The impacts of environmental conditions on inbreeding depression: a meta-analysis.

Yuan Fu Chan

PhD Thesis

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor of Philosophy by Yuan Fu Chan

31 August 2022

Abstract

Inbreeding depression, the reduction in fitness of inbred offspring, is a major concern in the conservation and management of small and isolated populations. While it is known to reduce individual fitness, the magnitude and severity of inbreeding depression is also known to vary, depending on how genotypes interact with the environment. Evidence synthesis is required for conservation practitioners to understand and predict this variation and to facilitate management of this genetic problem in wild populations of animals and plants. In this thesis, I conducted a meta-analysis to investigate how inbreeding depression was influenced by a range of factors, including environmental variation and stress, and to identify the most important sources of heterogeneity in inbreeding depression in populations of animals and plants. I carried out literature searches to retrieve 129 articles documenting phenotypic differences between outbred and inbred populations of animals and plants. First, automated-content analysis (ACA) and linear discriminant analysis (LDA) were tested as methods to assist abstract assessment in order to expedite the identification of relevant articles describing inbreeding depression, from a larger body of literature search results. In chapter 2, I demonstrated that analysis of article abstracts with ACA and LDA was a viable method for screening relevant articles for meta-analysis with a 45% reduction in workload accompanied by a loss of 6 % of potentially relevant articles. In chapter 3, I used meta-analysis to identify significant variation of inbreeding depression with different types of phenotypic traits and life history stages, with fecundity, fitness and traits at late life history stages suffering more inbreeding depression. However, I did not detect any significant variation among different taxonomic groups. In chapter 4, I showed that inbreeding depression occurred irrespective of environmental stress, both in fitness and non-fitness component traits. However, in contrast to earlier syntheses, I found

that the magnitude of inbreeding depression in fitness component traits was not greater in populations experiencing environmental stress than in those in benign environments. In conclusion, my results show that (i) inbreeding depression is a ubiquitous response to inbreeding in natural populations of conservation concern, and (ii) that variation in inbreeding depression can be anticipated by considering only a small number of sources of heterogeneity. When considering the conservation management of small and isolated populations, carefully managed geneflow is likely to prove beneficial in most cases, since moving individuals or populations to new (potentially stressful) environments does not intensify inbreeding depression, and could boost adaptive potential. Rather, my results suggest that environmental stress and inbreeding depression independently generate fitness costs that are expected to undermine population persistence.

Acknowledgement

This work was supported by the University of Liverpool, as a part of Ph D project. I thank David Atkinson and Ilik Saccheri for constructive criticisms on the structure of this meta-analysis. I also thank Philipp Antczak's kind help to provide high performic computer for literature analysis. Finally, I will thank my supervisors, Raj Whitlock and Jenny Hodgson gave guidance on design, methods, project conception and assistance with statistical methods and providing feedback.

Contents

Chapter 1. General introduction

It often occurred to me that it would be advisable to try whether seedlings from cross-fertilised flowers were in any way superior to those from self-fertilised flowers.

(Darwin, 1900)

In Darwin's experiments on 57 plant species, he documented that self-fertilised plants have inferior phenotypic traits compared with cross-fertilised plants, including smaller plant size and lower seed production. This was the first work to quantify phenotypic consequences from inbreeding depression. It has also shed light on the adverse effects of inbreeding on the fitness of organisms for future researchers.

Since Darwin's pioneering study, inbreeding depression has been widely studied in the wild and its impacts has been recognised as ecological importance (Crnokrak & Roff, 1999; Keller & Waller, 2002; O'Grady *et al*., 2006). Inbreeding depression, the reduction in fitness as a consequence of mating between relatives, has been well documented in populations of animals and plants by evolutionary biologists (Keller & Waller, 2002). The detrimental effects of inbreeding depression on fitness reductions in endangered species has also drawn the attention of conservation biologists and ecologists. Inbreeding depression along with climate changes and anthropogenic activities are major threats to population persistence. However, little is known about how inbreeding depression interacts with environmental changes influences demographic contexts. Evidence has shown that the magnitude and severity of inbreeding depression varies depending on how genotypes interact with the environment (Cheptou & Donohue, 2011). However, fitness variation of inbreeding depression across environments in natural populations has been little addressed. Efficient conservation management that takes into account the effects of genetic factors and environmental variation is still in its infancy, partly because the available evidence has not been synthesised and communicated to conservation practitioners. This thesis aims to provide an improved understanding of the factors—environmental, demographic and species' characteristics—that cause variation in inbreeding depression, and then incorporate this information into a conservation framework to predict inbreeding risks under global environmental changes.

In this chapter, I will first review the pivotal role of genetic variation in evolutionary potential. How new phenotypes are generated through mutation and their role in adaptation to environmental changes. How global climate changes and human activities could pressure population persistence by degrading genetic diversity will be also addressed. Following that, I will consider loss of evolutionary potential stemmed from the core genetic problems, inbreeding and genetic drift that lead to the decline in individual fitness and eventually threaten population sustainability. The discussion on how these genetic problems could be translated into extinction risk is provided. Additionally, I will review the potential variables to construct a generalisation between fitness reduction and inbreeding, specifically, the environmental variation and species/population characteristics that are required to provide comprehensive understanding of the impacts of inbreeding depression. Finally, I review the viable conservation measures that could alleviate inbreeding depression in wild populations and finish by detailing the goals and scope of this thesis.

1.1 Genetic variation and evolutionary adaptation

Mutation, which arises from errors in DNA repair, is the ultimate source of genetic variation (Frankham, Ballou, & Briscoe, 2010). The effects of a mutation on fitness can be classified as neutral, deleterious and adaptive, based on its selection coefficient (Cutter, 2019). When mutations occur in non-coding regions, they are expected to be neutral since they don't change the expression of phenotypes. However, if mutations occur in coding regions, they are likely to change protein function through the alteration of amino acids, consequently generating new phenotypes. Typically, these mutations are predominantly deleterious due to the disruption of efficiently functioning protein (Cutter, 2019). Occasionally, the fitness effects of a mutation can be flexible, depending on environmental background. In particular, a new phenotype generated by mutation could be beneficial to environmental changes depending on genotype × environment interactions (Kondrashov & Houle, 1994).

Genetic variation is a fundamental element that underpins evolutionary adaptation in natural populations. It is the raw material upon which natural selection can operate,

enabling evolutionary adaptation to changing environments (Hoffmann & Sgro, 2011). When evolutionary adaptation occurs, genetic variation allows populations to adapt to changing climatic conditions by shifting the frequency of genotypes underlying traits influencing fitness (Hoffmann & Sgro, 2011). Therefore, the persistence of populations through environmental changes depends on their genetic constitution (Hoffmann & Willi, 2008). Specifically, the extent of genetic variation in a population determines its evolutionary potential. Breeder's equation, which has been applied to predict population persistence to climate change (Sinervo *et al*., 2010), state that the response to selection (R, implying change in mean of a phenotypic trait) can be quantified by heritability (h^2, a measurement to quantify the genetic contribution to quantitative traits) and selection differential (S), as follows:

$$R = h^2 S \qquad \text{(Lush, 1943)}$$

Similarly, Li and co-workers (2000; 1999) showed that the genetic diversity in wild emmer wheat (*Triticum dicoccoides*) allowed this species to locally adapt to different environmental variation across geographic microsites. These examples further reinforce the role of genetic diversity in evolutionary adaptation when populations are experienced global environmental changes.

Some studies have demonstrated that natural populations are able to adapt to cope with climate changes. One of the most recognised examples of evolutionary adaption was discovered in Darwin's finches (*Geospiza fortis*). Given the high heritability of beak size, this trait is likely to subject to evolutionary response and natural selection. Evolutionary shift toward smaller beak size enables Darwin's finches to take advantage of small seeds when large seeds are scare due to being consumed by other species or during drought period. Grant and Grant (2006) have shown that drought conditions in combination with competition from *G. magnirostris* favoured smaller beaks in 2004, whereas larger beaks were selected as competition was absence in 1997. Meanwhile, Reale and colleagues (2003) detected that warmer spring temperature and the increase in food abundance advanced parturition dates of red squirrels in Canada (*Tamiasciurus hudsonicus*). Applying an animal model, they identified that evolution accounted for 13% of the phenotypic changes in parturition date.

Maintaining sufficient genetic diversity is a major concern in conservation biology to ensure the long-term sustainability of populations (Norberg, Urban, Vellend, Klausmeier, & Loeuille, 2012; Walsh & Blows, 2009). However, many wild populations are faced with threats from global climate change and other human-associated environmental changes, such as habitat fragmentation, overexploitation and pollution (Vitousek, Mooney, Lubchenco, & Melillo, 1997). The emission of greenhouse gases is predicted to increase global average temperature between 1.5 to 2 ℃ within this century (IPCC, 2022). These anthropogenic threats along with global warming could reduce habitat areas or fragment a large habitat into different patches, leading to isolated and fragmented populations. When population sizes remain small and isolated, species can suffer from inbreeding and genetic drift, consequently decreasing fitness and limiting evolutionary potential in nature populations (Frankham *et al.*, 1999). In a common garden experiment, Potvin and Tousignant (1996) observed that *Brassica juncea* failed to evolve in response to simulated climate change, due to low levels of genetic diversity. Another experiment carried out by Hoffmann (2003) illustrated that *Drosophila birchii* was unable to adapt to desiccated environment because of the lack of heritable variation. Consequently, low genetic diversity arising from genetic threats could hinder populations to respond to future environmental changes. In the following two sections, we will consider the two main genetic problems that small populations could face, genetic drift and inbreeding.

1.2 Genetic drift

When sexual reproduction occurs, the genotypes of offspring zygotes are derived from a subsampling of grand-parental alleles within each parent individual during gametogenesis. Allelic variants represented in offspring generations are therefore a product of the sample of gametes that united successfully to form zygotes. Genetic drift is the process of stochastic change in allele frequencies arising from the random sampling of alleles between generations. (Frankham *et al.*, 2010). For a population with effective population size N_e and allele frequency p, the predicted change in allele frequency can be quantified according to a binomial distribution:

$$2p\ (1\text{-}p)/2N_e$$

Specifically, small populations will have a greater probability of large changes in allele frequencies, and therefore are more likely to lose alleles, making them more vulnerable to genetic impoverishment (Cutter, 2019).

1.3 Inbreeding

Inbreeding is defined as the production of offspring from the mating of related individuals or an increase in homozygosity related to such mating (Keller & Waller, 2002). The magnitude of inbreeding can be quantified by the inbreeding coefficient *F*. Wright's F is defined as the probability that two alleles are derived from a common ancestor, which is termed identity by descent (Wright, 1922). Additionally, the inbreeding coefficient signifies the extent of deviation in homozygosity from that expected under Hardy-Weinberg equilibrium, and can be calculated as follows:

$$F = \frac{H_{exp} - H_{obs}}{H_{exp}}$$

where H_{exp} and H_{obs} indicate the expected and observed heterozygosity (Cutter, 2019). This measures the deviation of heterozygosity frequency from a random mating population under Hardy-Weinberg equilibrium. For example, an inbreeding coefficient of 0.5 implies that this population has only 50% of heterozygosity frequency compared to a population under Hardy-Weinberg equilibrium.

Inbreeding inevitably increases the frequencies of homozygotes, driving declines in heterozygosity (Keller & Waller, 2002). Under the dominance model, when the proportion of homozygotes increases under inbreeding, deleterious mutations are more likely to be expressed due to the unmasking of recessive alleles that confer detrimental phenotypes in the homozygous state. Alternatively, if loci are maintained by heterozygote advantage (overdominance model), the decline in heterozygotes upon inbreeding will reduce individual fitness (Charlesworth. & Willis, 2009). This fitness reduction relative to that expected under random mating is referred to as *inbreeding depression*. The magnitude of inbreeding depression can be quantified as the proportionate decline in fitness values resulting from inbreeding (Frankham *et al.,* 2010):

$$\text{Inbreeding depression} = 1 - \frac{\text{fitness of inbred offspirng}}{\text{fitness of outbred offspring}}$$

Fitness reduction arising from inbreeding depression could lower individual's survival rate and reproductive output, subsequently decreasing population viability. Concern arising whether inbreeding depression could threaten population sustainability in conservation context.

