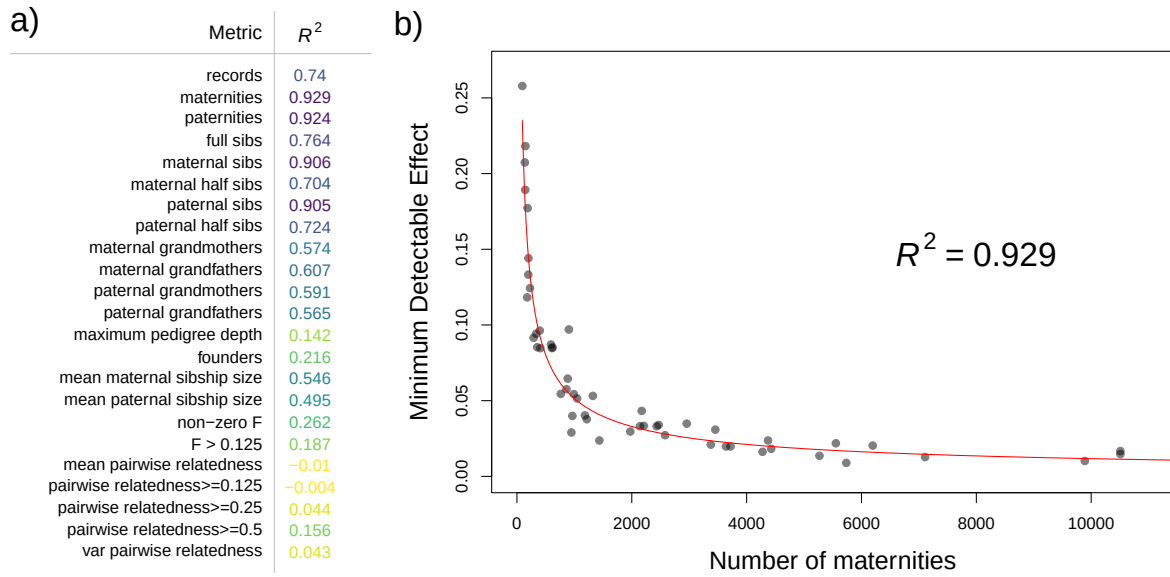


Figure 5: a) Strength of the relationship between the pedantics metrics and the minimum detectable effect size, across 53 wild pedigrees. Adjusted R^2 values are coloured based on the strength of the relationship. b) Relationship between the minimum detectable effect size and the number of maternities. Red line shows the model prediction from a simple linear model. For clarity, the X axis is limited to 10,000, and so the figure does not show the pedigree with >20,000 maternities.

3.3.2 Reporting metrics for phenotyped individuals

When we have phenotypic data for a subset of individuals in the pedigree but use metrics based on the whole pedigree (even after pruning), the metrics' ability to predict precision of $\hat{h}^2$ breaks down (blue points in Figure 6), particularly in smaller pedigrees. Using the number of phenotyped individuals (orange points), rather than the number of individuals in the pedigree (blue points), provides a similar relationship between the metric and precision to when the data is available for all individuals in the pedigree (black points). This demonstrates that when not all individuals in the pedigree have a phenotype, the metrics referring to the whole pedigree (even when pruned) will not always capture the expected precision of $\hat{h}^2$, especially when the pedigree is small. It also suggests that, for a typical wild pedigree, the number of phenotyped individuals alone provides a very crude measure of precision of $\hat{h}^2$. Note that in this scenario, all individuals with a phenotype also have a parent in the pedigree, so the number of phenotyped individuals is very close to the number of individuals with a parent. This would look very different if many individuals with a phenotype were founders, as

then this number would not link closely to the number of individuals that are linked into the pedigree/have a parent or close relative.

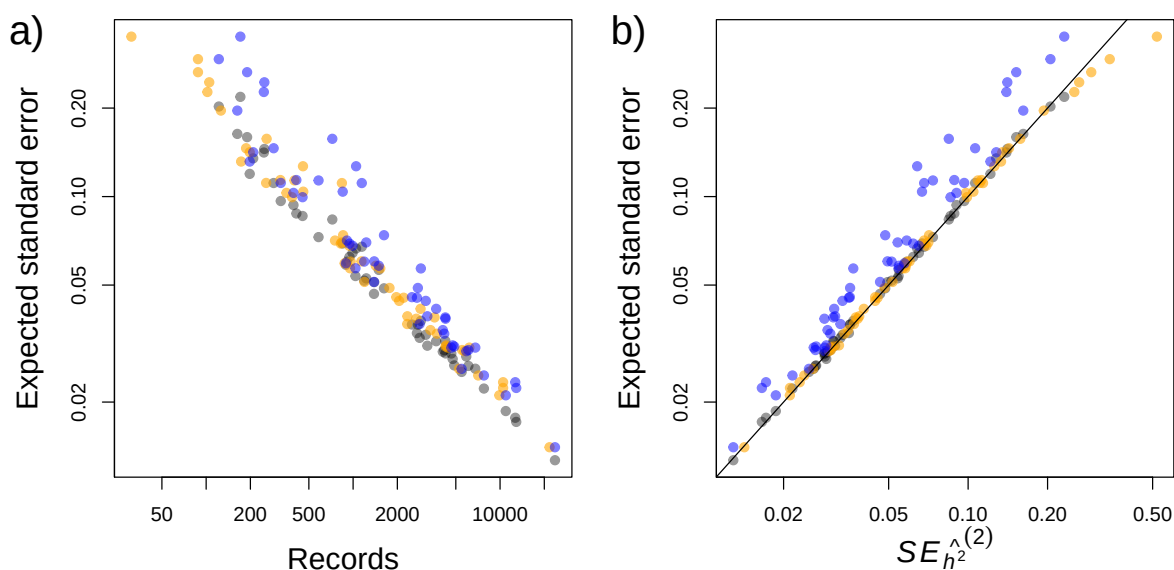


Figure 6: How well a) the number of individuals (records) and b) $SE_{h^2}^{(2)}$ predict the the expected SE of $\hat{h}^2$ based on two scenarios, 1) when all individuals in the pedigree have a phenotype (black points; similar to simulations in previous section), and 2) when only a subset of individuals have a phenotype, where expected SE of $\hat{h}^2$ was compared with both metrics for the full pedigree (blue points), and metrics from the phenotyped individuals only (orange points). $h^2 = 0.3$ in all the simulations. Black line in b) is the 1:1 line.

3.3.3 GRMs

When comparing pedigrees with GRMs, we can see that the number of records in a GRM is much less predictive of the expected standard error of $\hat{h}^2$ (Figure 7a), with much more variation in the standard error for a given size of GRM than pedigree. For example, taking only pedigrees/GRMs with between 400 and 600 records (five pedigrees, three GRMs), the expected SE of the pedigrees varied between 0.081 and 0.165 (twofold variation), whereas the GRMs varied between 0.058 and 0.344 (nearly sixfold variation). Generally, the GRMs have slightly lower SE, compared with pedigrees, which we might expect in systems with similar sampling to pedigree based systems, as the GRMs are dense matrices with information about the realised relatedness of all individuals, compared to the sparse pedigree based A matrix. This benefit is much greater in absolute terms at smaller sample sizes, and becomes

negligible at large sample sizes (note the axes in Figure 7a are on the log scale). However, the stickleback GRM (with very few related individuals) has a much higher standard error for its size, as does the Roe deer GRM from a mark-recapture system, although to a lesser extent. These disparities reflect the importance of sampling strategy on the precision in $\hat{h}^2$ from a GRM. Unsurprisingly, $SE_{\hat{h}^2}^{(2)}$ is also a very good predictor of SE with the GRMs, as with pedigrees (Figure 7b).

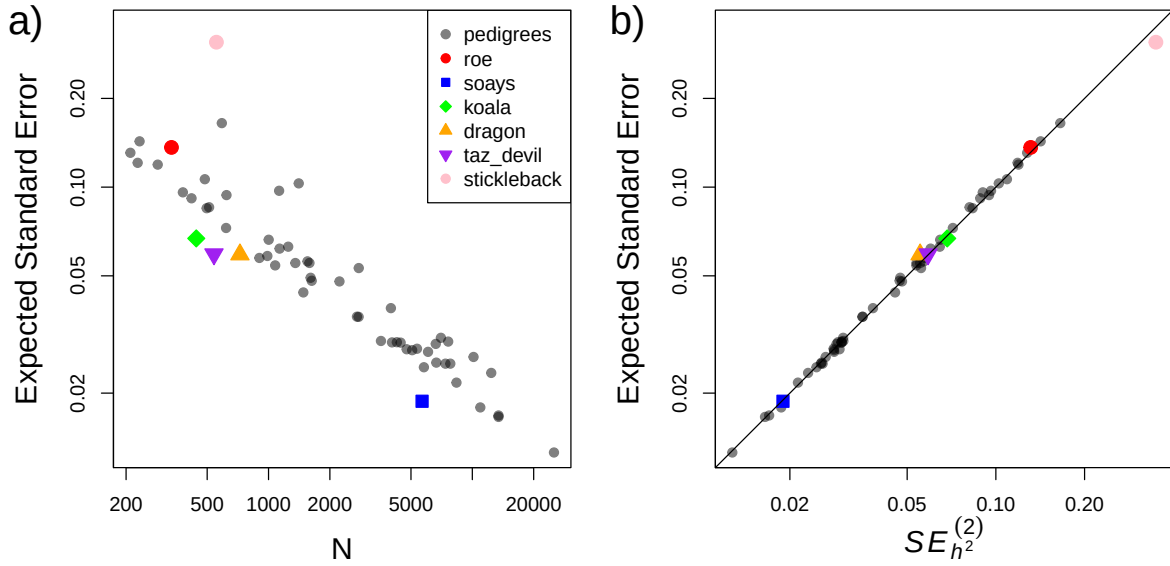


Figure 7: Relationship between the expected standard error of $\hat{h}^2$ based on simulations with a) number of records (N) and b) approximate standard error based on the [Visscher & Goddard 2015](#). Expected standard error and $SE_{\hat{h}^2}^{(2)}$ are based on a h^2 of 0.3. Values generated from pedigrees are in black and GRMs are in colour. Black line in b) is the 1:1 line.

3.4 Predictive power for bias in $\hat{h}^2$

The mean maternal (to a lesser extent paternal) sibship size was the metric most strongly related to the size of the bias (Figure 8). This is perhaps unsurprising because high maternal sibship size means that a large amount of information is coming from maternal sibling comparisons compared to parent-offspring comparisons, and the former are confounded with the maternal effects. Furthermore, in larger maternal sibships, more individuals share the same confounding maternal effects, than when sibships are smaller. More generally, this serves to demonstrate that these pedigree metrics can provide information about the risk of different

biases.

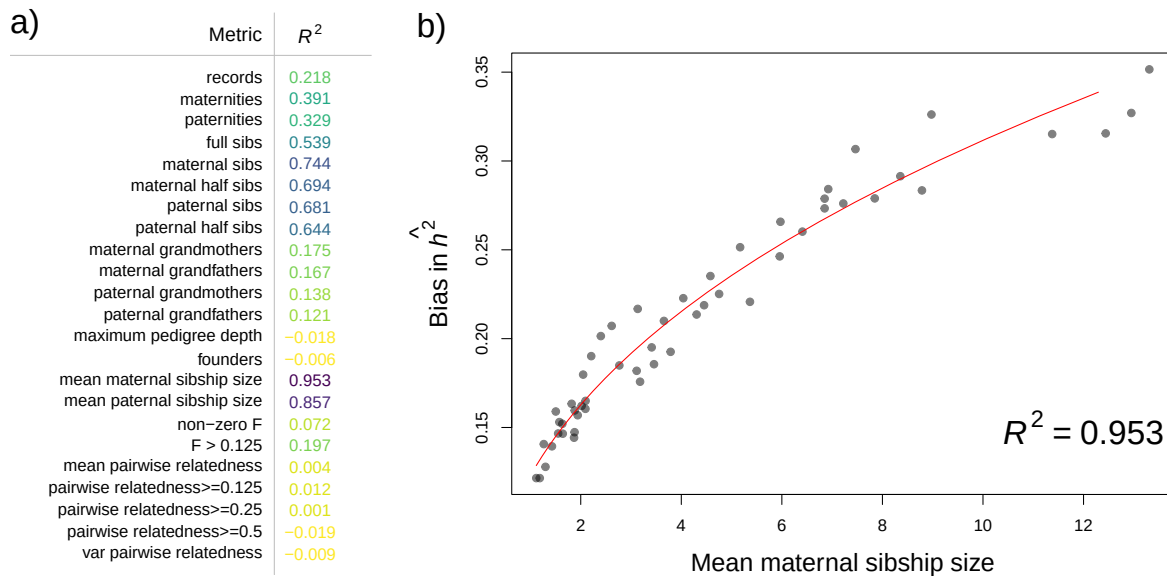


Figure 8: a) Strength of the relationship between the pedantics metrics and the bias in $\hat{h}^2$ (based on simulations with $h^2 = 0.1$ and $m^2 = 0.2$), across 53 wild pedigrees. R^2 values are coloured based on the strength of the relationship. b) Relationship between the expected standard error and the number of maternities. Red line shows the model prediction from a simple linear model.

4 Discussion

Here we show that, in typical pedigrees found in long-term studies of wild populations, the size of the pedigree is highly predictive of the precision of $\hat{h}^2$. Although we think of wild pedigrees as being highly variable, they are typically based on studies of close family units, meaning that there are large numbers of highly related individuals (parent-offspring and siblings; $r=0.25$ - 0.5), which are the most informative for estimating h^2 . It is likely that wild pedigrees vary most in the relative proportion of more distant relationship, which matter much less for estimating h^2 precisely. We also highlight a direct analytical approximation of the sampling variance of $\hat{h}^2$ ($SE_{\hat{h}^2}^{(2)}$) based on the relatedness (A) matrix, which has been seldom used in studies of wild populations. This analytical approximation almost perfectly predicts the sampling variance of $\hat{h}^2$, it can be used for both pedigrees and genomic relatedness matrices (GRMs), and remains accurate at low sample sizes. Whilst [Morrissey & Wilson \(2010\)](#) correctly point out that the

pruned pedigree is most relevant when presenting appropriate pedigree metrics, we further demonstrate that when not all individuals have a phenotype, the metrics are more relevant in the context of the individuals with phenotypes rather than the whole pedigree (even when pruned). Finally, we demonstrate that these pedigree metrics can be especially useful to assess the risk of bias in $\hat{h}^2$, as differences in pedigree structure likely reflect variation in key sources of confounding effects from unmodelled parameters.

4.1 Using pedigree metrics as an indicator of precision

Similar to a post-hoc power analysis, presenting metrics for the purpose of representing the expected precision of $\hat{h}^2$ from a given pedigree is largely redundant after the analysis of interest has been conducted, as the analysis already provides an estimate of the uncertainty (e.g. the SE of the estimate). This means that assessing the expected precision of $\hat{h}^2$ is mainly useful pre-analysis, as part of planning the study or analyses. It is important to note that the results we present here are based on the expected standard error for a h^2 of 0.3, from a model with a simple random effect structure, and Gaussian error distribution. Using more complex distributions will strongly affect the standard errors of variance components in mixed models (e.g. [Pick et al., 2023](#)), and so any rule of thumb taken here should not be directly extrapolated to more complex models.

There are several additional reasons these metrics may be useful post-hoc. First, when power is low, we are unlikely to find statistical support for small effect sizes. A metric related to power and precision is the minimum detectable effect size (also referred to as the critical effect size), which tells us what effect size would be detectable as statistically significant for a given pedigree. Such a metric can be useful post-hoc as it allows statistically “non-significant” results to be put into context of what could be detected with the pedigree at hand. Here we show that there is also a tight correspondence between the number of maternities and the minimum detectable $\hat{h}^2$, which may allow future studies to assess the minimum detectable effect without doing simulations, and equally allow us to assess this in published literature

where simulations are not an option. As with precision, the minimum detectable effect is also useful for planning studies, with the similar caveat that the specific patterns we present here are only relevant for the models we explored. Second, low power can also lead to systematic upward biases in variance estimation in mixed models (Pick *et al.*, 2023). This bias has a tight relationship with precision, and so these metrics may allow us to assess the risk of that bias in future work without doing simulations, and assess it in already published studies.

4.2 Presenting metrics as indicators of risk of bias

Perhaps most importantly, as well as telling us about precision, these metrics will likely tell us about the risk of bias in $\hat{h}^2$ from different confounding sources. Estimating h^2 in wild populations is difficult, as there are many confounding factors. Often these are not accounted for, either because they are unknown or unmeasurable. Often, different types of individuals encounter different confounding effects, and so the structure of a pedigree may tell us about the risk of bias. Here, we show that the mean maternal sibship size is a strong predictor of the bias from unmodelled maternal effects. This common environment is often (although not ubiquitously) accounted for and so may not be so much of a problem. However, there are many other sources of confounding effects that may not be commonly accounted for, and the other pedigree metrics typically reported may be descriptive of these biases. We therefore, advocate for continuing to publish these metrics alongside estimates of V_A and h^2 (for example as a supplementary table), as they may be useful for future synthesis work, especially when the respective pedigrees have not been archived (see next section). Future work focussing on how these metrics relate to different biases would be extremely useful.

A handful of other studies have also shown metrics that describe biases. For example, Kruuk & Hadfield (2007) showed that bias in $\hat{h}^2$ from unmodelled common environment effects is affected by both dispersal and pedigree depth in two pedigree structures (passerine bird and ungulate), but note that pedigree depth and pedigree size were confounded in this study. The impact of pedigree depth found may therefore relate to the maternal sibship size

effect we show here; with overlapping generations, as the number of years monitored increased (pedigree depth), the maternal sibship size will also increase (as females will have offspring over multiple years). [Pick et al. \(2025b\)](#) also investigated biases from unmodelled maternal *genetic* effects and found that the proportion of non-sibling maternal links (individual related via their mothers that are not siblings, e.g. cousins whose mothers are sisters) predicts the risk of bias from unmodelled maternal genetic effects, as the covariance between these individual is in part due to maternal genetic variation. In both cases, the risk of bias can be assessed by knowing the relative number of individuals who share confounding effects.

Dominance is another obvious source of confounding variation that may be predicted by pedigree structure. Dominance variance contributes to the covariance between both full siblings and double first cousin (e.g. [Lynch & Walsh, 1998](#); [Wolak & Keller, 2014](#)) (the latter likely being rare in natural populations)/ Consequently, not modelling dominance variance can inflate $\hat{h}^2$ ([Class & Brommer, 2020b](#)), although we are not aware of simulation studies that explore how the bias is affected by pedigree structure, but presumably the proportion of full siblings compared to other informative relationships that do not share dominance variance will be key to predicting risk of bias.

4.3 Reporting practices

Our data show that pedigrees have only begun to be made available in recent years, and there is still a high proportion of quantitative genetic studies that do not archive a pedigree alongside their study. The importance of these pedigrees as a resource cannot be understated. Not only do they allow for data re-use, full analytical transparency and computational reproducibility ([Pick et al., 2026](#)), but it may be that we do not currently report a metric that we will find to be useful with future work. For example, [Pick et al. \(2025b\)](#) show that the proportion of non-sibling maternal links predicts the risk of bias in V_A estimates from unmodelled maternal genetic effects, but this is not a metric that we commonly report. With a pedigree available, this metric can be calculated, and the risk of bias assessed. We therefore advocate that a

pruned pedigree should be archived alongside a given study, as this will allow biases to be fully assessed. Currently only a minority of studies do this (Figure 1c), despite many journals in ecology and evolution mandating data deposition. We understand maintainers of long-term studies may have reservations about this (e.g. Mills *et al.*, 2015), but note that a pruned pedigree can be used for little else than reproducing the study in question (for example, it cannot be used to calculate lifetime reproductive success because not all individuals are contained within it, only those that provide informative links). We also note that there is an increasing trend in archiving (inverse) A matrices derived from pedigrees (Figure 1c). Although these allow for computational reproducibility, they cannot be used to calculate any relationship-based metrics, limiting their future utility.

4.4 Genomic Relatedness Matrices

Given the increase in the availability of genomic data from many wild populations, there has been an increasing use of GRMs for quantitative genetic studies. This approach, in theory, allows researchers to go to any population, record phenotypes and take genetic samples, and then estimate h^2 with no further information (Frentiu *et al.*, 2008). In systems where a high proportion of individuals of known relatedness are phenotyped, GRMs can also provide gains in precision for the same number of phenotype individuals, due to the high resolution of relatedness information, although in absolute terms this gain reduces quickly as pedigree size increases. GRMs can also be useful in systems where pedigrees are available, but there is high pedigree error due to extra-pair paternities. Here, we show the analytical approximation of Visscher & Goddard (2015, $SE_{h^2}^{(2)}$) can be effectively used to predict precision from any relatedness matrix; both those derived from pedigrees and GRMs. However, the strong relationship between the number of phenotyped individuals and sampling variance does not necessarily hold for systems where genotyped individuals may represent a small, random sample of a population. Therefore, although we may be able to make rule of thumb suggestions about sample size in pedigree based studies, this will be much more problematic in studies

based on GRMs made from less structured population sampling (as in the case with the stickleback GRM presented here; Figure 7), without prior knowledge of the resulting relatedness structure. When planning a GRM based study, it is therefore important to consider sampling design, and to try and maximise sets of related individuals (for example by capturing individuals at close proximity of each other).

Whilst GRMs offer a lot of promise for being able to estimate h^2 in the wild without needing long-term individual based studies, we should take caution with a simple GRM approach. Without a pedigree and detailed population monitoring, we do not know the specific relationships among individuals, which allows to either correct for many common confounders (e.g. common environment effects), or infer the risk of bias. Such studies, therefore, come with a clear caveat that $\hat{h}^2$ may well be inflated, and we cannot assess the degree to which this might occur. Methods have been developed in human genetic to deal with this; for example, additionally fitting a restricted GRM capturing closer family relationships to remove confounders such as maternal effects and dominance, with relationship below a certain threshold set to 0 (Zaitlen *et al.*, 2013). To our knowledge, this method has currently only had limited application in wild populations (James *et al.*, 2024; Strickland *et al.*, 2024). There are several important caveats to this methods. First, it appears to be sensitive to the particular threshold used (James *et al.*, 2024; Strickland *et al.*, 2024). Second, its applicability at small sample sizes is not well known; James *et al.* (2024) used 8500 individuals and human studies use many more. Finally, this method removes part of the heritable variation that may be of direct interest to evolutionary biologists; for example, whilst $\hat{h}^2$ may be unbiased, this method removes maternal *genetic* effects, which are important to quantify evolutionary potential. We note that we might expect less biases in $\hat{h}^2$ from shared environments in scenarios where a large population is being randomly sampled (such as the stickleback population here).