1.4 The genetic consequences of small population size

As gametes are sampled randomly during sexual reproduction, genetic diversity might decline in small populations due to the stochastic loss of alleles. Loss of allelic variation inevitably increases the frequency of homozygotes. When populations remain small, the effects of drift on alleles frequencies are prominent, leading to the random fixation of deleterious variants. As a result, the impacts of genetic drift outweigh the forces of natural selection on allele frequencies, hindering the removal of deleterious mutations (Hedrick, 2001). Eventually, the increased frequency of deleterious alleles and homozygotes could reduce individual fitness, successively impeding population growth. This fitness reduction in individuals due to the fixation of detrimental variation within populations is called genetic load (Hedrick & Garcia-Dorado, 2016). Reduced population sizes increase the chance of inbreeding, resulting in the increase intensity of genetic loads. Persistently small population sizes can compound the effects of inbreeding depression, allowing opportunities for non-genetic factors, including environmental variation and demographic stochasticity, to negatively impact on population dynamics (Caro & Laurenson, 1994). Ultimately, this negative feedback, which successively erodes population size and growth rate, is referred to "mutational meltdown" a form of "extinction vortex" (Lynch, Conery, & Burger, 1995). Although how dose mutational meltdown contribute to extinction probability in the wild is still in debate, ignoring genetic factors has been suggested to hamper efficiency conservation measures (Frankham, 2005).

1.5 Inbreeding depression and extinction probability

Despite the fact that inbreeding depression universally reduces mean fitness at the population level, if these effects could be translated into elevated extinction possibility in nature populations is still controversial. It has been argued that in small populations, the impacts of genetic factors on population persistence will be secondary to ecological

factors, because ecological factors such as demographic and environmental stochasticity drive populations to extinction before the effects of genetic threats (Lande, 1988). Several researchers have used cheetah (*Acinonyx jubatus*) as an example to demonstrate that predation was the primary factor that caused the decline in population sizes instead of genetic defects arise from low genetic variation (Caro & Laurenson, 1994; Caughley, 1994). Similarly, Hoelzel and co-workers (1993) stated that the population of the Northern elephant seal (*Mirounga angustirostrus*) still prospered after experiencing a severe population bottleneck and low genetic diversity.

However, there are also conflicting studies, which indicate that genetic problems could still elevate extinction risk regardless of environmental and demographic stochasticity (O'Grady et al., 2006; Spielman, Brook, & Frankham, 2004). In 1998, Saccheri *et al.*, demonstrated that a decline in heterozygosity, an indication of inbreeding, was significantly associated with extinction risk in wild butterfly populations (*Melitaea cinxia*). Higher level of heterozygosity was corresponding to lower larval survival, adult longevity and egg-hatching rates. Brook and co-workers (2002) conducted realistic population viability analyses with 20 threatened species and concluded that inbreeding depression increased population extinction rates irrespective of initial population sizes.

Although the studies above emphasise the impacts of genetic problems on population extinction risk, majority of these studies shared some weakness and there are still some gaps in our understanding of inbreeding depression. To begin with, most studies use low genetic diversity as an indication of inbreeding except the Brook's (2002) study. Previous study has suggested the correlation between genetic diversity and inbreeding coefficient only occurs in narrow conditions (Balloux, Amos, & Coulson, 2004). Direct empirical analysis of the link between inbreeding depression and population its contribution to population demography and persistence in the wild is still rare. Additionally, the negative effects of genetic problems don't rule out the immediate threats of environmental and demographic stochasticity on population persistence. For example, Bijlsma and co-workers (2000) established an inbreeding experiment with *Drosophila melanogaste*r and detected that inbreeding depression not only increased population extinction risk but acted synergistically worse with environmental stress. Furthermore, many endangered species have been forced to inhabit sub-optimal

environments due to anthropogenic activity and habitat fragmentation (Vitousek *et al*., 1997). As a result, in order to get an improved understanding in the relationship between inbreeding depression and its contribution to population demography, we cannot ignore the impacts of environmental variation.

Our understanding of the effects of interactions between inbreeding depression and environmental factors in the wild is still poorly developed, for several reasons. First, the studies of this topic in natural populations have often suffered with low replication, yielding relatively weak statistical power. Second, accessible pedigree information in natural populations is often scarce, while the use of genetic variation as a surrogate for inbreeding typically doesn't reflect realised inbreeding levels (Balloux *et al*., 2004; Pemberton, 2004). Third, most inbreeding–environment interaction studies have been conducted in the lab or common garden, which raises questions about whether findings reflect population sustainability in the wild (Bijlsma *et al*., 2000; Kristensen, Barker, Pedersen, & Loeschcke, 2008; Reed, Briscoe, & Frankham, 2002; Springer, Messina, & Gompert, 2020; Waller, Dole, & Bersch, 2008). There are few studies that have used long-term pedigree data in wild populations to investigate the effects of environmental factors on the fitness costs of inbreeding (Keller, Grant, Grant, & Petren, 2002; Marr, Arcese, Hochachka, Reid, & Keller, 2006; Szulkin & Sheldon, 2007). However, all of these studies spanned several years and only a few environmental factors were chosen to investigate environment-dependent inbreeding depression, implying there are many environmental factors left to be studied across study years. Moreover, environments in these studies were observational not manipulated. Therefore, we can't infer causation. Additionally, all of these studies suffer from low sample sizes due to the rarity of natural inbreeding events in the wild. Lastly, three of these studies focused on bird species and only one trait was investigated in Keller and co-worker's study (2002, survival) and also in Szulkin and Sheldon's study (2007, reproduction). This implies the effects of environment-dependent inbreeding could be specific to from the characteristics of the species or population investigated.

Together, these studies indicate that we still lack a consistent comprehension of how environmental factors alter inbreeding outcomes in natural population. A full understanding of inbreeding consequences in a conservation context requires us to

disentangle intrinsic sources of heterogeneity (species characteristics, mating system, traits and demography) from extrinsic ones (including environmental factors). In the following section, we will briefly review how these extrinsic (environmental variation) and intrinsic factors (species characteristics and traits) shape the expression of inbreeding depression.

1.6 Impacts of environmental factors on inbreeding depression

Environmental variation influences inbreeding when phenotypic responses of inbred and outbred offspring vary across environmental gradient. Inbreeding by environment interaction occurs either by (1) altering phenotypic response to environments plastically without changing selection regime or (2) shifting adaptive values between inbred and outbred offspring under natural selection regime without changing phenotypes. Synergistic relationship between inbreeding depression and environmental factors are not universal. Early empirical and theoretical studies suggested that stress, defined as the fitness reduction of outbred individuals in a stressful environment, may increase the severity of inbreeding depression. In this scenario, inbreeding depression is more prominent in stressful environments since inbred and outbred offspring performed well in benign environments. In 2011, a meta-analysis with 33 studies carried out by Fox and Reed showed that the magnitude of inbreeding depression in survival scales positively with in more stressful environments. Armbruster and Reed (2005) undertook another review by comparing lethal equivalents between benign and stressful environments. Lethal equivalent indicates the extent of inbreeding depression by quantifying mutant genes that would cause death if homozygous (Frankham *et al*., 2010). Their review illustrated that inbreeding depression increased under stressful environments in 76% of studies with only 48% of studies were significant.

Despite this general pattern, a number of studies have found that environmental stress lowers the intensity of inbreeding depression. It has been reported that in song sparrow, *Melospiza melodia*, inbreeding depression decreased hatching success under longer intervals of rain (stressful environment) but decreased in response to lower temperature (stressful environment) for egg laying date (Marr *et al*., 2006). A similar effect was detected by Henry and colleagues (2003) who showed that in a freshwater snail (*Physa acuta*), inbreeding depression for survival was higher in a laboratory

environment (stable and benign) than in a stressful field environment characterised by fluctuating temperature and humidity. From these cases it is clear that the extent to which environmental conditions modify inbreeding depression is likely to depend on different types of environments or stress and on the identity of responding traits. However, we still have limited understanding what influences the heterogeneity of inbreeding depression across environmental gradient.

1.7 Species characters

The magnitude of inbreeding depression might differ based on species' characteristics, such as life form (e.g. tree, shrub or herbaceous plants), longevity or mating system (Angeloni, Ouborg, & Leimu, 2011). How these species characteristics affect inbreeding depression has been widely examined in plants. A meta-analysis conducted by Angeloni and co-workers (2011) reported that the levels of inbreeding depression were significantly higher in perennial compared with annual plants. Inbreeding depression was three times more severe in trees, and approximately twice as high in shrubs in comparison with herbaceous species. One possible explanation might be that the accumulation of mutations during the lifetime of long-lived plants leads to higher inbreeding depression (Scofield & Schultz, 2006).

The expression of inbreeding depression might also be influenced by mating systems, resulting in the interaction among life from (e.g. herbaceous plant, shrub or tree), inbreeding and mating system (Duminil, Hardy, & Petit, 2009; Lande, 1988; Morgan, 2001; Winn *et al*., 2011). Theory predicts that natural selection favours self-fertilization by optimal allocation of resources into reproduction, instead of sex function and mitigation of constraints on reproduction time in annual plants (Zhang, 2000). This could explain why self-fertilization features annual species whereas outcrossing predominates in perennial plant. Numerous empirical and theoretical studies have investigated the influence of life expectancy and mating system on inbreeding depression. In 2009, a systematic review carried out by Duminal *et al.* indicated the strong relationship between mating system and perenniality on inbreeding depression. In particular, long-lived plants had higher outcrossing rates and lower inbreeding coefficient than short-lived plants. Another genetic model predicted that annual, small-statured plants possess a variety of mating systems and exhibit relatively lower inbreeding

depression whereas perennial, large-statured plants are mainly outcrosser and display higher inbreeding depression (Scofield & Schultz, 2006). Together, these results suggest the joint effects of inbreeding depression and mating system might have profound impacts on the evolutionary dynamics of iteroparous and semelparous organisms.

1.8 Trait

Inbreeding depression has been observed in a wide variety of traits, including fitness-related traits and non-fitness-related traits (Angeloni *et al.*, 2011; Thiele, Hansen, Siegismund, & Hauser, 2010). The degree of inbreeding depression might vary among different phenotypes depending on their genetic architecture or patterns of natural selection (Merila & Sheldon, 1999; Scofield & Schultz, 2006). It is well known that traits related to fitness, such as survival or reproduction might experience more severe inbreeding depression (Falconer, 1989). The underlying reason might be that these traits harbour more dominance variance and have lower heritability, making them more vulnerable to inbreeding depression (Merila & Sheldon, 1999). Moreover, morphological traits might display lower levels of inbreeding depression in comparison with fitness-related traits (or life-history traits). This is because most morphological traits might undergo stabilizing selection for an intermediate optimum (DeRose & Roff, 1999). In these traits, recessive alleles that lead to shifts in mean phenotype in either direction would be selectively equivalent and therefore display slightly dominance (Lynch & Walsh, 1998).

Some studies tried to generalise the relationship between inbreeding depression and trait types. In a review focussing on nondomestic animals, DeRose and Roff (1999) found that inbreeding depression had more severe impacts on life-history traits, which showed an 11.8% reduction in fitness compared with a 2.2% reduction for morphological traits. A further example was shown in a meta-analysis carried out by Angeloni *et al.* (2011), which indicated significant variation of inbreeding depression in different stages of life-history traits, such as reproduction, biomass or survival. Their results reflected the potential variation of genetic structures or selection patterns for different traits, which could underpin the heterogeneity of inbreeding depression. However, both of these studies limited their scopes in either animals or plants. DeRose and Roff study applied Kruskal-Wallis test, which failed to control between studies variation.

Thus, there is a need to update the synthesis to provide a comprehensive understanding between inbreeding depression and different phenotypic traits.