4.5 Use of these metrics outside of estimating h^2

Our focus here has been entirely on the estimation of h^2 . Although this is perhaps the most commonly estimated quantitative genetic parameter using pedigree information, many other parameters can also be estimated using such information. The pedigree metrics considered here will undoubtedly provide information about the estimation of these other parameters. Perhaps the most obvious additional parameter is dominance variance, although the metrics that best predict this would need to be identified. Whilst this has not been explicitly tested, it seems a safe assumption that the analytical approximation provided by [Visscher & Goddard \(2015\)](#), here referred to as $SE_{\hat{h}^2}^{(2)}$ would also work as an approximation of the SE for estimates of dominance variance, as this modelled in the same way, using a dominance matrix instead of a relatedness matrix as a covariance structure for a set of random effects. Indeed, this metric should work to provide an approximation of the SE for any variance estimates derived from a covariance matrix included in a similar way in a mixed model (e.g. accounting for spatial autocorrelation, home-range overlap, social similarity [Thomson et al., 2018](#)).

Similarly, these metrics may also provide information about the precision of genetic correlation estimates from multi-response animal models. Whilst these parameters are estimated with much more uncertainty, and their precision additionally depends on the heritability of both traits, we would expect that metrics presented here would scale in a similar way with the precision of genetic correlation estimates as with the SE of $\hat{h}^2$. In other words, the precision of genetic correlation estimates likely scales closely with the number of individuals with a known parent in the pedigree. This is because, as with the precision of h^2 estimates, their estimation comes down to how well individual breeding values can be estimated, and the information available to do this. This may be further complicated by not all individuals having phenotypes for both traits, with between sex covariances being the extreme example in which no individual expresses both phenotypes. [Visscher et al. \(2014\)](#) provide an approximation of the sampling variance of genetic correlation estimates based on the variance in relatedness, sample size and heritabilities of the two traits. As with their approximation of the sampling variance of $\hat{h}^2$

626 $(SE_{\hat{h}^2}^{(1)})$, this is designed for GRMs generated from unrelated samples, and is less accurate at
627 the small sample sizes we expect in natural populations (Visscher *et al.*, 2014). No extension
628 of second approximation of sampling variance of $\hat{h}^2$ ($SE_{\hat{h}^2}^{(2)}$; Visscher & Goddard, 2015) to
629 genetic correlations has yet been made. It would be interesting for future work to explore how
630 all these metrics may relate to precision in these more complex models.

Box 1: A Roadmap for Pedigree reporting

Planning a Study

- Ideally, perform simulations with your pedigree to assess precision and risk of bias in $\hat{h}^2$.
- If simulation is not an option:
 - The analytical approximation of [Visscher & Goddard \(2015\)](#) provides a good estimate of the expected SE of $\hat{h}^2$ for simple models. With more complex models and distributions the SE will be larger, but this will give a lower bound.
 - The number of individuals with a known parent will also give a reasonable approximation of expected SE of $\hat{h}^2$.
 - Metrics associated with known biases can be used to identify the risk of bias when certain confounding factors are not modelled.
- Metrics may give information about other parameters, such as dominance variance, but more work is needed. The analytical approximation of [Visscher & Goddard \(2015\)](#) likely predicts the expected SE of variance estimates associated with any covariance structure.
- Compute metrics and analytical approximations on the phenotyped individuals, not the whole (pruned) pedigree.

Good reporting practices

- Archive the pruned pedigree alongside the data.
 - This allows any new metric to be calculated post-hoc.
 - It allows for simulations to assess risk of bias or minimum detectable effect size.
- Relatedness matrices allow the models to be re-run, but do not provide the information about *how* individuals are related that is needed to assess risk of bias. Pruned pedigrees should therefore be archived over relatedness matrices.
- Pedigree metrics should be reported alongside the study (for example in the Supplements), as they provide an indication of risk of bias from unmodelled confounders.

Assessing previously published papers

- Metrics can tell you about risk of bias from unmodelled confounding effects, and may allow you to assess the minimum detectable effect size.

Future simulations studies

- Future studies examining biases in $\hat{h}^2$ (and other parameters) should take pedigree structure into account, and identify metrics that describe risk of bias.
- Future studies could explore how well these metrics and approximations predict sampling variance in more complex models and for more parameters, e.g., dominance variance, genetic correlations etc.

4.6 Conclusions

Due to the way that we typically collect data in wild populations, the precision of $\hat{h}^2$ is very highly correlated with the number of maternities (which, in these systems, strongly represents the number of individuals with at least one parent). We further highlight an analytical approximation of the sampling variance of $\hat{h}^2$, that works extremely well with wild pedigrees and GRMs, and may provide an alternative to simulation based power analysis for simple animal models. Other commonly reported pedigree metrics provide little additional information about the precision of $\hat{h}^2$. However, we recommend that researchers continue to present detailed metrics as they can likely tell us about risk of bias for $\hat{h}^2$. Generally, more work needs to be done to link bias in $\hat{h}^2$ to pedigree structure. Archiving pruned pedigree alongside every study will allow us maximum flexibility in the future to assess risk of bias. Although GRMs offer a powerful way of estimating quantitative genetic parameters in wild populations, care needs to be taken when additional information about the potential confounding factors is missing.

Acknowledgments

This is a contribution of the Statistical Quantification of Individual Differences (SQulD) working group, and we'd like to thank the other members of SQulD for valuable feedback on the ideas presented here. Work on this project at SQulD workshops in 2022-4 was funded by a Research Council of Norway INTPART grant (project number 309356). JLP and KS were funded by a Royal Society Professorship (RSRP-R1-211017) and ERC grant (101020503 — EVOWILD — ERC-2020-ADG) to Loeske Kruuk. NAD was supported by the US National Science Foundation (IOS-2222929).

Conflict of Interest statement

The authors declare no conflict of interest.

Author Contributions

Conceptualization - JLP, BC and YAA Methodology - JLP, BC and YAA Software - JLP and JGAM Data curation - JLP Formal analysis - JLP Investigation - JLP, YGAA, BC, PA, AC, DR, HS and CT Visualization - JLP Writing – original draft - JLP and YGAA Writing – review & editing - All authors Supervision - YGAA, DR, NJD

Data and code availability

All data and code are this project are available at https://github.com/squidgroup/pedigree_metrics

References

- Acker, P., Daunt, F., Wanless, S., Burthe, S.J., Newell, M.A., Harris, M.P., Swann, R.L., Gunn, C., Morley, T.I. & Reid, J.M. (2023) Additive genetic and environmental variation interact to shape the dynamics of seasonal migration in a wild bird population. *Evolution*, **77**, 2128–2143. <https://dx.doi.org/10.1093/evolut/qpad111>.
- Bates, D., Maechler, M. & Jagan, M. (2025) *Matrix: Sparse and Dense Matrix Classes and Methods*. R package version 1.7-3.
- Bauch, C., Boonekamp, J.J., Korsten, P., Mulder, E. & Verhulst, S. (2021) Data from: High heritability of telomere length and low heritability of telomere shortening in wild birds. Artwork Size: bytes Pages: bytes.

674 Becker, P.J.J., Reichert, S., Zahn, S., Hegelbach, J., Massemin, S., Keller, L.F., Postma,
675 E. & Criscuolo, F. (2015) Data from: Mother-offspring and nest mate resemblance but
676 no heritability in early-life telomere length in white-throated dippers. Artwork Size: 44180
677 bytes Pages: 44180 bytes.

678 Bolund, E., Bouwhuis, S., Pettay, J.E. & Lummaa, V. (2013) Data from: Divergent selection
679 on, but no genetic conflict over, female and male timing and rate of reproduction in a
680 human population. Artwork Size: 2507068 bytes Pages: 2507068 bytes.

681 Bonnet, T., Wandeler, P., Camenisch, G. & Postma, E. (2017) Bigger Is Fitter? Quantitative
682 Genetic Decomposition of Selection Reveals an Adaptive Evolutionary Decline of Body Mass
683 in a Wild Rodent Population.

684 Bourret, A. & Garant, D. (2016) Data from: An assessment of the reliability of quantita-
685 tive genetics estimates in study systems with high rate of extra-pair reproduction and low
686 recruitment. Artwork Size: 22798047 bytes Pages: 22798047 bytes.

687 Bérénos, C., Ellis, P.A., Pilkington, J.G. & Pemberton, J.M. (2014) Data from: Estimating
688 quantitative genetic parameters in wild populations: a comparison of pedigree and genomic
689 approaches. Artwork Size: 692816901 bytes Pages: 692816901 bytes.

690 Class, B. & Brommer, J. (2020a) Data from: Can dominance genetic variance be ignored
691 in evolutionary quantitative genetic analyses of wild populations? Artwork Size: 1192320
692 bytes Pages: 1192320 bytes.

693 Class, B. & Brommer, J.E. (2020b) Can dominance genetic variance be ignored in evolu-
694 tionary quantitative genetic analyses of wild populations? *Evolution*, **74**, 1540–1550.
695 <https://dx.doi.org/10.1111/evo.14034>.

696 Class, B., Brommer, J.E. & van Oers, K. (2020) Data from: Exploratory behavior undergoes
697 genotype-age interactions in a wild bird. Artwork Size: 1083561 bytes Pages: 1083561
698 bytes.

699 Class, B., Klueen, E. & Brommer, J.E. (2019) Data from: Tail color signals performance in
700 blue tit nestlings. Artwork Size: 2445165 bytes Pages: 2445165 bytes.

701 Cristescu, R.H., Strickland, K., Schultz, A.J., Kruuk, L.E.B., de Villiers, D. & Frère,
702 C.H. (2022) Susceptibility to a sexually transmitted disease in a wild koala popula-
703 tion shows heritable genetic variance but no inbreeding depression. *Molecular Ecol-*
704 *ogy*, **31**, 5455–5467. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/mec.16676>,
705 <https://dx.doi.org/10.1111/mec.16676>.

706 Dochtermann, N.A., Sekhar, M.A. & Nakagawa, S. (2026) Ungeneralizable generalizations?
707 A meta-meta-analysis of the influence of taxonomic bias on the study of behaviour. *Animal*
708 *Behaviour*, **236**, 123542. <https://dx.doi.org/10.1016/j.anbehav.2026.123542>.

709 Edme, A., Zobač, P., Korsten, P., Albrecht, T., Schmoll, T. & Krist, M. (2018) Data from:
710 Moderate heritability and low evolvability of sperm morphology in a species with high risk
711 of sperm competition, the collared flycatcher *Ficedula albicollis*. Artwork Size: 1077386
712 bytes Pages: 1077386 bytes.

713 Firth, J.A., Hadfield, J.D., Santure, A.W., Slate, J. & Sheldon, B.C. (2015) Data from:
714 The influence of nonrandom extra-pair paternity on heritability estimates derived from wild
715 pedigrees. Artwork Size: 53956 bytes Pages: 53956 bytes.

716 Fisher, D. (2019) Data and ASREML code for "Social effects of territorial neighbours on the
717 timing of spring breeding in North American red squirrels".

718 Fletcher, Q.E., Speakman, J.R., Boutin, S., Lane, J.E., McAdam, A.G., Gorrell, J.C., Coltman,
719 D.W. & Humphries, M.M. (2015) Data from: Daily energy expenditure during lactation is
720 strongly selected in a free-living mammal. Artwork Size: 75413 bytes Pages: 75413 bytes.

721 Fletcher, R. (2017) Rapid Morphological Change of a Top Predator with the Invasion of a
722 Novel Prey.

723 Foley, N., Petit, E., Brazier, T., Finarelli, J., Hughes, G., Touzalin, F., Puechmaille, S. &
724 Teeling, E. (2020) Drivers of longitudinal telomere dynamics in the long-lived bat species,
725 *Myotis myotis*. Artwork Size: 143787 bytes Pages: 143787 bytes.

726 Fracasso, G., Matthysen, E. & Heylen, D. (2021) Heritable variation in host quality as mea-
727 sured through an ectoparasite's performance. Artwork Size: 286658 bytes Pages: 286658
728 bytes.

729 Frentiu, F.D., Clegg, S.M., Chittock, J., Burke, T., Blows, M.W. & Owens, I.P.
730 (2008) Pedigree-free animal models: the relatedness matrix reloaded. *Proceedings of*
731 *the Royal Society B: Biological Sciences*, **275**, 639–647. Publisher: Royal Society,
732 <https://dx.doi.org/10.1098/rspb.2007.1032>.

733 Gauzere, J., Walling, C.A., Pemberton, J.M., Kruuk, L.E.B., Morris, S. & Morris, A. (2020)
734 Data from: The genetic architecture of maternal effects across ontogeny in the red deer.
735 Artwork Size: 893709 bytes Pages: 893709 bytes.

736 Germain, R.R., Wolak, M.E., Arcese, P., Losdat, S. & Reid, J.M. (2017) Data from: Direct
737 and indirect genetic and fine-scale location effects on breeding date in song sparrows.
738 Artwork Size: 33344194 bytes Pages: 33344194 bytes.

739 Gervais, L., Hewison, A.J.M., Morellet, N., Bernard, M., Merlet, J., Cargnelutti, B., Chaval,
740 Y., Pujol, B. & Quéméré, E. (2020a) Pedigree-free quantitative genetic approach provides
741 evidence for heritability of movement tactics in wild roe deer. *Journal of Evolutionary*
742 *Biology*, **33**, 595–607. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/jeb.13594>,
743 <https://dx.doi.org/10.1111/jeb.13594>.

744 Gervais, L., Hewison, A.M., Morellet, N., Bernard, M., Merlet, J., Cargnelutti, B., Chaval,
745 Y., Pujol, B. & QUEMERE, E. (2020b) Data from: Pedigree-free quantitative genetic
746 approach provides evidence for heritability of movement tactics in wild roe deer. Artwork
747 Size: 1075723 bytes Pages: 1075723 bytes.

748 Godoy, I., Korsten, P. & Perry, S. (2022) Data from: Genetic, maternal, and environmental
 749 influences on sociality in a pedigreed primate population. Artwork Size: 1440046 bytes
 750 Pages: 1440046 bytes.

751 Henderson, C.R. (1988) Theoretical basis and computational methods for a number of different
 752 animal models. *Journal of Dairy Science*, **71**, 1–16.

753 Husby, A., Schielzeth, H., Forstmeier, W., Gustafsson, L. & Qvarnström, A. (2012) Data
 754 from: Sex chromosome linked genetic variance and the evolution of sexual dimorphism of
 755 quantitative traits. Artwork Size: 886077 bytes Pages: 886077 bytes.

756 James, C., Pemberton, J.M., Navarro, P. & Knott, S. (2024) Investigating pedigree- and
 757 SNP-associated components of heritability in a wild population of Soay sheep. *Heredity*,
 758 **132**, 202–210. <https://dx.doi.org/10.1038/s41437-024-00673-6>.

759 Kim, S.Y., Fargallo, J.A., Vergara, P. & Martínez-Padilla, J. (2013) Data from: Multivariate
 760 heredity of melanin-based coloration, body mass and immunity. Artwork Size: 658256 bytes
 761 Pages: 658256 bytes.

762 Kimock, C., KDubuc, C., Brent, L.J.N. & Higham, J.P. (2019) Rhesus Morphometrics.xlsx.

763 Kruuk, L.E.B. & Hadfield, J.D. (2007) How to separate genetic and environmental causes
 764 of similarity between relatives. *Journal of Evolutionary Biology*, **20**, 1890–1903. ISBN:
 765 1010-061X, <https://dx.doi.org/10.1111/j.1420-9101.2007.01377.x>.

766 Kruuk, L.E.B. (2004) Estimating genetic parameters in natural populations using
 767 the 'animal model'. *Philosophical Transactions of the Royal Society of Lon-*
 768 *don Series B: Biological Sciences*, **359**, 873–890. Publisher: Royal Society,
 769 <https://dx.doi.org/10.1098/rstb.2003.1437>.

770 Kvalnes, T., Sæther, B.E., Haanes, H., Røed, K.H., Engen, S., Solberg, E.J. & Saether, B.E.
 771 (2016a) Data from: Harvest-induced phenotypic selection in an island population of moose,
 772 *Alces alces*. Artwork Size: 52416 bytes Pages: 52416 bytes.

773 Kvalnes, T., Sæther, B.E., Haanes, H., Røed, K.H., Engen, S., Solberg, E.J. & Sæther, B.E.
774 (2016b) Data from: Harvest-induced phenotypic selection in an island population of moose,
775 *Alces alces*. Artwork Size: 52416 bytes Pages: 52416 bytes.

776 Lynch, M. & Walsh, B. (1998) *Genetics and Analysis of Quantitative Traits*. Sinauer Asso-
777 ciates, Inc., Sunderland, MA.

778 Malenfant, R.M., Davis, C.S., Richardson, E.S., Lunn, N.J. & Coltman, D.W. (2022) Data
779 from: Heritability of body size in the polar bears of Western Hudson Bay. Artwork Size:
780 2587048 bytes Pages: 2587048 bytes.

781 Martin, J., Wolak, M., Johnston, S., Pick, J. & Morrissey, M. (2026) *pedtricks: Visualize,*
782 *Summarize and Simulate Data from Pedigrees*. R package version 0.5.0.

783 Martin, J.G., Pirotta, E., Petelle, M.B. & Blumstein, D.T. (2017) Data from: Genetic basis
784 of between-individual and within-individual variance of docility. Artwork Size: 528839 bytes
785 Pages: 528839 bytes.

786 Mazé-Guilmo, E., Loot, G., Páez, D.J., Lefèvre, T., Blanchet, S., Lefevre, T., Paez, D.J. &
787 Maze-Guilmo, E. (2015) Data from: Heritable variation in host tolerance and resistance
788 inferred from a wild host– parasite system. Artwork Size: 86433 bytes Pages: 86433 bytes.

789 McFarlane, S.E., Boutin, S., Humphries, M.M., McAdam, A.G., Gorrell, J.C. & Coltman,
790 D.W. (2015) Data from: Very low levels of direct additive genetic variance in fitness and
791 fitness components in a red squirrel population. Artwork Size: 646460 bytes Pages: 646460
792 bytes.

793 Messel Nafstad, , Rønning, B., Aase, K., Ringsby, T.H., Hagen, I., Sjolte Ranke, P., Kvalnes,
794 T., Stawski, C., Räsänen, K., Sæther, B.E., Muff, S. & Jensen, H. (2023) Spatial variation
795 in the evolutionary potential and constraints of basal metabolic rate and body mass in a
796 wild bird. Artwork Size: 5564868550 bytes Pages: 5564868550 bytes.

Mills, J.A., Teplitsky, C., Arroyo, B., Charmantier, A., Becker, P.H., Birkhead, T.R., Bize, P.,
 Blumstein, D.T., Bonenfant, C., Boutin, S., Bushuev, A., Cam, E., Cockburn, A., Côté,
 S.D., Coulson, J.C., Daunt, F., Dingemanse, N.J., Doligez, B., Drummond, H., Espie,
 R.H.M., Festa-Bianchet, M., Frentiu, F., Fitzpatrick, J.W., Furness, R.W., Garant, D.,
 Gauthier, G., Grant, P.R., Griesser, M., Gustafsson, L., Hansson, B., Harris, M.P., Jiguet,
 F., Kjellander, P., Korpimäki, E., Krebs, C.J., Lens, L., Linnell, J.D.C., Low, M., McAdam,
 A., Margalida, A., Merilä, J., Møller, A.P., Nakagawa, S., Nilsson, J., Nisbet, I.C.T., van
 Noordwijk, A.J., Oro, D., Pärt, T., Pelletier, F., Potti, J., Pujol, B., Réale, D., Rockwell,
 R.F., Ropert-Coudert, Y., Roulin, A., Sederberg, J.S., Swenson, J.E., Thébaud, C., Visser,
 M.E., Wanless, S., Westneat, D.F., Wilson, A.J. & Zedrosser, A. (2015) Archiving Primary
 Data: Solutions for Long-Term Studies. *Trends in Ecology & Evolution*, **30**, 581–589.
<https://dx.doi.org/10.1016/j.tree.2015.07.006>.

Morrissey, M.B., Wilson, A.J., Pemberton, J.M. & Ferguson, M.M. (2007) A framework
 for power and sensitivity analyses for quantitative genetic studies of natural populations,
 and case studies in Soay sheep (*Ovis aries*). *Journal of Evolutionary Biology*, **20**, 2309–
 2321. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1420-9101.2007.01412.x>,
<https://dx.doi.org/10.1111/j.1420-9101.2007.01412.x>.

Morrissey, M.B., Walling, C.A., Wilson, A.J., Pemberton, J.M., Clutton-Brock, T.H. & Kruuk,
 L.E.B. (2022) Data from: Genetic analysis of life-history constraint and evolution in a wild
 ungulate population. Artwork Size: 28821 bytes Pages: 28821 bytes.

Morrissey, M.B. & Wilson, A.J. (2010) pedantics: an r package for pedigree-based genetic
 simulation and pedigree manipulation, characterization and viewing. *Molecular Ecology*
Resources, **10**, 711–719. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1755-0998.2009.02817.x>,
<https://dx.doi.org/10.1111/j.1755-0998.2009.02817.x>.

Munds, R.A., Cooper, E., Janiak, M.C., Lam, L.G., DeCasien, A., Baumann Surratt, S.,
 Montague, M., Martinez, M., Kawamura, S., Higham, J.P. & Melin, A.D. (2022) Rhesus
 macaque cone ratio heritability. Artwork Size: 614670 bytes Pages: 614670 bytes.

Nakagawa, S., Ivimey-Cook, E.R., Grainger, M.J., O'Dea, R.E., Burke, S., Drobniak, S.M., Gould, E., Macartney, E.L., Martinig, A.R., Morrison, K., Paquet, M., Pick, J.L., Pottier, P., Ricolfi, L., Wilkinson, D.P., Willcox, A., Williams, C., Wilson, L.A.B., Windecker, S.M., Yang, Y. & Lagisz, M. (2023) Method Reporting with Initials for Transparency (MeRIT) promotes more granularity and accountability for author contributions. *Nature Communications*, **14**, 1788. Publisher: Nature Publishing Group, <https://dx.doi.org/10.1038/s41467-023-37039-1>.