1.9 Purging

Purging refers to the decline in the frequency of deleterious alleles, which are exposed by inbreeding via selection. These deleterious alleles are lost due to failure to reproduce, subsequently lowing the genetic loads. In natural populations, purging could alleviate the fitness reduction due to inbreeding and are recognised as fundamentally relevant to conservation (Hedrick & Garcia-Dorado, 2016). The efficiency of purging depends on the architecture of inbreeding load, the effects of deleterious mutations and the speeds of inbreeding (Charlesworth & Charlesworth, 1999; Hedrick & Garcia-Dorado, 2016; Keller & Waller, 2002). Inbreeding caused by overdominance will be immune to selection and purging due to heterozygote advantage, and therefore genetic loads will be maintained in populations (Keller & Waller, 2002). Additionally, mildly deleterious alleles are less efficiently purged (Garcia-Dorado, 2012). Theoretical models predict that recessive alleles with minor selective effects (i.e. s(0.5-h) < $1/2N_e$, where h is the dominance coefficient) will be effectively neutral and the effects of genetic drift will dominate any effects of natural selection (Garcia-Dorado, 2012). Lastly, purging will be more efficient when inbreeding occurs gradually over several generations. Rapid and intense inbreeding drastically reduces N_e, subsequently leading to the random fixation of genetic load (Keller & Waller, 2002). As a result, the Hill-Robertson effect, which describes the associations among alleles at different loci in small populations, will be more prevalent, which hampers the removal of deleterious alleles (Hill & Robertson, 2007).

Purging has been proposed as a viable method to reverse the harmful effects of inbreeding depression in managed populations. In 1983, Templeton and Read devised a careful, controlled breeding scheme to eradicate inbreeding depression in an endangered species, Speke's gazelle (*Gazella spekei*) (Schonewaldcox, Chambers, MacBryde, & Thomas, 1984). They claimed success in their purging programme after they detected that the second and third generations of inbred Speke's gazelles had higher survival and viability than the first generation of inbred offspring. However, subsequent investigation showed that the fitness improvement of second and third

generations of Speke's gazelles was due to improvements in husbandry, rather than the purging itself (Kalinowski, Hedrick, & Miller, 2000). This case emphasises the difficulty in correctly attributing phenotypic and fitness changes to inbreeding depression and its interaction with environmental factors. Despite the failure of Speke's gazelle's programme of genetic load purging, Templeton and Read's study has fuelled interests in purging as a conservation intervention for captive populations.

Most of studies on purging have centred on laboratory populations. Empirical evidence for purging in natural populations remains scare due to the limitation of typical field-based surveys to detect and attribute genetic/ phenotypic changes to purging (Bryant, Meffert, & McCommas, 1990; Fowler & Whitlock, 1999; Saccheri *et al.*, 1998). A meta-analysis conducted by Byers and Waller (1999) identified purging in 4 studies out of 11 that were reviewed. Another valuable study that utilised a 20-year pedigree of the Chatham Island black robin (*Petroica traversi*) in the wild documented the effects of purging of close breeding between highly-inbred parents, whereas inbreeding depression still reduced fledging survival for close breeding of less-inbred parents (Weiser, Grueber, Kennedy, & Jamieson, 2016). Together, these studies suggest purging occurs in the wild but is not a universal phenomenon. Deliberate purging to mitigate the effects of inbreeding depression in captive populations is still risky, for the following reasons. First, purging is always accompanied with an early fitness reduction before benefits to fitness in later generations takes place (Hedrick & Garcia-Dorado, 2016). Additionally, continued inbreeding might eliminate lethal mutations but deleterious alleles with mild effects could still be fixed permanently, subsequently lowering population fitness (Wang, Hill, Charlesworth, & Charlesworth, 1999). Therefore, it is questionable if purging is a viable conservation intervention. *Ex-situ* conservation (genetic rescue) might still be necessary for extremely endangered species.

1.10 Genetic Rescue

When populations remain small and isolated, the effects of genetic drift dominate natural selection, resulting in the increase of drift load by random fixation of deleterious alleles. Gene flow can counter the development of drift load, enabling isolated populations behave as a single large population (Frankham *et al*., 2010). Additionally, fitness reduction resulting from inbreeding depression can be mitigated by the

introduction of individuals from separate subpopulations, a conservation intervention called genetic rescue (Frankham, 2015). Genetic rescue recovers individual fitness by creating "heterosis" either through the regeneration of heterozygote advantage (over-dominance) or by masking deleterious alleles in the heterozygous state (Whitlock, Ingvarsson, & Hatfield, 2000). Translocation of non-local individuals into isolated, inbred populations has been carried out effectively to restore individual fitness in endangered wild populations.

For example, the remnant population of Florida panther (*Felis concolor coryi*) had suffered from precipitously declining population size, severe low genetic diversity and genetic health problems (e.g., spermatozoa1 defects, cryptorchidism and cardiac abnormalities). These latter health defects were due to the accumulation and expression of deleterious alleles caused by inbreeding (Roelke, Martenson, & Obrien, 1993). Minisatellite genetic data indicated that the Florida panther had 85% less genetic variation compared with related western pumas (*F. c. coryi*). In 1995, genetic rescue was attempted in order to reverse inbreeding depression. After the introduction of female panthers (*Puma concolor stanleyana*) from Texas, the frequency of inbreeding-related deleterious traits in the Florida panther population declined drastically and population size had tripled by the year 2002 (Johnson *et al.,* 2010).

A further example of genetic rescue comes from an isolated population of adders (*Vipera berus*) in southern Sweden (Madsen, Stille, & Shine, 1996). This population had declined in size since 1992 as a result of habitat fragmentation and habitat destruction (Madsen *et al.,* 1996). The resultant inbreeding increased the accumulation of detrimental phenotypes (e.g., smaller litter size and stillborn offspring) and reduced genetic variation, threatening population sustainability. The average effective population sizes had dwindled to 11.6 and 4.42 for male and female respectively during 1984 to 1990. In 1992, after 20 immigrant male adders were translocated from a northern Swedish population, the adder population size bounced back to its pre-inbreeding level (Madsen, Shine, Olsson, & Wittzell, 1999).

In 2015, Frankham conducted a meta-analysis and provided a comprehensive overview, demonstrating the consistent beneficial effects of genetic rescue in wild populations. In his study, he detected beneficial effects of outcrossing between inbred populations in 95.9% of 156 cases, with little variation in the success of genetic rescue among different taxonomy groups (Frankham, 2015).

1.11 The opposite effects of genetic rescue: outbreeding depression

However, inter-mating between individuals from isolated populations in order to reverse inbreeding depression is still relatively infrequently undertaken, due to concerns over outbreeding depression and the loss of local adaptation. Outbreeding depression refers to a decline in fitness of hybrid offspring arising from crossing between populations (Whitlock *et al*., 2013). The extent of outbreeding depression is thought to be related to the genetic divergence between immigrant and local populations (Tallmon, Luikart, & Waples, 2004). During long term isolation, a subset of alleles at different loci might evolve to work well together, allowing intrinsic co-adaption to the current environment (Orr, 1995). On outcrossing between different subpopulations, these co-adapted gene complexes could break down, reducing fitness, and resulting in outbreeding depression (Frankham *et al*., 2011). In their systematic review and meta-analysis, Whitlock and colleagues (2013) reported consistent costs to outbreeding in later generation crosses, despite heterogeneity in outbreeding effects among studies. Frankham and colleagues (2011) constructed a decision tree and generalised that populations are more likely to suffer from outbreeding depression if they have fixed chromosome difference or did not exchange genes for more than 500 years or inhabit different habitats. Understanding the costs of inbreeding depression is still a major concern in conservation biology and placing them in the context of threats from outbreeding depression enhances our knowledge when and genetic rescue might be appropriate.

1.12 Study system

In order to understand the effects of ongoing climate changes on population persistence in small and isolated populations via fitness reduction of inbreeding, there is a need to synthesise empirical studies on phenotypic outcomes of inbreeding in related to environmental shifts. A comprehensive, evidence-based literature database (a systematic

map of the literature) was assembled by Neaves and co-workers in 2015, comprising empirical studies of the phenotypic and fitness consequences of inbreeding in natural populations. The objective, rigorous assessment of the scope and quality of inbreeding studies was conducted in full-text scale, providing informative searchable map for conservation biologist. Articles within this systematic map are research papers retrieved from three online databases (Web of sciences, Scopus and JSTOR). The search strings were inbreeding and fitness-related search strings identified by references cited in review papers (appendix 1) and refined by subject-based experts. The systematic map comprises of 703 individual articles, reporting inbreeding depression in a variety of taxonomic groups, 66.7 % of which were plants, 7.3 % insects and 6.1 % birds (Neaves *et al*., 2015). The majority of studies in the systematic map investigated the phenotypic consequences of inbreeding on fitness components (e.g., fecundity, survival or viability; 92.2 %) and were undertaken in laboratory or greenhouse environments (69.1 %). The factors relating to potential sources of heterogeneity in the outcomes of inbreeding (e.g. inbreeding coefficient, mating system, population size, life history traits) were also recorded in this database (Neaves *et al*., 2015). This systematic map not only highlights gaps in our understanding of the consequences of inbreeding in natural populations but also provides a valuable guideline for future researchers who seek to synthesise the impacts of inbreeding depression in a conservation context by meta-analysis.

Through this thesis, I have updated Neaves' *et al*. (2015) systematic map, using the same literature search strings, and have utilised this updated systematic map resource to carry out a series of meta-analyses into the fitness and phenotypic consequence of inbreeding. The goal is to disentangle the causes and consequence of inbreeding depression in populations of animals and plants. Lastly, I wish to generalise the relationship between inbreeding depression to environmental variation, trait types and species characteristics. Through this work, I aim to provide insights into the potential solutions for conservation practitioners to prevent the extinction of small populations via inbreeding in face of global environmental changes.

1.13 This thesis

A key challenge in understanding the role of inbreeding depression in driving the extinction risks of threatened populations and species is unpick how environmental variation shapes responses to inbreeding in different demographic contexts. In this thesis, I will conduct Bayesian meta-analysis to synthesise and evaluate the evidence-base, facilitating a general view of the interaction between inbreeding depression and environmental and demographic characters. I will use meta-analysis to overcome low statistical power that is often present in individual papers, giving a general view, and to provide quantitative synthesis of the empirical studies. Ultimately, this work aims to provide guidelines for conservation management regarding genetic factors and environmental variation.

In chapter 2, I developed and tested an automated process to quickly identify articles relevant to inbreeding depression from the systematic map database. Automated content analysis (ACA) is applied via the structural topic model (stm) package from R, to analyse texts and identify the topics in our literature database. I evaluate the specificity and sensitivity of ACA in relation to manual literature assessment. How ACA could be applied to reduce workload and accelerate screening relevant articles for systematic reviews was also discussed.

In chapter 3, a meta-analysis is performed to understand the relationship between inbreeding depression and species/ demographic characteristics. Several characteristics have been identified to cause variation in inbreeding depression, including types of traits, mating system, life expectancy and taxonomic group and are incorporate in this study. The results of this meta-analysis update our understanding of the factors creating heterogeneity in inbreeding depression and improve the evidence base informing conservation.

In chapter 4, I assess, using meta-analysis, how abiotic and biotic environmental conditions influence fitness costs associated with inbreeding. The levels of stress experienced by organisms are also quantified to evaluate their association with inbreeding depression. Here, I focus on changes to inbreeding depression stemming from environmental changes in phenotypic expression and dominance. I assume that

environmental factors and stress intensity could be used to predict the severity of inbreeding depression in natural populations.

Chapter 5 synthesises the findings to develop an integrated understanding of how inbreeding, environmental conditions, stress and contextual factors interact to shape inbreeding depression. Furthermore, I discussed the risks and effectiveness of conservation measures utilised to mitigate the adverse effects of inbreeding depression and how these measures will be influenced by future climate changes. The goal is to provide suggestions on how current conservation measures could be improved and applied to recover fitness and genetic diversity from inbreeding depression under a changing world.