Nicolaus, M., Piau, R., Ubels, R., Tinbergen, J.M. & Dingemanse, N.J. (2016) Data from: The correlation between colouration and exploration behaviour varies across hierarchical levels in a wild passerine bird. Artwork Size: 705265 bytes Pages: 705265 bytes.

Olsson, M., Lindsay, W., Bererhi, B., Ljungström, G. & Erik, W. (2023) Pedigree and phenotypic variables used for estimation of heritability of sand lizard badge size and condition. Artwork Size: 390502 bytes Pages: 390502 bytes.

Pemberton, J.M., Kruuk, L.E.B. & Clutton-Brock, T. (2022) The Unusual Value of Long-Term Studies of Individuals: The Example of the Isle of Rum Red Deer Project. *Annual Review of Ecology, Evolution, and Systematics*, **53**, 327–351. <https://dx.doi.org/10.1146/annurev-eolsys-012722-024041>.

Perrier, C., Delahaie, B. & Charmantier, A. (2018) Heritability estimates from genomewide relatedness matrices in wild populations: Application to a passerine, using a small sample size. *Molecular Ecology Resources*, **18**, 838–853. <https://dx.doi.org/10.1111/1755-0998.12886>.

Petelle, M.B., Martin, J.G.A. & Blumstein, D.T. (2015) Data from: Heritability and genetic correlations of personality traits in a wild population of yellow-bellied marmots (*Marmota flaviventris*). Artwork Size: 785128 bytes Pages: 785128 bytes.

Pick, J.L., Allegue, H., Ajoy, Y.G.A., Class, B., Dochtermann, N.A., Ivimey-Cook, E.R., Laskowski, K., Réale, D., Wright, J., Dingemanse, N.J. & Westneat, D.F. (2025a) squidSim:

a flexible R package for structured and reproducible simulations in Ecology and Evolutionary
Biology. *EcoEvoRxiv*. Publisher: EcoEvoRxiv.

Pick, J.L., Allen, B.J., Bachelot, B., Bairos-Novak, K.R., Brand, J.A., Class, B., Dal-
las, T., D'Amelio, P.B., Fenollosa, E., Fernández-Juricic, E., Gomes, D.G.E., Grainger,
M.J., Guillemaud, T., John, C., Krasnow, R., Lagisz, M., Lequime, S., Maynard, D.S.,
Nakagawa, S., O'Dea, R.E., Paquet, M., Petitjean, Q., Sánchez-Tójar, A., van Dis,
N.E., Wilson, L.A.B. & Ivimey-Cook, E.R. (2026) The SORTEE guidelines for data and
code quality control in ecology and evolutionary biology. *Peer Community Journal*, **6**.
<https://dx.doi.org/10.24072/pcjournal.687>.

Pick, J.L., Kasper, C., Allegue, H., Dingemanse, N.J., Dochtermann, N.A., Laskowski,
K.L., Lima, M.R., Schielzeth, H., Westneat, D.F., Wright, J. & Araya-Ajoy, Y.G.
(2023) Describing posterior distributions of variance components: Problems and the
use of null distributions to aid interpretation. *Methods in Ecology and Evolution*, **14**,
2557–2574. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/2041-210X.14200>,
<https://dx.doi.org/10.1111/2041-210X.14200>.

Pick, J.L., Walling, C.A. & Kruuk, L.E.B. (2025b) Simple maternal effects animal models may
provide biased estimates of additive genetic and maternal variation. *Journal of Evolutionary
Biology*, **38**, 1556–1572. <https://dx.doi.org/10.1093/jeb/voaf104>.

Pigeon, G., Festa-Bianchet, M., Coltman, D.W. & Pelletier, F. (2022) Data from: Intense
selective hunting leads to artificial evolution in horn size. Artwork Size: 296462 bytes Pages:
296462 bytes.

Poissant, J., Morrissey, M.B., Gosler, A.G., Slate, J. & Sheldon, B.C. (2016) Data from:
Multivariate selection and intersexual genetic constraints in a wild bird population. Artwork
Size: 922120 bytes Pages: 922120 bytes.

Postma, E. (2014) Four decades of estimating heritabilities in wild vertebrate populations:

Improved methods, more data, better estimates? A. Charmantier, D. Garant & L.E.B. Kruuk, eds., *Quantitative Genetics in the Wild*, pp. 16–33. Oxford University Press, Oxford.

Quinn, J.L., Charmantier, a., Garant, D. & Sheldon, B.C. (2006) Data depth, data completeness, and their influence on quantitative genetic estimation in two contrasting bird populations. *Journal of Evolutionary Biology*, **19**, 994–1002. ISBN: 1010-061X (Print)\r1010-061X (Linking), <https://dx.doi.org/10.1111/j.1420-9101.2006.01081.x>.

R Core Team (2024) R: A Language and Environment for Statistical Computing. Place: Vienna, Austria.

Ramakers, J.J., Culina, A., Visser, M.E. & Gienapp, P. (2019) Data from: Environmental coupling of heritability and selection is rare and of minor evolutionary significance in wild populations. Artwork Size: 526940 bytes Pages: 526940 bytes.

Reed, T.E., Gienapp, P. & Visser, M.E. (2016) Data from: Testing for biases in selection on avian reproductive traits and partitioning direct and indirect selection using quantitative genetic models. Artwork Size: 1039977 bytes Pages: 1039977 bytes.

Regan, C., Pemberton, J., Pilkington, J., Smiseth, P. & Wilson, A. (2019) Data from: Linking genetic merit to sparse behavioral data: does behavior explain genetic variation for maternal care in Soay sheep? Artwork Size: 387508565 bytes Pages: 387508565 bytes.

Regan, C.E., Pemberton, J.M., Pilkington, J.G., Smiseth, P.T. & Wilson, A.J. (2020) Linking genetic merit to sparse behavioral data: behavior and genetic effects on lamb growth in Soay sheep. *Behavioral Ecology*, **31**, 114–122. <https://dx.doi.org/10.1093/beheco/arz166>.

Reid, J.M., Arcese, P., Keller, L.F. & Losdat, S. (2014) FEMALE AND MALE GENETIC EFFECTS ON OFFSPRING PATERNITY: ADDITIVE GENETIC (CO)VARIANCES IN FEMALE EXTRA-PAIR REPRODUCTION AND MALE PATERNITY SUCCESS IN SONG SPARROWS (MELOSPIZA MELODIA). *Evolution*, **68**, 2357–2370. <https://dx.doi.org/10.1111/evo.12424>.

899 Rosenthal, M.F., Gertler, M., Hamilton, A.D., Prasad, S. & Andrade, M.C.B. (2017)
 900 Taxonomic bias in animal behaviour publications. *Animal Behaviour*, **127**, 83–89.
 901 <https://dx.doi.org/10.1016/j.anbehav.2017.02.017>.

902 Sakaluk, S.K., Wilson, A.J., Bowers, E.K., Johnson, L.S., Masters, B.S., Johnson, B.G., Vogel,
 903 L.A., Forsman, A.M. & Thompson, C.F. (2014) Data from: Genetic and environmental
 904 variation in condition, cutaneous immunity, and haematocrit in house wrens. Artwork Size:
 905 231921 bytes Pages: 231921 bytes.

906 Sauve, D., Divoky, G. & Friesen, V.L. (2019) Data from: Phenotypic plasticity or evolutionary
 907 change? An examination of the phenological response of an arctic seabird to climate change.
 908 Artwork Size: 348611 bytes Pages: 348611 bytes.

909 Sparks, A., Spurgin, L., van der Velde, M., Fairfield, E., Komdeur, J., Burke, T., Richardson,
 910 D. & Dugdale, H. (2022) Data from: Telomere heritability and parental age at conception
 911 effects in a wild avian population. Artwork Size: 531564 bytes Pages: 531564 bytes.

912 Stedman, J.M., Hallinger, K.K., Winkler, D.W. & Vitousek, M.N. (2017) Data from: Heritable
 913 variation in circulating glucocorticoids and endocrine flexibility in a free-living songbird.
 914 Artwork Size: 163731 bytes Pages: 163731 bytes.

915 Strickland, K., Cristescu, R., Schultz, A., Kruuk, L., de Villiers, D. & Frere, C. (2022)
 916 Data from: Susceptibility to a sexually transmitted disease in a wild koala population shows
 917 heritable genetic variance but no inbreeding depression. Artwork Size: 2204782 bytes Pages:
 918 2204782 bytes.

919 Strickland, K., Jones, M.E., Storfer, A., Hamede, R.K., Hohenlohe, P.A., Margres,
 920 M.J., McCallum, H.I., Comte, S., Lachish, S. & Kruuk, L.E.B. (2024) Adaptive po-
 921 tential in the face of a transmissible cancer in Tasmanian devils. *Molecular Ecol-*
 922 *ogy*, **33**, e17531. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/mec.17531>,
 923 <https://dx.doi.org/10.1111/mec.17531>.

924 Strickland, K., Mitchell, D., Delmé, C. & Frère, C. (2021a) Data from: Repeatability and
 925 heritability of social reaction norms in a wild agamid lizard. Artwork Size: 8272538 bytes
 926 Pages: 8272538 bytes.

927 Strickland, K., Mitchell, D.J., Delmé, C. & Frère, C.H. (2021b) Repeatability and her-
 928 itability of social reaction norms in a wild agamid lizard. *Evolution*, **75**, 1953–1965.
 929 <https://dx.doi.org/10.1111/evo.14298>.

930 Strickland, K., Räsänen, K., Kristjánsson, B., Phillips, J., Einarsson, A., Snorradóttir, R.,
 931 Bartrons, M. & Jónsson, Z.O. (2023a) Data from: Genome-phenotype-environment asso-
 932 ciations identify signatures of selection in a panmictic population of threespine stickleback.
 933 Artwork Size: 110341 bytes Pages: 110341 bytes.

934 Strickland, K., Räsänen, K., Kristjánsson, B.K., Phillips, J.S., Einarsson,
 935 A., Snorradóttir, R.G., Bartrons, M. & Jónsson, Z.O. (2023b) Genome-
 936 phenotype-environment associations identify signatures of selection in a pan-
 937 mictic population of threespine stickleback. *Molecular Ecology*, **32**, 1708–
 938 1725. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/mec.16845>,
 939 <https://dx.doi.org/10.1111/mec.16845>.

940 Tarka, M., Åkesson, M., Hasselquist, D. & Hansson, B. (2013) Data from: Intralocus sexual
 941 conflict over wing length in a wild migratory bird. Artwork Size: 103305 bytes Pages:
 942 103305 bytes.

943 Teplitsky, C., Mouawad, N.G., Balbontín, J., de Lope, F. & Møller, A.P. (2011) Data from:
 944 Quantitative genetics of migration syndromes: a study of two barn swallow populations.
 945 Artwork Size: 236248 bytes Pages: 236248 bytes.

946 Thomson, C.E., Winney, I.S., Salles, O.C. & Pujol, B. (2018) A guide to using a multiple-
 947 matrix animal model to disentangle genetic and nongenetic causes of phenotypic variance.
 948 *PLOS ONE*, **13**, e0197720. <https://dx.doi.org/10.1371/journal.pone.0197720>.

949 Vedder, O. (2021) Full data on telomere length and lifespan. Artwork Size: 286030 bytes
 950 Pages: 286030 bytes.

951 Visscher, P.M. & Goddard, M.E. (2015) A General Unified Framework to Assess the Sampling
 952 Variance of Heritability Estimates Using Pedigree or Marker-Based Relationships. *Genetics*,
 953 **199**, 223–232. <https://dx.doi.org/10.1534/genetics.114.171017>.

954 Visscher, P.M., Hemani, G., Vinkhuyzen, A.A.E., Chen, G.B., Lee, S.H., Wray, N.R., Goddard,
 955 M.E. & Yang, J. (2014) Statistical Power to Detect Genetic (Co)Variance of Complex Traits
 956 Using SNP Data in Unrelated Samples. *PLOS Genetics*, **10**, e1004269. Publisher: Public
 957 Library of Science, <https://dx.doi.org/10.1371/journal.pgen.1004269>.

958 vonHoldt, B.M., DeCandia, A.L., Heppenheimer, E., Janowitz-Koch, I., Shi,
 959 R., Zhou, H., German, C.A., Brzeski, K.E., Cassidy, K.A., Stahler, D.R. &
 960 Sinsheimer, J.S. (2020) Heritability of interpack aggression in a wild pedi-
 961 greed population of North American grey wolves. *Molecular Ecology*, **29**,
 962 1764–1775. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/mec.15349>,
 963 <https://dx.doi.org/10.1111/mec.15349>.

964 Wei, T. & Simko, V. (2021) *R package 'corrplot': Visualization of a Correlation Matrix*.
 965 (Version 0.92).

966 Westneat, D., Young, R., Cones, A., Kucera, A., Anacleto, A. & Heidinger, B. (2023) Early
 967 life telomeres are influenced by environments acting at multiple temporal and spatial scales.
 968 Artwork Size: 738640 bytes Pages: 738640 bytes.

969 Wheelwright, N.T., Keller, L.F. & Postma, E. (2014) Data from: The effect of trait type and
 970 strength of selection on heritability and evolvability in an island bird population. Artwork
 971 Size: 570849 bytes Pages: 570849 bytes.

972 Wilson, A.J., Réale, D., Clements, M.N., Morrissey, M.M., Postma, E., Walling, C.A., Kruuk,
 973 L.E.B. & Nussey, D.H. (2010) An ecologist's guide to the animal model. *Journal of Ani-*

mal *Ecology*, **79**, 13–26. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2656.2009.01639.x>, <https://dx.doi.org/10.1111/j.1365-2656.2009.01639.x>.

Winney, I.S., Schroeder, J., Nakagawa, S., Hsu, Y.H., Simons, M.J., Sánchez-Tójar, A., Mannarelli, M.E. & Burke, T. (2019) Data from: Heritability and social brood effects on personality in juvenile and adult life-history stages in a wild passerine. Artwork Size: 1102111 bytes Pages: 1102111 bytes.

Wolak, M.E. (2012) nadiv : an R package to create relatedness matrices for estimating non-additive genetic variances in animal models. *Methods in Ecology and Evolution*, **3**, 792–796. <https://dx.doi.org/10.1111/j.2041-210X.2012.00213.x>.

Wolak, M.E. (2024) gremlin: R package for mixed-effects model reml incorporating generalized inverses. *Zenodo*.

Wolak, M.E. & Keller, L.F. (2014) Dominance genetic variance and inbreeding in natural populations. A. Charmantier, D. Garant & L.E.B. Kruuk, eds., *Quantitative Genetics in the Wild*, p. 0. Oxford University Press.

Wolak, M.E. & Reid, J.M. (2017) Accounting for genetic differences among unknown parents in microevolutionary studies: how to include genetic groups in quantitative genetic animal models. *Journal of Animal Ecology*, **86**, 7–20. <https://dx.doi.org/10.1111/1365-2656.12597>.

Yamamura, K. (1999) Transformation using $(x + 0.5)$ to stabilize the variance of populations. *Population Ecology*, **41**, 229–234. ISBN: 0034-5466, <https://dx.doi.org/10.1007/s101440050026>.

Young, E.A. & Postma, E. (2023a) Low interspecific variation and no phylogenetic signal in additive genetic variance in wild bird and mammal populations. *Ecology and Evolution*, **13**, e10693. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/ece3.10693>, <https://dx.doi.org/10.1002/ece3.10693>.

- 999 Young, E.A. & Postma, E. (2023b) Replication Data for: Low interspecific variation and no
1000 phylogenetic signal in additive genetic variance in wild bird and mammal populations.
- 1001 Zaitlen, N., Kraft, P., Patterson, N., Pasaniuc, B., Bhatia, G., Pollack, S. &
1002 Price, A.L. (2013) Using Extended Genealogy to Estimate Components of Heritabil-
1003 ity for 23 Quantitative and Dichotomous Traits. *PLOS Genetics*, **9**, e1003520.
1004 <https://dx.doi.org/10.1371/journal.pgen.1003520>.
- 1005 Zedrosser, A. (2019) R-Data for: Heritability of head size in a hunted large carnivore, the
1006 brown bear (*Ursus arctos*).
- 1007 Zub, K., Piertney, S., Szafrńska, P.A. & Konarzewski, M. (2012) Environmen-
1008 tal and genetic influences on body mass and resting metabolic rates (RMR) in
1009 a natural population of weasel *Mustela nivalis*. *Molecular Ecology*, **21**, 1283–
1010 1293. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-294X.2011.05436.x>,
1011 <https://dx.doi.org/10.1111/j.1365-294X.2011.05436.x>.

Supplementary Material for:

Pedigree metrics: What are they good for?

Joel L. Pick, Barbara Class, Paul Acker, Anne Charmantier, Niels J. Dingemanse , Ned A. Dochtermann, Julien G.A. Martin, Denis Réale, Jane M. Reid, Holger Schielzeth, Celine Teplitsky, Kasha Strickland, Yimen G. Araya-Ajoy

S1 Reporting over time

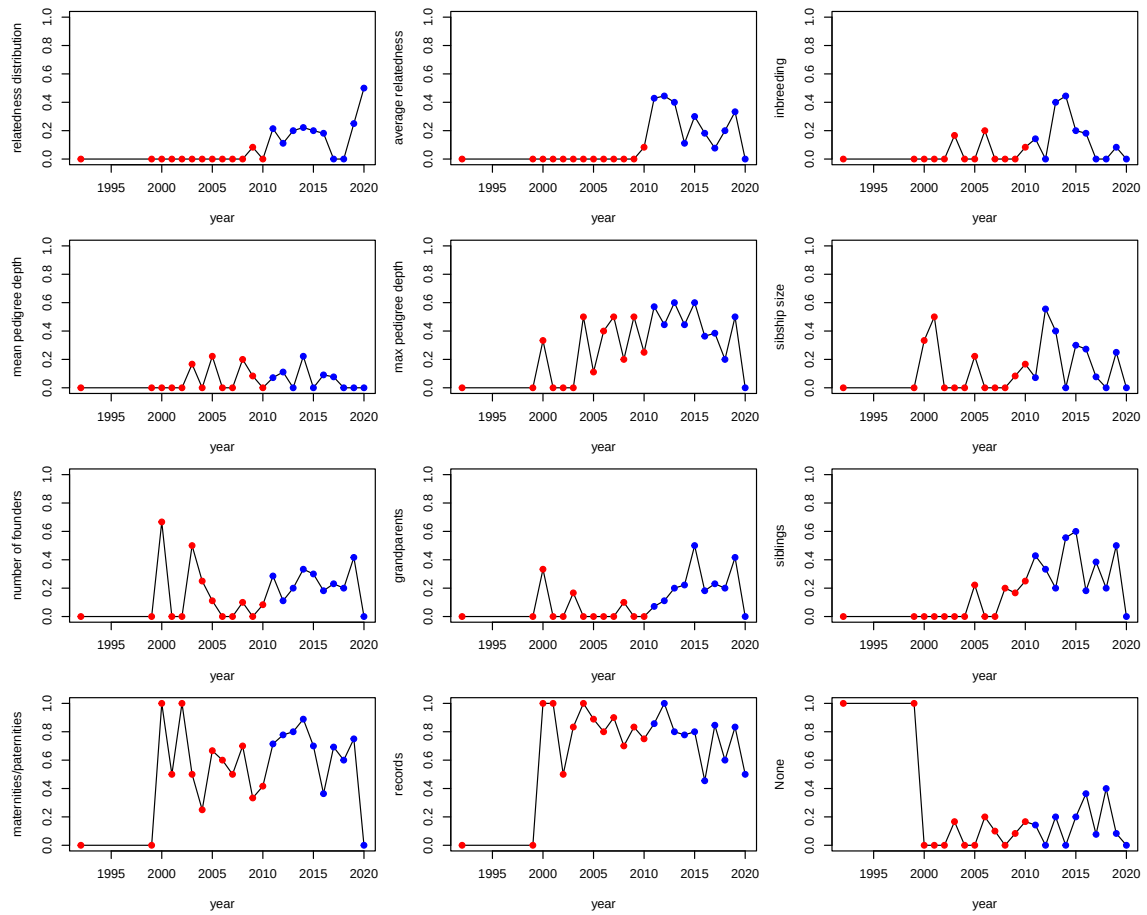


Figure S1: Changes in the reporting of different metrics over time. Red point represent up until 2010 when pedantics was published, and blue represents after 2010.