1.14 References

Angeloni, F., Ouborg, N. J., & Leimu, R. (2011). Meta-analysis on the association of population size and life history with inbreeding depression in plants. *Biological Conservation, 144*(1), 35-43. doi:10.1016/j.biocon.2010.08.016

Armbruster, P., & Reed, D. H. (2005). Inbreeding depression in benign and stressful environments. *Heredity, 95*(3), 235-242. doi:10.1038/sj.hdy.6800721

Balloux, F., Amos, W., & Coulson, T. (2004). Does heterozygosity estimate inbreeding in real populations? *Molecular Ecology, 13*(10), 3021-3031. doi:10.1111/j.1365-294X.2004.02318.x

Bijlsma, R., Bundgaard, J., & Boerema, A. C. (2000). Does inbreeding affect the extinction risk of small populations? predictions from Drosophila. *Journal of Evolutionary Biology, 13*(3), 502-514. Retrieved from <Go to ISI>://WOS:000087030000015

Brook, B. W., Tonkyn, D. W., Q'Grady, J. J., & Frankham, R. (2002). Contribution of inbreeding to extinction risk in threatened species. *Conservation Ecology, 6*(1). Retrieved from <Go to ISI>://WOS:000177892600007

Bryant, E. H., Meffert, L. M., & McCommas, S. A. (1990). Fitness rebound in serially bottlenecked populations of the house-fly. *American Naturalist, 136*(4), 542-549. doi:10.1086/285112

Byers, D. L., & Waller, D. M. (1999). Do plant populations purge their genetic load? Effects of population size and mating history on inbreeding depression. *Annual Review of Ecology and Systematics, 30*, 479-513. doi:10.1146/annurev.ecolsys.30.1.479

Caro, T. M., & Laurenson, M. K. (1994). Ecological and genetic-factors in conservation - a cautionary tale. *Science, 263*(5146), 485-486. doi:10.1126/science.8290956

Caughley, G. (1994). Directions in conservation biology. *Journal of Animal Ecology, 63*(2), 215-244. doi:10.2307/5542

Charlesworth, B., & Charlesworth, D. (1999). The genetic basis of inbreeding depression. *Genetical Research, 74*(3), 329-340. doi:10.1017/s0016672399004152

Charlesworth., & Willis, J. H. (2009). Fundamental concepts in genetics the genetics of inbreeding depression. *Nature Reviews Genetics, 10*(11), 783-796. doi:10.1038/nrg2664

Cheptou, P. O., & Donohue, K. (2011). Environment-dependent inbreeding depression: its ecological and evolutionary significance. *New Phytologist, 189*(2), 395-407. doi:10.1111/j.1469-8137.2010.03541.x

Crnokrak, P., & Roff, D. A. (1999). Inbreeding depression in the wild. *Heredity, 83*, 260-270. doi:10.1038/sj.hdy.6885530

Cutter, A. D. (2019). *A primer of molecular population genetics* (1 st ed.): Oxford University Press.

Darwin, C. (1900). *The effects of cross and self fertilisation in the vegetable kingdom* (2d. ed. 5th impression ed.). London :: J. Murray.

DeRose, M. A., & Roff, D. A. (1999). A comparison of inbreeding depression in life-history and morphological traits in animals. *Evolution, 53*(4), 1288-1292. doi:10.2307/2640831

Duminil, J., Hardy, O. J., & Petit, R. J. (2009). Plant traits correlated with generation time directly affect inbreeding depression and mating system and indirectly genetic structure. *Bmc Evolutionary Biology, 9*. doi:10.1186/1471-2148-9-177

Falconer, D. S. (1989). *An introduction to quantitative genetics* (3 rd ed.): Wiley.

Fowler, K., & Whitlock, M. C. (1999). The variance in inbreeding depression and the recovery of fitness in bottlenecked populations. *Proceedings of the Royal Society B-Biological Sciences, 266*(1433), 2061-2066. doi:10.1098/rspb.1999.0887

Fox, C. W., & Reed, D. H. (2011). Inbreeding depression increases with environmental stress: an experimental study and meta-analysis. *Evolution, 65*(1), 246-258. doi:10.1111/j.1558-5646.2010.01108.x

Frankham, R. (2005). Genetics and extinction. *Biological Conservation, 126*(2), 131-140. doi:10.1016/j.biocon.2005.05.002

Frankham, R. (2015). Genetic rescue of small inbred populations: meta-analysis reveals large and consistent benefits of gene flow. *Molecular Ecology, 24*(11), 2610-2618. doi:10.1111/mec.13139

Frankham, R., Ballou, J. D., & Briscoe, D. A. (2010). *Introduction to conservation genetics* (2 nd ed.): Cambridge University press.

Frankham, R., Ballou, J. D., Eldridge, M. D. B., Lacy, R. C., Ralls, K., Dudash, M. R., & Fenster, C. B. (2011). Predicting the probability of outbreeding depression. *Conservation Biology, 25*(3), 465-475. doi:10.1111/j.1523-1739.2011.01662.x

Frankham, R., Lees, K., Montgomery, M. E., England, P. R., Lowe, E. H., & Briscoe, D. A. (1999). Do population size bottlenecks reduce evolutionary potential? *Animal Conservation, 2*(4), 255-260. doi:10.1111/j.1469-1795.1999.tb00071.x

Garcia-Dorado, A. (2012). Understanding and predicting the fitness decline of shrunk populations: inbreeding, purging, mutation, and standard selection. *Genetics, 190*(4), 1461-1476. doi:10.1534/genetics.111.135541

Grant, P. R., & Grant, B. R. (2006). Evolution of Character Displacement in Darwin's Finches. *Science, 313*, 224 - 226.
Hedrick, P. W. (2001). Conservation genetics: where are we now? *Trends in Ecology & Evolution, 16*(11), 629-636. doi:10.1016/s0169-5347(01)02282-0
Hedrick, P. W., & Garcia-Dorado, A. (2016). Understanding inbreeding depression, purging, and genetic rescue. *Trends in Ecology & Evolution, 31*(12), 940-952. doi:10.1016/j.tree.2016.09.005
Henry, P. Y., Pradel, R., & Jarne, P. (2003). Environment-dependent inbreeding depression in a hermaphroditic freshwater snail. *Journal of Evolutionary Biology, 16*(6), 1211-1222. doi:10.1046/j.1420-9101.2003.00629.x
Hill, W. G., & Robertson, A. (2007). The effect of linkage on limits to artificial selection (Reprinted). *Genetics Research, 89*(5-6), 311-336. doi:10.1017/s001667230800949x
Hoelzel, A. R., Halley, J., Obrien, S. J., Campagna, C., Arnbom, T., Leboeuf, B., . . . Dover, G. A. (1993). Elephant seal genetic-variation and the use of simulation-models to investigate historical population bottlenecks. *Journal of Heredity, 84*(6), 443-449. doi:10.1093/oxfordjournals.jhered.a111370
Hoffmann, A. A., Hallas, R. J., Dean, J. A., & Schiffer, M. (2003). Low potential for climatic stress adaptation in a rainforest Drosophila species. *Science, 301*(5629), 100-102. doi:10.1126/science.1084296
Hoffmann, A. A., & Sgro, C. M. (2011). Climate change and evolutionary adaptation. *Nature, 470*(7335), 479-485. doi:10.1038/nature09670
Hoffmann, A. A., & Willi, Y. (2008). Detecting genetic responses to environmental change. *Nature Reviews Genetics, 9*(6), 421-432. doi:10.1038/nrg2339
IPCC. (2022). *Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Retrieved from Cambridge, UK
New York, NY, USA:
Johnson, W. E., Onorato, D. P., Roelke, M. E., Land, E. D., Cunningham, M., Belden, R. C., . . . O'Brien, S. J. (2010). Genetic restoration of the florida panther. *Science, 329*(5999), 1641-1645. doi:10.1126/science.1192891
Kalinowski, S. T., Hedrick, P. W., & Miller, P. S. (2000). Inbreeding depression in the Speke's gazelle captive breeding program. *Conservation Biology, 14*(5), 1375-1384. doi:10.1046/j.1523-1739.2000.98209.x
Keller, L. F., Grant, P. R., Grant, B. R., & Petren, K. (2002). Environmental conditions affect the magnitude of inbreeding depression in survival of Darwin's finches. *Evolution, 56*(6), 1229-1239. doi:10.1111/j.0014-3820.2002.tb01434.x
Keller, L. F., & Waller, D. M. (2002). Inbreeding effects in wild populations. *Trends in Ecology & Evolution, 17*(5), 230-241. doi:10.1016/s0169-5347(02)02489-8
Kondrashov, A. S., & Houle, D. (1994). Genotype-environment interactions and the estimation of the genomic mutation-rate in Drosophila-melanogaster. *Proceedings of the Royal Society B-Biological Sciences, 258*(1353), 221-227. doi:10.1098/rspb.1994.0166
Kristensen, T. N., Barker, J. S. F., Pedersen, K. S., & Loeschcke, V. (2008). Extreme temperatures increase the deleterious consequences of inbreeding under laboratory and semi-natural conditions. *Proceedings of the Royal Society B-*

Biological Sciences, 275(1646), 2055-2061. doi:10.1098/rspb.2008.0426
Lande, R. (1988). Genetics and demography in biological conservation. *Science, 241*(4872), 1455-1460. doi:10.1126/science.3420403
Li, Y.-C., Fahima, T., Krugman, T., Beiles, A., Röder, M. S., Korol, A. B., & Nevo, E. (2000). Parallel microgeographic patterns of genetic diversity and divergence revealed by allozyme, RAPD, and microsatellites in Triticum dicoccoides at Ammiad, Israel. *Conservation Genetics, 1*(3), 191-207. doi:10.1023/A:1011545403198
Li, Y. C., Fahima, T., Beiles, A., Korol, A. B., & Nevo, E. (1999). Microclimatic stress and adaptive DNA differentiation in wild emmer wheat, Triticum dicoccoides. *Theoretical and Applied Genetics, 98*(6-7), 873-883. doi:10.1007/s001220051146
Lush, J. L. (1943). *Animal breeding plans*: Read Books Ltd.
Lynch, M., Conery, J., & Burger, R. (1995). Mutational meltdowns in sexual populations. *Evolution, 49*(6), 1067-1080. doi:10.1111/j.1558-5646.1995.tb04434.x
Lynch, M., & Walsh, B. (1998). *Genetics and Analysis of Quantitative Traits*: Sinauer.
Madsen, T., Shine, R., Olsson, M., & Wittzell, H. (1999). Conservation biology - Restoration of an inbred adder population. *Nature, 402*(6757), 34-35. doi:10.1038/46941
Madsen, T., Stille, B., & Shine, R. (1996). Inbreeding depression in an isolated population of adders Vipera berus. *Biological Conservation, 75*(2), 113-118. doi:10.1016/0006-3207(95)00067-4
Marr, A. B., Arcese, P., Hochachka, W. M., Reid, J. M., & Keller, L. F. (2006). Interactive effects of environmental stress and inbreeding on reproductive traits in a wild bird population. *Journal of Animal Ecology, 75*(6), 1406-1415. doi:10.1111/j.1365-2656.2006.01165.x
Merila, J., & Sheldon, B. C. (1999). Genetic architecture of fitness and nonfitness traits: empirical patterns and development of ideas. *Heredity, 83*, 103-109. doi:10.1046/j.1365-2540.1999.00585.x
Morgan, M. T. (2001). Consequences of life history for inbreeding depression and mating system evolution in plants. *Proceedings of the Royal Society B-Biological Sciences, 268*(1478), 1817-1824. doi:10.1098/rspb.2001.1741
Neaves, L. E., Eales, J., Whitlock, R., Hollingsworth, P. M., Burke, T., & Pullin, A. S. (2015). The fitness consequences of inbreeding in natural populations and their implications for species conservation - a systematic map. *Environmental Evidence, 4*(1). doi:10.1186/s13750-015-0031-x
Norberg, J., Urban, M. C., Vellend, M., Klausmeier, C. A., & Loeuille, N. (2012). Eco-evolutionary responses of biodiversity to climate change. *Nature Climate Change, 2*(10), 747-751. doi:10.1038/nclimate1588
O'Grady, J. J., Brook, B. W., Reed, D. H., Ballou, J. D., Tonkyn, D. W., & Frankham, R. (2006). Realistic levels of inbreeding depression strongly affect extinction risk in wild populations. *Biological Conservation, 133*(1), 42-51. doi:10.1016/j.biocon.2006.05.016
Orr, H. A. (1995). The population-genetics of speciation - the evolution of hybrid incompatibilities. *Genetics, 139*(4), 1805-1813. Retrieved from <Go to ISI>://WOS:A1995QN97500029

Pemberton, J. (2004). Measuring inbreeding depression in the wild: the old ways are the best. *Trends in Ecology & Evolution, 19*(12), 613-615. doi:10.1016/j.tree.2004.09.010

Potvin, C., & Tousignant, D. (1996). Evolutionary consequences of simulated global change: Genetic adaptation or adaptive phenotypic plasticity. *Oecologia, 108*(4), 683-693. doi:10.1007/bf00329043

Reale, D., McAdam, A. G., Boutin, S., & Berteaux, D. (2003). Genetic and plastic responses of a northern mammal to climate change. *Proceedings of the Royal Society B-Biological Sciences, 270*(1515), 591-596. doi:10.1098/rspb.2002.2224

Reed, D. H., Briscoe, D. A., & Frankham, R. (2002). Inbreeding and extinction: The effect of environmental stress and lineage. *Conservation Genetics, 3*(3), 301-307. doi:10.1023/a:1019948130263

Roelke, M. E., Martenson, J. S., & Obrien, S. J. (1993). The consequences of demographic reduction and genetic depletion in the endangered florida panther. *Current Biology, 3*(6), 340-350. doi:10.1016/0960-9822(93)90197-v

Saccheri, I., Kuussaari, M., Kankare, M., Vikman, P., Fortelius, W., & Hanski, I. (1998). Inbreeding and extinction in a butterfly metapopulation. *Nature, 392*(6675), 491-494. doi:10.1038/33136

Schonewaldcox, C. M., Chambers, S. M., MacBryde, B., & Thomas, L. (1984). Genetic and conservation - a reference for managing wild animal and plant-populations. *Benjamin/Cummings Publishing, 34*(11), 719-719. doi:10.2307/1309668

Scofield, D. G., & Schultz, S. T. (2006). Mitosis, stature and evolution of plant mating systems: low-Phi and high-Phi plants. *Proceedings of the Royal Society B-Biological Sciences, 273*(1584), 275-282. doi:10.1098/rspb.2005.3304

Sinervo, B., Mendez-de-la-Cruz, F., Miles, D. B., Heulin, B., Bastiaans, E., Cruz, M. V. S., . . . Sites, J. W. (2010). Erosion of Lizard Diversity by Climate Change and Altered Thermal Niches. *Science, 328*(5980), 894-899. doi:10.1126/science.1184695

Spielman, D., Brook, B. W., & Frankham, R. (2004). Most species are not driven to extinction before genetic factors impact them. *Proceedings of the National Academy of Sciences of the United States of America, 101*(42), 15261-15264. doi:10.1073/pnas.0403809101

Springer, A. L., Messina, F. J., & Gompert, Z. (2020). Measuring the effect of environmental stress on inbreeding depression alone obscures the relative importance of inbreeding-stress interactions on overall fitness inCallosobruchus maculatus. *Evolutionary Applications, 13*(10), 2597-2609. doi:10.1111/eva.13060

Szulkin, M., & Sheldon, B. C. (2007). The Environmental Dependence of Inbreeding Depression in a Wild Bird Population. *Plos One, 2*(10). doi:10.1371/journal.pone.0001027

Tallmon, D. A., Luikart, G., & Waples, R. S. (2004). The alluring simplicity and complex reality of genetic rescue. *Trends in Ecology & Evolution, 19*(9), 489-496. doi:10.1016/j.tree.2004.07.003

Thiele, J., Hansen, T., Siegismund, H. R., & Hauser, T. P. (2010). Genetic variation of inbreeding depression among floral and fitness traits in Silene nutans.

Heredity, 104(1), 52-60. doi:10.1038/hdy.2009.103
Vitousek, P. M., Mooney, H. A., Lubchenco, J., & Melillo, J. M. (1997). Human domination of Earth's ecosystems. *Science, 277*(5325), 494-499. doi:10.1126/science.277.5325.494
Waller, D. M., Dole, J., & Bersch, A. J. (2008). Effects of stress and phenotypic variation on inbreeding depression in Brassica rapa. *Evolution, 62*(4), 917-931. doi:10.1111/j.1558-5646.2008.00325.x
Walsh, B., & Blows, M. W. (2009). Abundant Genetic Variation plus Strong Selection = Multivariate Genetic Constraints: A Geometric View of Adaptation. *Annual Review of Ecology Evolution and Systematics, 40*, 41-59. doi:10.1146/annurev.ecolsys.110308.120232
Wang, J. L., Hill, W. G., Charlesworth, D., & Charlesworth, B. (1999). Dynamics of inbreeding depression due to deleterious mutations in small populations: mutation parameters and inbreeding rate. *Genetics Research, 74*(2), 165-178. doi:10.1017/s0016672399003900
Weiser, E. L., Grueber, C. E., Kennedy, E. S., & Jamieson, I. G. (2016). Unexpected positive and negative effects of continuing inbreeding in one of the world's most inbred wild animals. *Evolution, 70*(1), 154-166. doi:10.1111/evo.12840
Whitlock, R., Stewart, G. B., Goodman, S. J., Piertney, S. B., Butlin, R. K., Pullin, A. S., & Burke, T. (2013). A systematic review of phenotypic responses to between-population outbreeding. *Environmental Evidence, 2*(1), 13. doi:10.1186/2047-2382-2-13
Whitlock, M. C., Ingvarsson, P. K., & Hatfield, T. (2000). Local drift load and the heterosis of interconnected populations. *Heredity, 84*(4), 452-457. doi:10.1046/j.1365-2540.2000.00693.x
Winn, A. A., Elle, E., Kalisz, S., Cheptou, P. O., Eckert, C. G., Goodwillie, C., . . . Vallejo-Marin, M. (2011). Analysis of inbreeding depression in mixed-mating plants provides evidence for selective interference and stable mixed mating. *Evolution, 65*(12), 3339-3359. doi:10.1111/j.1558-5646.2011.01462.x
Wright, S. (1922). Coefficients of inbreeding and relationship. *American Naturalist, 56*, 330-338. doi:10.1086/279872
Zhang, D. Y. (2000). Resource allocation and the evolution of self-fertilization in plants. *American Naturalist, 155*(2), 187-199. doi:10.1086/303310

Chapter 2: Assessment of automated content analysis for literature evaluation during meta-analysis

2.1 Abstract

In the digitised era, the explosive velocity with which literature is being published has made literature synthesis labour-intensive and time consuming. Traditional literature synthesis requires researchers to screen thousands of articles in order to identify relevant studies for synthesis. Automated content analysis (ACA), a text-mining approach offers a potential solution to accelerate this process, by classifying the underlying topics for a large text corpus. In this study, I applied ACA to a large database of articles relating to inbreeding within animal and plant species. My goal was to construct and evaluate an end-to-end pipeline to retrieve relevant articles from a literature corpus, for systematic review. ACA was used to describe the structure of this literature corpus. Subsequently, linear discriminant analysis (LDA) was performed on ACA topic scores for each article to test the predictive power of ACA for assessing article relevance, versus manual assessment, using a training-set test-set approach. I evaluated the efficiency and viability of this ACA-LDA article evaluation process for facilitating assessment of article relevance for meta-analysis. I identified 20 topics by ACA within the database. 6 of these topics were relevant to my meta-analysis whereas 11 of these topics were irrelevant. In addition, I detected a trade-off between sensitivity (detecting relevant articles) and efficiency (workload reduction) when applying different LD1 thresholds with ACA-LDA prediction. The best of the ACA-LDA result (-0.6 LD1) indicated 45% of the workload reduction with the risk of losing 6 % of relevant articles. I conclude that the application of ACA-LDA process on screening literature is promising with -0.6 LD1 on review update. However, the lack of replication and assessment of ACA-LDA process on different database make it difficult to draw a conclusion on whether -0.6 LD1 is a suitable threshold universally. Furthermore, how large should be the training set if we wish to apply this ACA-LDA process on a brand review is still unsolved.

2.2 Introduction

Systematic reviews assemble data from published and unpublished resources, synthesising existing evidence to offer comprehensive answers to formulated scientific

questions (Stewart, Coles, & Pullin, 2005). It typically follows a strict guideline and statistical protocol to minimise bias and enhance transparency and reliability (Stewart *et al*., 2005). It is pivotal for the progression of ecology and evolutionary biology as it allows us to validate and generalise concepts and findings from the primary literature facilitating the development of general theories (Arnqvist & Wooster, 1995). However, the information age has been marked by the exponential growth of scientific literature due to the increasing volume and speed of publication. This phenomenon, called "big literature", has posed a great challenge to meta-analysis and systematic review. Specifically, researchers are required to read thousands of titles and abstracts of irrelevant articles, a process called screening, in order to retrieve a small proportion of relevant articles for their reviews. Since the rapid growth in scientific corpora has not been paralleled by an increase in the available resources available to human reviewers, it is extremely difficult for a single researcher to keep up with ever-expanding literature data (Stockwell, Colomb, Smith, & Wiles, 2009). Therefore, the need exists to develop an efficient method to classify and distil the growing body of scientific information for systematic review.

Application of algorithms for analysing text from digital sources has given researchers a mean to process the ever-increasing scientific literature in a data-driven, objective way. Automated Content Analysis (ACA) refers to the statistical algorithm program that applies "topic models" or "concept mapping modes" to detect the thematic composition of text data (Blei, 2012; Nunez-Mir, Iannone, Pijanowski, Kong, & Fei, 2016). It is a type of text-mining approaches that utilises text-parsing and machine learning to identify topics and their relationships based on word frequency and exclusivity within a body of literature. In the context of the ACA, topic is defined as a mixture of words and each word has its own probability within a topic. The fundamental process of the ACA is accomplished through topic identification, topic definition and text classification (Nunez-Mir *et al*., 2016). Topics are identified as focal semantic subjects based on word co-occurrence and correlation in a text corpus that have similar concepts (Nunez-Mir *et al*., 2016). Next, topic definition is achieved by compiling a thesaurus with the ACA build-in function, concept mapping algorithm (Nunez-Mir *et al*., 2016). During the last step, text classification, each part of text corpus (sentences or line segments) is analysed for evidence of the occurrence of the identified topics.

Finally, these topics are then used as categories to classify the literature on the basis of different research themes or interests (Nunez-Mir *et al*., 2016).

In summary, a critical feature of ACA is the ability to discover the information content and structure of text data (Ananiadou, McNaught, & Karamanis, 2005; Hearst, 1999). The applications of ACA in the fields of ecology and evolutionary biology are still in their infancy. Most of studies have utilised ACA descriptively to explore trends of scientific publications within a particular field over a period, or to discover the underlying topics in textual data from online database. For example. Altaweel *et al*., (2019) applied ACA on government documents to investigate policy responses to mountain pine beetle outbreaks in the United stated. Their finding identified an increase in literature in related to warmer temperature recently. McCallen *et al*.,(2019) employed ACA to explore the trends of topics in ecological journals during the past four decades. Their results suggested a decline in theoretical research and an increase in data-intensive research over the past four decades. Nunez-Mir *et al*., operated ACA on 15 forestry journals with 29766 abstracts to identify the knowledge gaps and trends within restoration ecology (Nunez-Mir, Iannone, Curtis, & Fei, 2015).

The application of ACA for screening relevant articles for meta-analysis has not yet been investigated systematically in the field of ecology. However, using text-mining approach to facilitate article screening for systematic review has been applied in the field of biomedicine. A review study conducted by O'Mara-Eves and co-workers (2015) evaluated the potential application of text-mining for systematic review in biomedicine and suggested a possible workload reduction of between 30% to 70% during literature assessment. The workload reduction in this approach is typically referred to the proportion of articles that doesn't need to be assessed. Typically, there are two approaches for applying text mining to screen for relevant articles for systematic reviews from enormous database. The first approach prioritises (or sorts) the list of articles for manual assessment, so that relevant articles are more likely to be encountered early in manual article assessment, increasing efficiency (Thomas, McNaught, & Ananiadou, 2011). The second approach applies the use of manual assessment as a training set for automatic classification tools (Thomas *et al*., 2011). Reviewers will be able to automatic apply exclusion/inclusion criteria on literature database for

systematic reviews. However, the following question remains unclear: (1) How efficient is ACA to facilitate article screening for systematic review? (2) How can ACA be adapted to the field of ecology and evolutionary biology?

In this study, I used an ACA approach called structural topic modelling (stm) package in R developed by Roberts *et al.,* (2019), to assist in the description and classification of a large database containing literature on inbreeding in animals and plants. A prototype of the literature database was constructed by Neaves *et al*., in 2015. I have expanded and updated this database comprehensively by extracting articles from online databases using the same key words. My main objectives are (1) to examine the specificity and sensitivity of ACA for identifying by applying stm to an inbreeding literature corpus (2) and to utilise stm models to test viable methods to accelerate article assessment in meta-analysis. Ultimately, my goal is to develop an end-to-end pipeline method to efficiently retrieve relevant articles for meta-analysis in the era of "big literature" era.