S2 Pedigrees

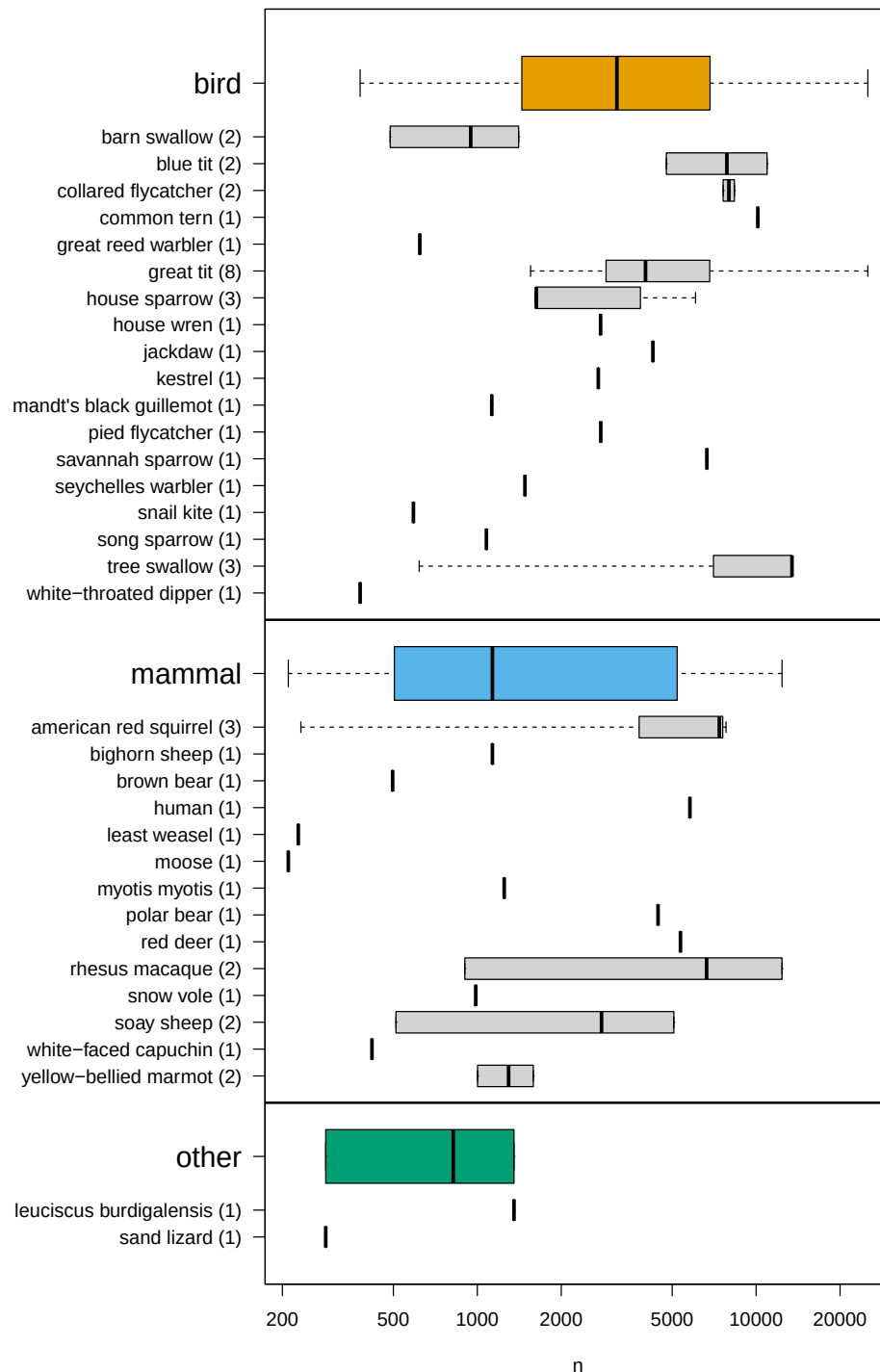


Figure S2: Size and taxa of the 53 pedigrees used in the study. Individual boxplots show the range of sample sizes within each taxa (coloured boxes) and species (grey boxes). Numbers in brackets show how many pedigrees there were per species.

S3 Metrics and precision

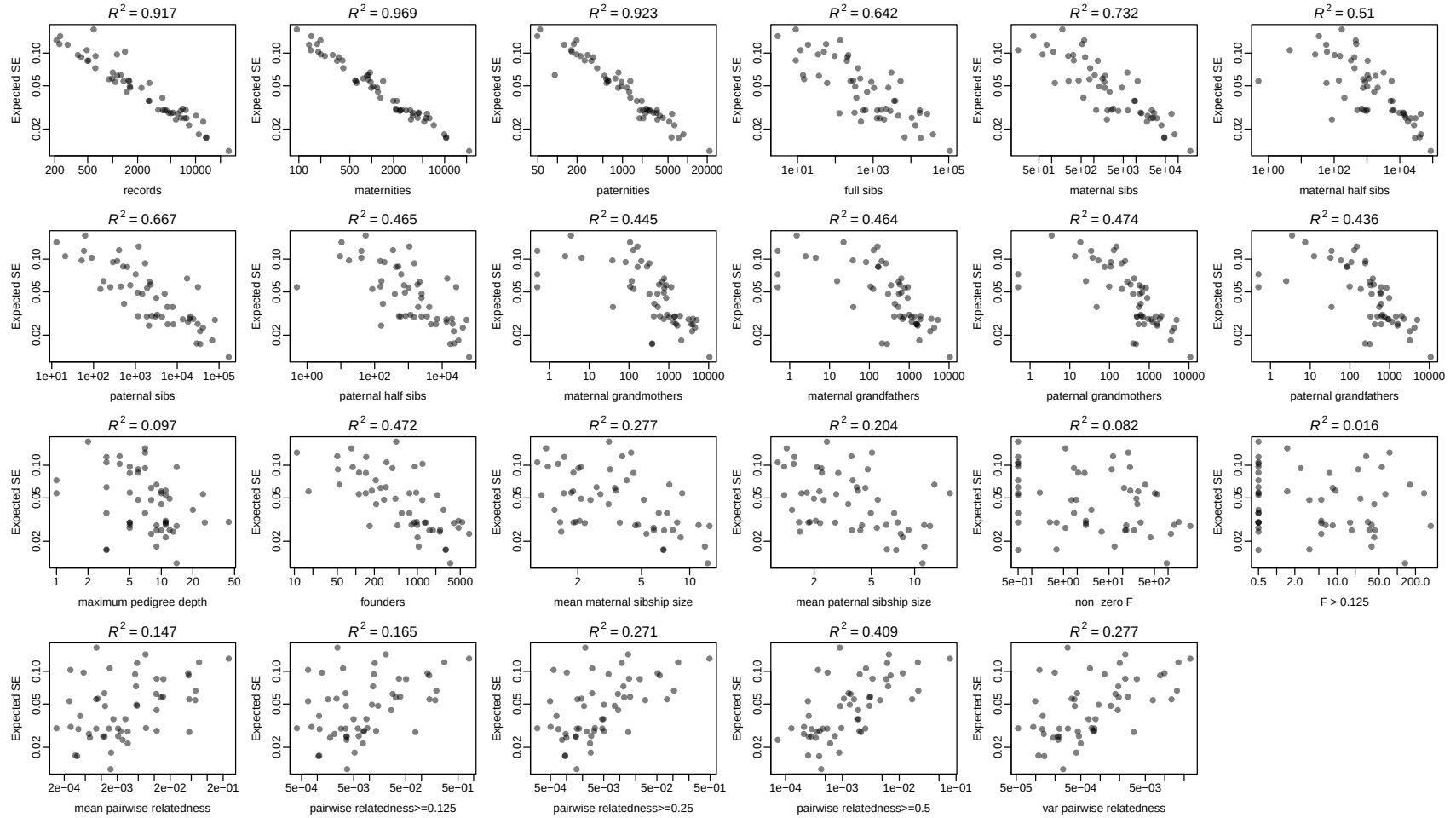


Figure S3: Relationship between expected standard error of $\hat{h}^2$ (based on simulations, where $h^2 = 0.3$) and different descriptive pedigree metrics, across 53 wild pedigrees. Note that axes are on the log scale.

S3.1 Sensitivity Analyses

To assess that the relationships between the different metrics and expected standard error of $\hat{h}^2$ are not dependent on the heritability we used in our simulations, we additionally ran simulations with heritabilities of 0.1 and 0.5.

Metric	h^2		
	0.1	0.3	0.5
records	0.872	0.917	0.933
maternities	0.987	0.969	0.946
paternities	0.949	0.923	0.908
full sibs	0.71	0.642	0.605
maternal sibs	0.835	0.732	0.68
maternal half sibs	0.622	0.51	0.463
paternal sibs	0.783	0.667	0.621
paternal half sibs	0.588	0.465	0.423
maternal grandmothers	0.509	0.445	0.423
maternal grandfathers	0.529	0.464	0.444
paternal grandmothers	0.536	0.474	0.455
paternal grandfathers	0.5	0.436	0.42
maximum pedigree depth	0.109	0.097	0.097
founders	0.372	0.472	0.529
mean maternal sibship size	0.404	0.277	0.228
mean paternal sibship size	0.324	0.204	0.163
non-zero F	0.139	0.082	0.067
F > 0.125	0.066	0.016	0.003
mean pairwise relatedness	0.068	0.147	0.184
pairwise relatedness>=0.125	0.08	0.165	0.206
pairwise relatedness>=0.25	0.167	0.271	0.321
pairwise relatedness>=0.5	0.305	0.409	0.464
var pairwise relatedness	0.172	0.277	0.324

Figure S4: R^2 of regression of expected SE of $\hat{h}^2$ (from simulations) and pedigree metrics, across varying simulated h^2

We also assessed how well $SE_{\hat{h}^2}^{(1)}$ and $SE_{\hat{h}^2}^{(2)}$ predicted across varying h^2 .

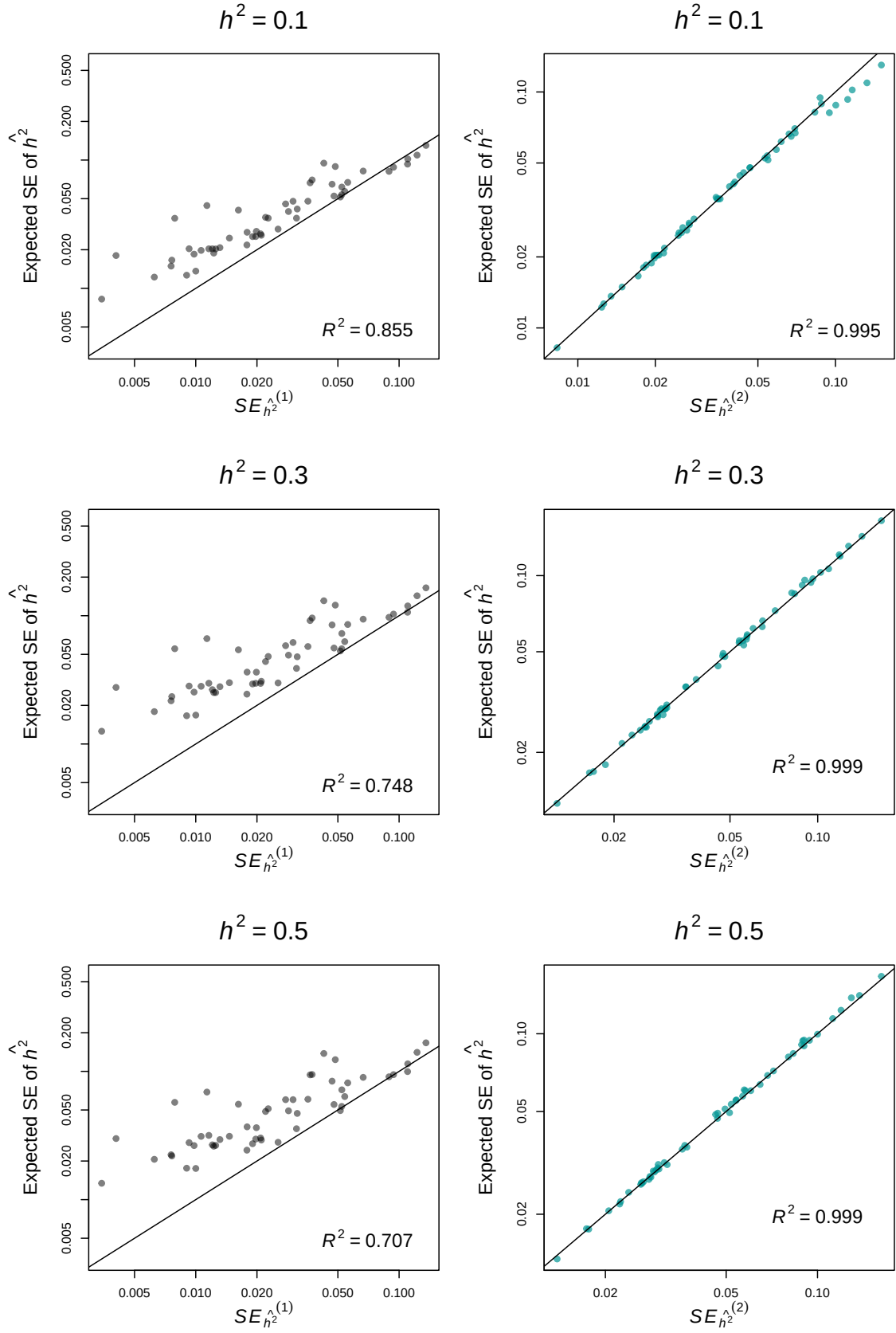


Figure S5: Relationship between the two analytical approximations of the SE of h^2 ($SE_{h^2}^{(1)}$ and $SE_{h^2}^{(2)}$) and the expected standard error of h^2 based on simulations, across 53 wild pedigrees. $SE_{h^2}^{(2)}$ and the simulations were run with across three h^2 s: 0.1, 0.3 and 0.5. The black line represents the 1:1 relationship.