2.3 Methods

2.3.1 Data

Before describing ACA deployed, a brief summary of the database is provided. In 2015, Neaves and co-workers did a literature search aiming to construct a database synthesising evidence-based articles describing phenotypic outcomes of inbreeding depression in natural populations. This database was built with systematic mapping techniques – an approach aim to collate, describe and classify the literature according to topics or questions of interests – (James, Randall, & Haddaway, 2016). The search strings (Table 1) identified by references cited in review papers (Appendix 1) and refined by subject-based experts from the review team were used to retrieved inbreeding-relevant articles from three online databases (Web of sciences, Scopus and JSTOR). The retrieved articles were assessed on subject-based exclusion/inclusion criteria (see 2.2.3 Exclusion/ Inclusion criteria) to compile a searchable map. The systematic map consists of 703 distinct inbreeding-relevant articles, with a majority of studies reporting inbreeding depression in plants (469 articles), followed by insects (52 articles) and birds (43 articles). Contextual information that was expected to be related to the

outcomes of inbreeding depression (e.g., inbreeding coefficient, mating system, population size, life history traits) was also recorded in this database. The extent and quality of evidence for inbreeding depression was assessed for each article by reading the article full-text, providing an informative searchable map for conservation biologists interested in the consequences of inbreeding. The master library was Neaves' original literature search record, containing all the original articles retrieved from the three online databases. It's the precursor of the systematic map, documenting the relevance/ irrelevance of articles assessed by Neaves and her colleagues. This master library is used in this chapter as a training data set in the ACA analyses described later.

Since the Neaves *et al*. (2015) database only included articles before 12th August 2013, I carried out a literature search and assessment update from 4th April 2013 to 2nd May 2018. I used identical search strings from the same online databases, in order to include inbreeding-related articles after 2013.

Table 1. Search strings used by Neaves *et al*. (2015) to retrieve articles from online database.

Group	Search string
Inbreeding related strings	"In*breeding coefficient$" "Cost$ of in*breeding" (inbred SAME mating*) NOT (("Quantitative trait loc*")OR(QTL*)) (inbred SAME (offspring OR progeny)) Selfed SAME out* "Optimal outcrossing" OR "Outcrossing distance" "Benefit* of dispersal" "Cost* of dispersal" ("Natal dispersal" AND (inbred OR in*breeding OR heterosis OR self* OR fitness)) (Philopatr* AND (inbred OR in*breeding OR heterosis OR self* OR fitness))

Fitness related strings	(Depression SAME in*bre*) (Depression SAME fitness) (Heterosis AND in*breeding) "Genetic load"

Boolean syntax above is based on Web of sciences template, and changes were made to adapt to different database.

2.3.2 Inclusion and exclusion criteria

I followed the topic scope and criteria for relevance for potentially relevant articles that were established by Neaves and co-workers (2015). This allowed me to retrieve the articles that assessed the phenotypic outcomes of inbreeding depression in natural and laboratory populations and assess their relevance against standardised criteria. The following section recapitulates details how Neaves and her colleagues (Neaves *et al*., 2015) defined the relevance criteria for articles by assessing the study organisms, phenotypic outcomes, the presence of comparators and types of studies.

- **Type of study**

  Empirical studies were considered as relevant where inbred groups could be identified based on pedigree data and phenotypic outcomes were well-recorded. Pedigree data defining inbred and outbred progeny were used when they were based on experimental crosses or marker-based or observational records of wild populations. Book sections and conference papers were excluded. Review and meta-analysis articles were not included but missing articles from their bibliographies were assessed with the same criteria. Their bibliographies were used to assess the specificity of my literature search and any missing articles was included in my inbreeding database.

- **Types of subjects**

  Relevant subjects were defined as natural populations of plants and animals at any location worldwide. Natural populations were defined to include native natural populations or introduced, naturalised populations that inhabit natural habitats and that persist in the absence of human intervention. I excluded studies documenting inbreeding effects resulting from mating among agricultural cultivars or zoo

populations since these populations have often been subjected to human-mediated regime of artificial selection (either knowingly or unknowlingly). Microorganisms were also excluded from my study for the following reasons. First, many microorganisms are poorly described taxonomically. Second, the meaning of inbreeding for microorganisms is different from eukaryotic organisms since they have unique mechanisms of reproduction and inheritance. Finally, microorganisms are not the target for conservation measures, such as genetic rescue that are employed to mitigate inbreeding depression.

- **Types of phenotypic outcome**

  Relevant outcomes were phenotypic data from the offspring of inbred and outbred crosses. This included fitness-related traits such as fecundity, viability and survival, and traits that are more distantly related to fitness, such as body size, growth rate and behaviour.

- **Types of comparators**

  The comparisons used to calculate inbreeding depression effect sizes were phenotypic comparisons between offspring from inbred and non-inbred groups. These inbred and non-inbred groups were defined based on pedigree information manipulated in the laboratory or observed in the wild. Inbreeding coefficient *FIS*, the deviation of the observed heterozygosity of an individual relative to expected heterozygosity under random mating was not included. The inbred groups are referred to the progenies stemming from inbreeding between related parents. The corresponding outbred groups are arising from either random mating within population or outcrossing between populations. The outbred groups acted as a reference allowing us to quantify the magnitude of inbreeding depression via differences in mean phenotype.

2.3.3 Structural topic modelling

Structural topic modelling is a form of topic modelling that utilises machine-learning technique to analyse textual data. Structural topic modelling is derived from the latent Dirichlet allocation, a three-level hierarchical Bayesian approach that models each segment of text as a finite mixture of words over an underlying set of topics (Blei, Ng,

& Jordan, 2003). In R, the Structural Topic Model stm package (Roberts *et al*., 2019) allows researchers to identify and explore underlying topics in a text corpus and flexibly assess their correlation with document-level metadata, which contains information about documents of research interest (R Core Team, 2019). In my study, text corpus was referred to all the relevant/ irrelevant abstracts retrieved from the literature searches in the master library. It contains a list/vector of text segments, forming a single section of corpus. The general workflow of the stm package is presented in Fig.1 below (Roberts *et al*., 2019). In overview, a text corpus will be ingested and trimmed, by dropping punctuation and stop words, creating a list/vector/ array of words, which contain information on the topics relevant to each text segment. This step ensures the text corpus is prepared for further analysis. Next, the text corpus is classified by the topics identified in the model. Finally, users are able to evaluate between-topic correlations among text segments and their relationships to text metadata.

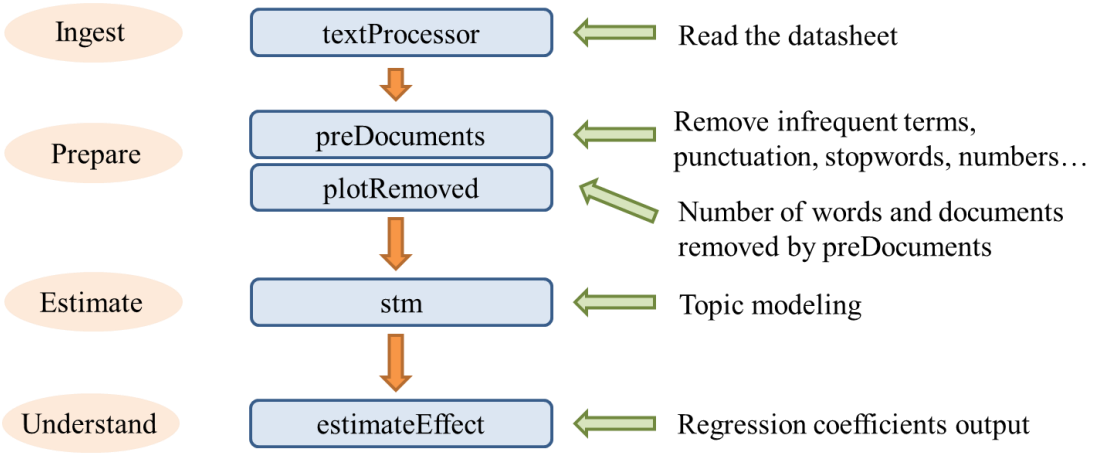


Figure 1. Principle workflow of stm package and some selected features and functions that were applied in this study (Roberts *et al*., 2019).

I performed ACA on the abstracts of Neaves's master library (training set), updated database (testing set) and both the mastery library and updated database described above using stm package (version 1.3.6) in R. The procedure and the functions of the stm package are illustrated in Fig. 1 above. In summary, the whole datasheet (both training and testing sets) was read in to R programme with textProcessor (Roberts et al., 2019). Infrequent words, punctuation, stop words and numbers were removed using the prepDocuments function (Roberts et al., 2019). The lower.thresh parameter in

prepDocuments was used to remove infrequent terms based on the minimum time of appearance that required a word to be kept during the processing (Roberts *et al*., 2019).

Next, topic modelling was carried out with stm, utilising topical prevalence that estimates covariates–the relevance of articles in related to systematic review assessed by Neaves– which could influence the frequency with which a topic is mentioned (Fig. 1). Neaves and co-workers' (205) assessment of article relevance was used as a covariate in topical prevalence in this analysis. The argument K in stm allows the user to decide the desired number of topics to be assessed in the model. Different combinations of lower.thresh (15, 20, 25, 30, 35, 40) and K values (15, 20, 25, 35, 40) and their corresponding results were inspected to evaluate the minimum number of adequate topics that were able to represent the content of the text corpus. Threshold number and K were set to 25 and 25, respectively for the final analysis. This led to the generation of 25 topics following analysis using the topic modelling function.

Afterwards, the estimateEffect function was used to construct a regression model, allowing assessment of the relationship between topics and their known relevance assessed by Neaves and co-workers (2015). To quantify the topic relevance to my study, the effect sizes (d) for each of 25 topics were calculated from the following equation (Nakagawa & Cuthill, 2007):

$$d = \frac{t(n_1+n_2)}{\sqrt{n_1 n_2}\sqrt{df}} \quad (1)$$

Where n_1 and n_2 are sample size in each group. In this case, these sample sizes are the numbers of articles which were assessed by Neaves *et al.* (2015) as relevant and irrelevant respectively. *df*) corresponds to the *t* value in the linear regression for topic relevance in the stm model. Topics with significant positive effect sizes indicate that those topics are relevant to the literature review, whereas topics with significant negative effect sizes are those that are likely to be irrelevant. The topics with the highest and lowest effect sizes and significant *t* values are likely to be those that are most reliably linked with relevance, and therefore are the most discriminating among relevant and irrelevant articles.

2.3.4 Linear discriminant analysis

Linear discriminant analysis (LDA) is a statistical methodology that allows researchers to examine linear combinations of candidate variables (topic proportion in this case) in order to determine the linear combinations of variables that best separate groups and to predict group membership for observations where their group has not been determined (Maindonald & Braun, 2010). To predict the relevance of articles within the updated database (testing data set), whose relevance was not known, the master library from original Neaves' inbreeding datasheet was used as a training set in a linear discriminant analysis (relevance of these articles had already been assessed). We obtained topic proportion information (the variable scores on which LDA would be carried out on) by retrieving theta from the stm model and merging it with the master library. Theta is the component of stm model indicating the number of documents by number of topic matrix that presents topic proportions for each article in my database (Roberts *et al*., 2019). In my dataset, theta listed the proportions of topics contributed to each article. I used LDA to construct a linear equation of topic proportion explanatory variables in relation to relevance (0, 1) for each article as groups. LDA is able to give a group assignment to articles (relevant/ not relevant) but also gives a linear predictor score for each article (on which group assignments are based). The following three subsets of topic-proportion variables were applied in separate LDA models applied to the training data set:

1) All topics were included which were significant predictors of article abstract relevance (*all topic model*)
2) All topics with significant positive effect sizes for abstract relevance (*positive topic model*)
3) All topics with significant negative effect sizes for abstract relevance (*negative topic model*)

Subsequently, linear predictors from these models were applied to the updated literature database (testing set), allowing us to predict relevance for each article, based on topic information within each abstract. Finally, I evaluated the predictions for abstract relevance in the test data set by comparing them to those obtained in a manual assessment of article relevance. LD1 values from the 4 combinations of article relevance (relevance/ irrelevance assessed manually or LDA) were assessed to understand how

these values differed between LDA prediction and manual assessment across three topic models (*positive, negative and all topic models*). Chi-square analysis was applied to test if manual assessment and LDA prediction were independent to each other. I. e. whether there were statistically significant deviations of observed frequencies from the expected frequencies for different combinations of LDA predictions and manual assessment. If Chi-square is statistically significant, then the LDA prediction and manual assessment are dependent in some way. Then, Cohen's kappa (Cohen, 1968) was used to measure the agreement between LDA predicted relevance and manual assessment of relevance.