S4 Metrics and minimum detectable effect

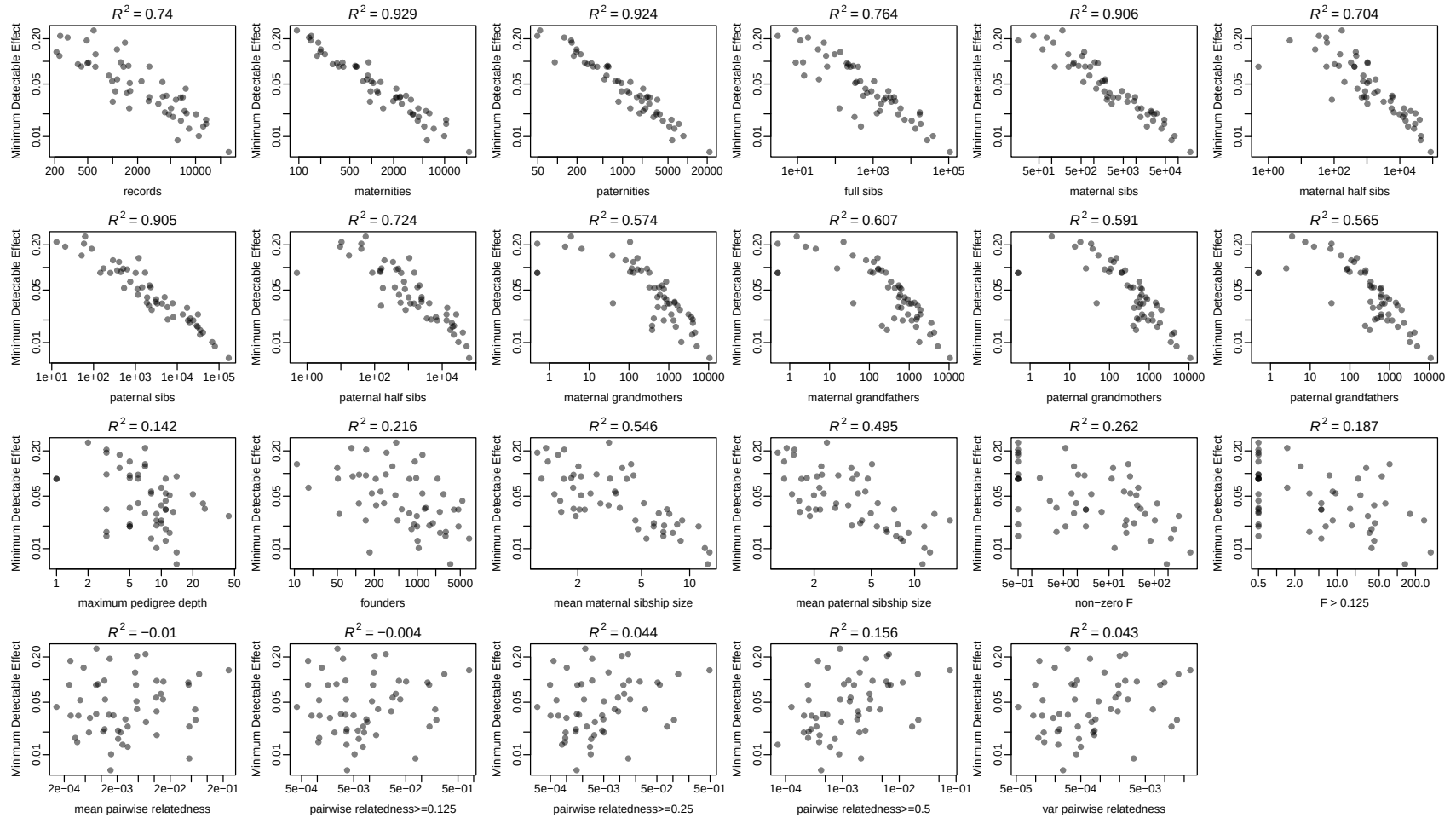


Figure S6: Relationship between minimum detectable effect (based on simulations) and different descriptive pedigree metrics, across 53 wild pedigrees. Note that axes are on the log scale.

S5 GRMs

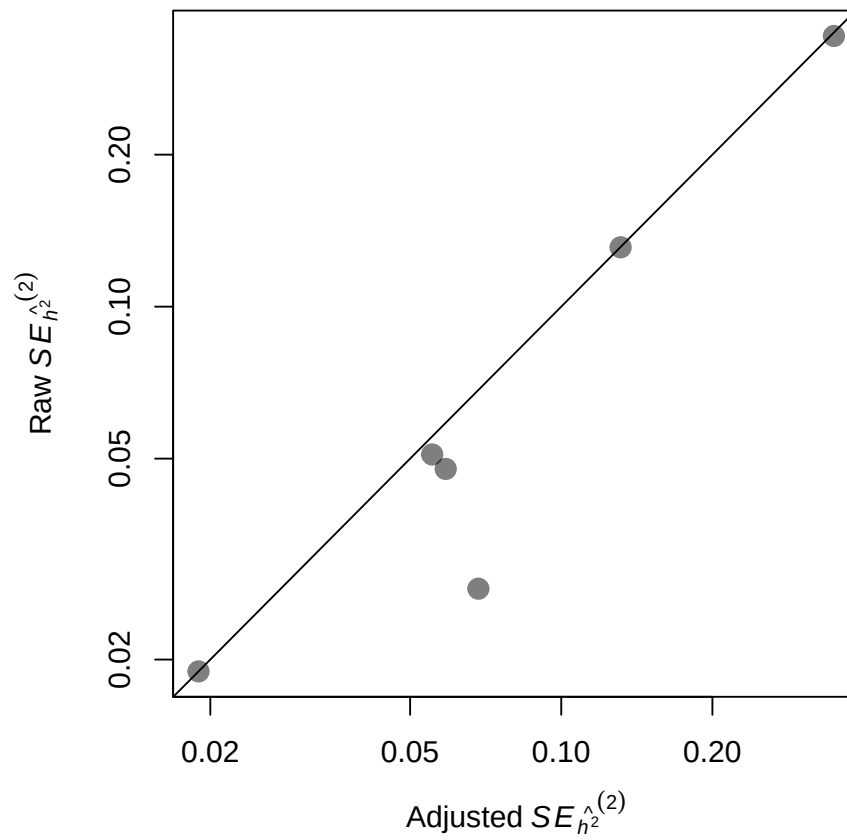


Figure S7: Comparison between $SE_{h^2}^{(2)}$ calculated with raw GRMs and $SE_{h^2}^{(2)}$ calculated from GRMs that have been made positive definite (adjusted). The difference here comes from the raw GRMs have negative eigenvalues, which inflates the variance in the eigenvalues, whereas the adjustment to make them positive definite caps the eigenvalue at 0.

S6 Metrics and bias

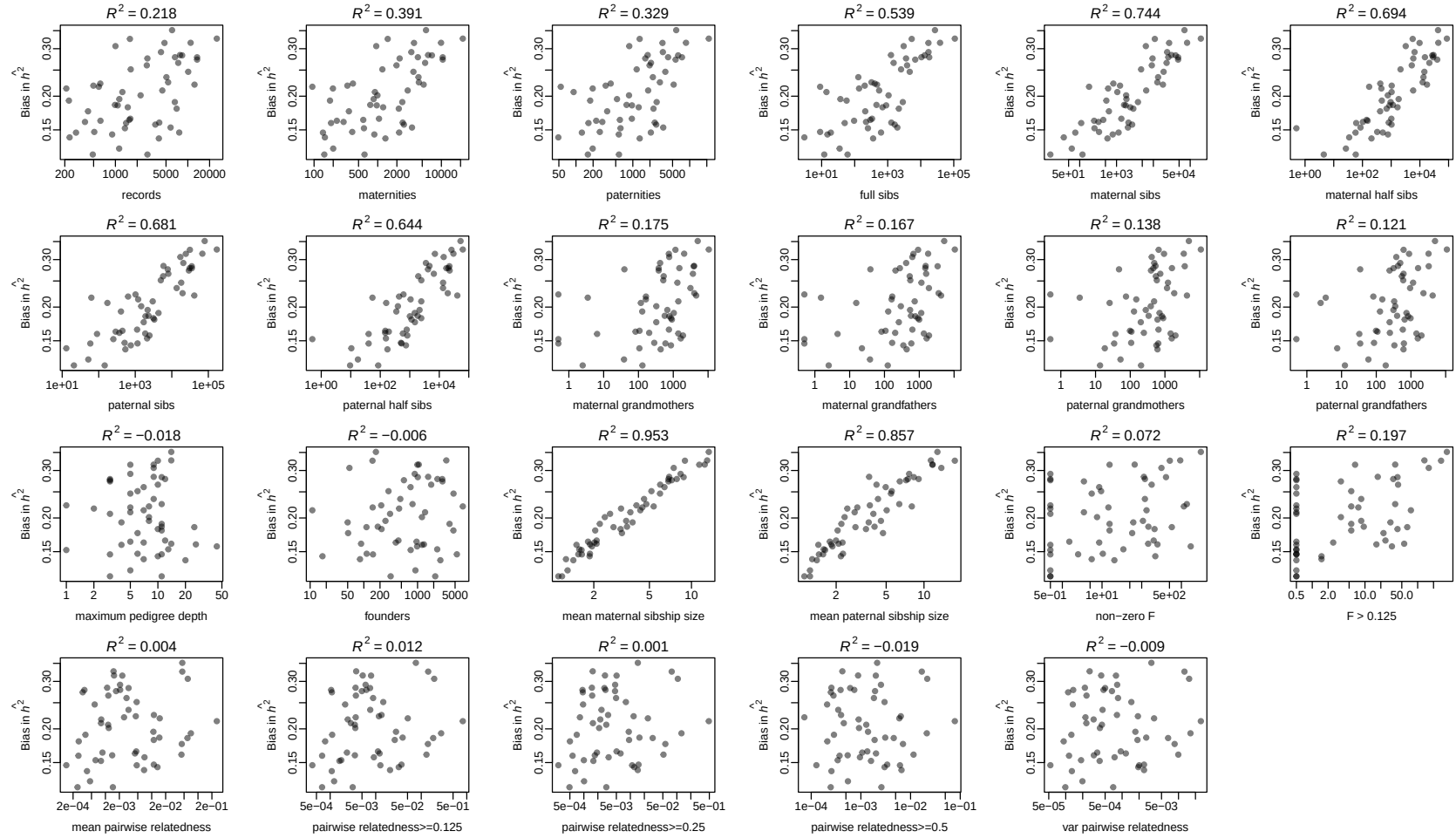


Figure S8: Relationship between the bias in $\hat{h}^2$ from unmodelled maternal effects, based on simulations with $h^2 = 0.1$ and $m^2 = 0.2$, and different descriptive pedigree metrics, across 53 wild pedigrees. Note that axes are on the log scale.