The distributions of LD1 values for relevant and irrelevant articles were plotted in the histograms with three thresholds for further evaluation. Relevant articles distributed on the left-hand side of the threshold represent false negative prediction that will be missed in the meta-analysis. Alternatively, irrelevant articles distributed on the left-hand side of the threshold represent articles that will be correctly eliminated. Workload reduction was calculated as the proportion of articles eliminated by ACA-LDA prediction. Except from the LD1 value applied by LDA, I also chose two candidate LD1 values (0 and -0.6) to test if choosing different thresholds can improve the performance of LDA prediction in all topic model. The candidate thresholds were chosen based on the distribution of LD1 values, to minimise the loss of relevant articles (around 5%) while maintaining workload reduction above 30% (O'Mara-Eves *et al*., 2015).

2.3.5 Manual literature assessment

Literature assessment was carried out manually with the assistance of theta matrix from the stm model. I combined the theta matrix with the updated inbreeding data set, sorting the articles by topic proportion. This allowed me to assess the articles with similar subjects simultaneously. The topics with negative effect sizes (indicating the most irrelevant topics) from stm model were sorted to the top of the list and the irrelevant articles were able to be eliminated rapidly. The relevant and irrelevant articles with their LD1 values summarised by LDA were plotted as histograms to summarise the level of agreement between ACA-LDA abstract assessment and manual assessment.

2.4 Results

2.4.1 Data

My updated literature search retrieved 29863 articles from 1941 to 2018. The number of articles returned increased sharply after 1990 and continued to rise steadily until it reached the peak at 2012 (Fig. 2). Thereafter the number of published articles decreased. The small number of articles returned in 2018 is due to the fact that the literature search was completed in May 2018. Therefore, papers published after May 2018 were not included in this study. The systematic map synthesised by Neaves *et al.*, recorded 11111 inbreeding-relevant articles before 12th August 2013. The number of articles retrieved by my literature search was 25586 with same search terms in the same period and literature databases. This difference might be partially explained by the fact that Neaves' literature search included articles up until August 2013, whereas my search included all articles in 2013. In addition, I also noticed that the articles retrieved from some search terms increased drastically in the same period compared with Neave's results. In particular, 10269 articles were retrieved by both literature searches while 15317 articles were recovered by me, and 842 articles were recovered only by Neaves and her colleagues.

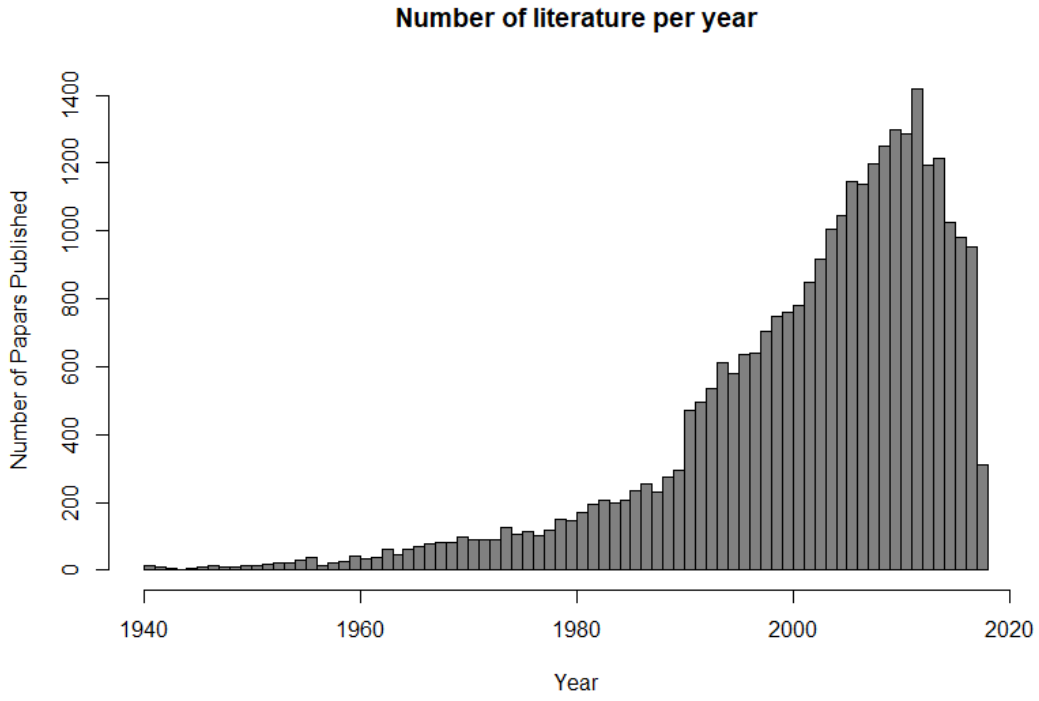


Figure 2. Numbers of published articles returned by the inbreeding depression literature search (updated to June 2018) by year of publication. The histogram includes all returned articles and therefore may include articles whose main focus is not on inbreeding depression.

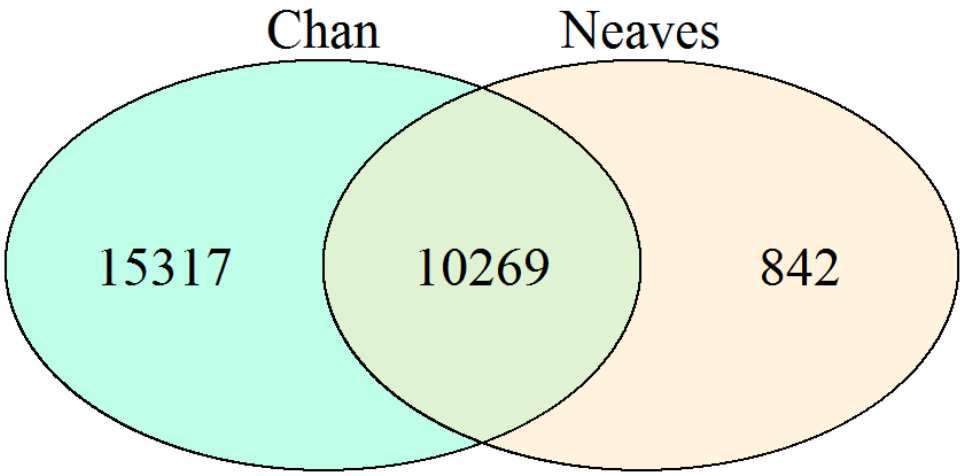


Figure 3. Venn diagram showing the overlap in literature search returns for articles retrieved by Neaves *et al.* (2015) and in this work. The diagram only includes articles returned up to 2013, to ensure comparability between the two literature searches.

2.4.2 Structural topic modelling

The 25 topics identified by the structural topical model (applied to Neaves *et al.* article assessment, the training set) are listed in table 3 with their corresponding effect sizes for article relevance. The words listed beside to each topic are the top 5 words which have the highest probability to occur in that topic. The effect sizes with significantly positive values ($p < 0.05$) indicate topics that are significant predictors of abstract (article) relevance and vice versa. The two most relevant topics were topics 6 and 20 with the effect sizes of 0.475 and 0.472, respectively. The key words associated with topic 6 (plant, seed, flower, pollen and pollin) indicate botanical studies that use artificial pollination to manipulate inbreeding, while the key words in topic 20 (inbreed, depress, effect, generat, inbr) indicate general terms associated with inbreeding depression studies. Alternatively, topics most negatively associated with article relevance were topics 3, 22 and 23 with corresponding effect sizes of -0.2896, -0.315 and -0.347,

respectively. Key words associated with Topic 3 (such as family, consanguinity, and marriage) suggest studies in human populations (irrelevant to this thesis). Likewise, key words in topic 22 indicate clinical or biomedical papers that investigate the relationship between a patient's mental health depression and physical exercise. Moreover, topic 23 could be related to genetic studies focus on chromosomes, genomic and alleles. This topic was less relevant to my study due to the lack of phenotypic measurements.

Table. 2 25 topics identified by ACA and their corresponding effect sizes for article relevance calculated from equation (1).

Topic	Key words	Effect sizes
Topic 1	densiti, increas, cost, size, resource	0.00441
Topic 2	genet, popul, divers, among, structure	-0.07033*
Topic 3	famili, consanguin, studi, marriag, group	**-0.28956*** **
Topic 4	use, genet, wild, develop, strategi	-0.22361***
Topic 5	resist, host, parasit, infect, diseas	0.05887
Topic 6	plant, seed, flower, pollen, pollin	0.47476***
Topic 7	model, popul, select, estim, rate	-0.19428***
Topic 8	group, social, femal, male, behavior	-0.09036**
Topic 9	mice, strain, rat, effect, behavior	-0.19718***
Topic 10	breed, use, estim, anim, pedigree	-0.22896***
Topic 11	dispers, habitat, seed, spatial, speci	0.0255
Topic 12	line, hybrid, cross, parent, yield	-0.16726***
Topic 13	studi, can, provid, model, understand	-0.16745***
Topic 14	male, femal, mate, sex, offspr	0.17351***
Topic 15	growth, time, life, surviv, rate	0.23392***
Topic 16	breed, nest, bird, dispers, surviv	0.01662
Topic 17	cell, mice, express, gene, mutat	-0.14803
Topic 18	que, los, las, des, les	0.00248
Topic 19	speci, hybrid, reproduct, morpholog, two	0.01259
Topic 20	inbreed, depress, effect, generat, inbr	0.47209***
Topic 21	popul, speci, size, effect, habitat	0.15003***
Topic 22	physic, exercis, depress, patient, studi	-0.31455***
Topic 23	gene, chromosom, genom, marker, map	-0.34682***
Topic 24	popul, island, region, rang, nativ	-0.02454
Topic 25	select, trait, variat, genet, fit	0.32415***

Red texts indicate the topics associated with article relevance whereas purple texts indicate the topics associated with irrelevance articles. *** indicates significant *p* values less than 0.001; ** indicates significant *p* values larger than 0.001 but less than 0.01; * indicates significant *p* values range from 0.05 to 0.01.

2.4.3 Linear discriminant analysis

Across three subsets of topic-proportion variables applied to LDA models, articles predicted as irrelevant by LDA had lower median LD1 (below 1) whereas articles predicted as relevant had higher median LD1, lying between 3 and 4 (Fig. 4). In addition, articles predicted as relevant by LDA had similar distribution of LD1 values in *positive* and *all topic models* (Fig. 4a; 4b). When categorising the articles assessed as relevant and irrelevant manually or by LDA (4 categories), *positive topic* and *all topic models* had similar numbers of articles among these 4 categories while *negative model* had small numbers of articles predicted as relevant by LDA, 12 and 1 for articles manually assessed as irrelevant and relevant separately (Table. 3). The *negative topic model* categorised 99.9% of articles as irrelevant. Furthermore, the article numbers at the bottom left cells of the table (1040,1181 and 1053 for three models) indicated the false negative predictions, which would eventually lead to the loss of relevant articles for my meta-analysis (Table. 3).

Chi-squared test supported the significant difference between the expected frequency and observed frequency, indicating manual assessment was associated with LDA prediction (χ-squared = 512.23, p-value < 0.001, for *positive topic model*; χ-squared < 0.001, p-value = 1, for *negative topic model*; χ-squared = 416.56, p-value < 0.001, for *all topic model*). Then, Cohen's kappa analysis was applied to test the agreement between ACA-LDA predictions and manual assessment, with *k* values of 0.135, 0.0006 and 0.122 for *positive*, *negative* and *all topic models* respectively. The relatively small *k* values from Kappa analysis indicated the disagreement between manual assessment and LDA predictions.

Threshold applied from *all topic model* has highest LD1 values, 1.973 (red line), leading to 89.1% of relevant articles being eliminated, accompanying by 97.7% of workload reduction (Fig. 5). Applying different LD1 thresholds did improve the ACA-LDA prediction by reducing the loss of relevant articles (false negative) but at the cost of the increase in workload (Fig. 5). Specifically, there was a trade-off between sensitivity (detecting relevant articles) and efficiency (workload reduction). In *all topic model,*

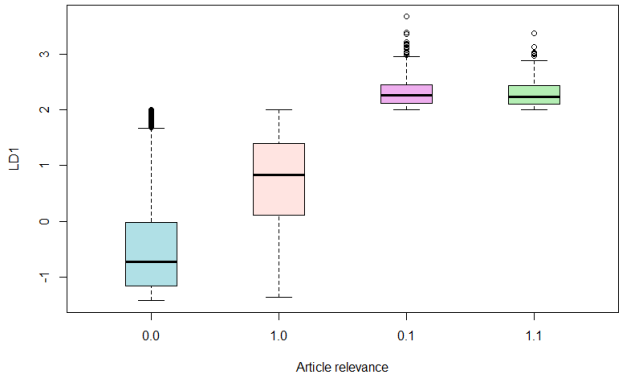
setting LD1 threshold to 0 (Fig. 5, blue lines) resulted in the loss of 17.1% of relevant articles with 70% of workload reduction. Alternatively, setting LD1 threshold to -06 (Fig. 5, purple lines) decreased the loss of relevant articles from 89.1% to 6.1%, with 45% of workload reduction. There were 8.3% and 13.4% of articles predicted correctly as relevant when applying -0.6 and 0 LD1 threshold in comparison to 23.3 % of the default threshold (1.973).

Table. 3 Contingency tables showing the numbers of relevant and irrelevant articles as predicted by ACA-LDA and manual assessment of abstracts.

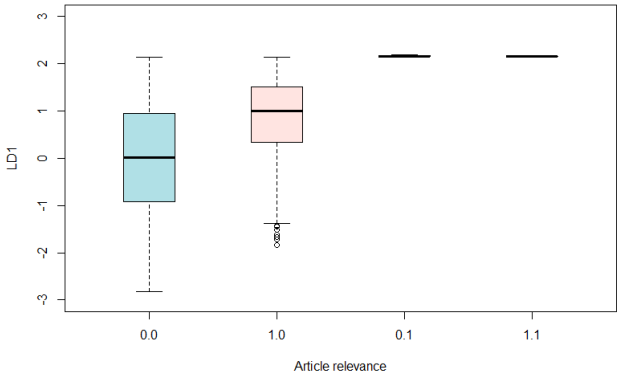
Manual assessment of article relevance	ACA-LDA prediction of article relevance					
	Positive effect size topics		Negative effect size topics		All topics	
	0	1	0	1	0	1
0	22847	424	23259	12	22847	424
1	1040	142	1181	1	1053	129

Three contingency tables are given, one for each LDA analysis carried out (see methods). The first analysis (positive effect size topics) included only topic predictors that were positively associated with article relevance (Table 2). The second analysis (negative effect size topics) included only topic predictors that were negatively associated with article relevance (Table. 2). The "all topics" LDA included all topics (both the positive and negative effect size topics).

(a)



(b)



(c)

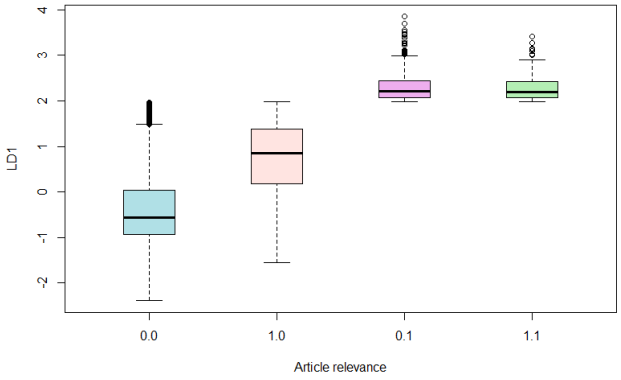
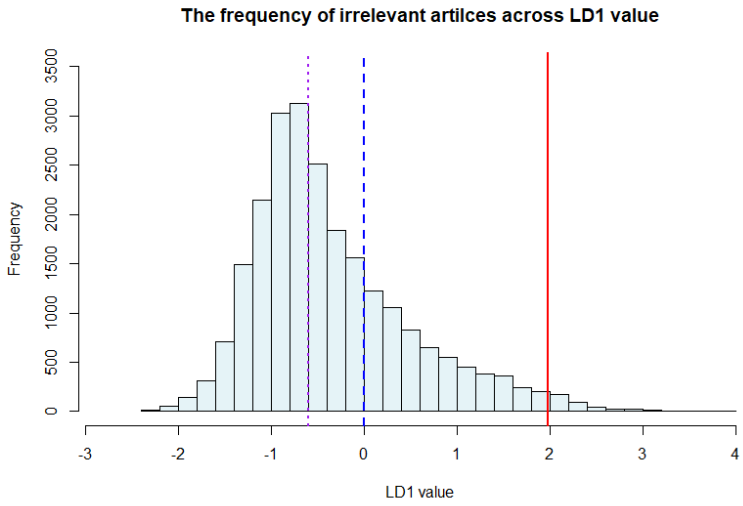


Fig.4 Boxplots showing the distributions of LD1 scores across 4 combinations of article relevance based on manual assessment and ACA-LDA prediction. (a) *positive topic model*; (b) *negative topic model*; (c) *all topic model*. Since there were only two article relevance groups (relevant and irrelevant), there is only one linear discriminant axis (LD1). 0.0 and 1.1 represent articles for which both ACA-LDA and manual assessment agreed on article relevance (either both irrelevant, 0.0 or both relevant, 1.1). 1.0 represents the articles assessed as relevant manually but predicted as irrelevant by ACA-LDA. In contrast, 0.1 represents articles assessed as irrelevant manually but predicted as relevant by ACA-LDA.

(a)



(b)

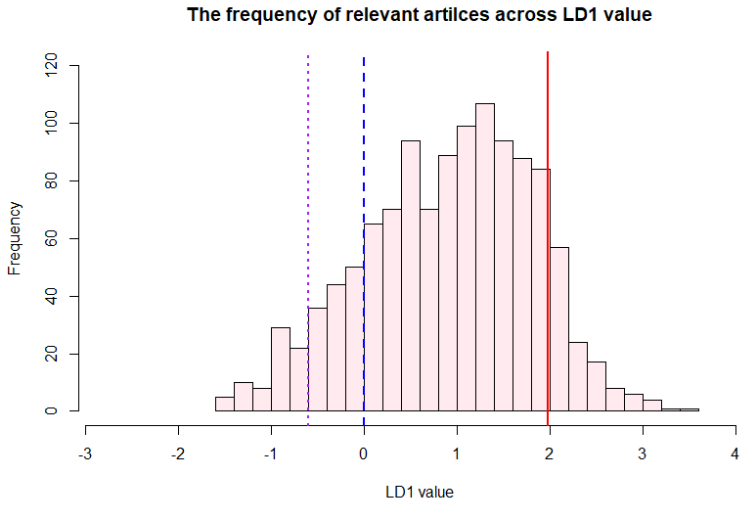


Fig. 5 The distribution of LD1 scores for irrelevant (a) and relevant (b) articles assessed manually. Red line represents the threshold (1.973) for relevance prediction applied by LDA: articles above this threshold would be predicted the be relevant, whilst those with LD1 scores below this value would be predicted to be irrelevant. Blue and purple vertical lines represent additional candidate LD1 thresholds for assessing article relevance. (a) Articles on the right-hand side of lines indicates irrelevant articles incorrectly predicted as relevant (false positive). (b) Articles on the left-hand side of lines indicates relevant articles incorrectly predicted as irrelevant (false negative).

2.5 Discussion

In the field of ecology, the application of text-mining approaches has been primarily focussed on exploring the trends within an investigated subject area through a period of time (McCallen *et al*., 2019; Millard, Freeman, & Newbold, 2020; Nunez-Mir *et al*., 2015). My study has provided the first example of utilising machine-learning process to screen relevant literature from a large number of irrelevant articles for a meta-analysis in the field of ecology. First, I showcased the use of ACA, in combination with LDA to accelerate the identification of relevant articles for meta-analysis by reducing the workload or screening prioritisation. Second, I detected a trade-off between workload reduction and sensitivity (for detecting relevant articles) by applying LDA as automatic classification. Finally, I identified two candidate LD1 thresholds (-0.6 and 0), instead of default LD1 (1.973) to reduce the loss of relevant articles (6 % and 17 %) with still 45% and 70 % of workload reductions respectively.

In the following section, I discussed the two approaches that ACA could assist the identification of the relevant articles. The first approach is by reducing the workload with automatic classification. The aim of the first approach is to reduce researcher workload by decreasing the number of papers or text items that have to be manually assessed. The second approach is to accelerate the speed of article assessment at the screening stage by screening prioritisation.

2.5.1 Automatic classification

The first approach emphasised decreasing the number of articles that are required to screen manually. Automatic classification achieves this goal by automatically applying exclusion and inclusion criteria to the articles in a database (Thomas *et al*., 2011). Application of the ACA-LDA approach proposed in this work would reduce the workload only when topics that were positively and significantly associated with article relevance were used as explanatory variables in LDA. Conducting LDA with only topics negatively associated with article relevance resulted in the majority of the papers being classified as irrelevant (99.9 %, Table. 3), suggesting that negative topics were unable to identify relevant literature. The possible reason might be that the irrelevant topics told the model little about whether it was an inbreeding relevant article. Further

investigation suggested that articles classified as relevant by *negative topic model* had extremely low proportions (mostly below 0.4%) of negative topics. This was a rare occasion that had to occur among all 11 negative topics simultaneously, leading to small proportion of articles classified as relevant. However, concerns still arise when topics were assigned as explanatory variables that were positive indicators of article relevance, since the numbers of false negative cases accounted for 87.0% and 89.1% of the total relevant articles for *positive* and *all topic models* respectively. Unlike false positives (articles predicted to be relevant that are actually irrelevant), which will be corrected in the later stage in article assessment, false negative articles are the main concern in automatic classification and the primary reason for losses of relevant article.

The possible reason for the large loss of relevant articles during assessment using ACA-LDA might be due to the high threshold value which LDA uses to make group predictions (article relevance or irrelevance). Specifically, articles with LD scores lower than 1.973 were categorised as irrelevant (red line, Fig. 5), resulting 89.1% loss of relevant articles. Therefore, it might be impractical to use default LD1 threshold to assess article relevance in my ACA-LDA process for systematic review. However, applying lower LD1 threshold provided a promising solution. In *all topic model*, setting LD1 threshold to -0.6, 45% of workload reduction can be achieve with the risk of losing only 6% of relevant articles (74) as false negative (relevant articles that are predicted to be irrelevant). In comparison to the systematic review carried out by O'Mara-Eves and co-workers (2015), which suggested using text-mining approaches to detect relevant articles for systematic review in the field of biomedicine resulted in 30% - 70% of workload reduction accompanied by a 5% loss of relevant articles. Therefore, in my ACA-LDA approach, setting threshold value at -0.6 might be a more viable option to utilise ACA on meta-analysis.

Further investigation on the relevant articles with the most negative LD scores suggested a lack of proper information regarding whether populations experienced artificial selection or not in their abstracts. Particularly, these abstracts were often related to agricultural species, with high proportions of topics 4, 10, 12 (Table 2). Details regarding whether these agricultural species were derived from natural populations or experienced any artificial selection are typically described in the method sections. In

my case with -0.6 LD1 threshold, 56.9% of the false negative articles with ACA-LDA prediction were related to agricultural populations, which usually require full-text assessment to determine their relevance. Furthermore, 9.7 % of the false negative articles are associated with genetic studies. These studies had high proportions of topic 2 and 23. Some of these genetic studies investigated how inbreeding depression affected genetic diversity. Other investigated the use of genetic markers to infer the magnitude of inbreeding. Both of these studies rarely recorded any phenotypic outcome of inbreeding depression and thus, were highly possible irrelevant to my meta-analysis. In my manual assessment, I could apply a more conservative method than Neaves' in article screening, keeping majority of abstracts mentioned agricultural species in case their studies used non-agricultural populations. Likewise, genetic studies were included for full-text assessment if their abstracts mentioned inbred offspring and phenotypes. Therefore, this result only implied the false negative cases at the stage of abstract screening. It is still likely that the putatively relevant articles with the lowest LD1 scores could be rejected at full-text assessment level. As a result, the disagreement between ACA-LDA relevance prediction and manual assessment doesn't necessarily indicate the unreliability of ACA. In practice, the percentage of false negative cases is likely to drop after full-text assessment, making the ACA-LDA process a reliable method for article screening in systematic review.

2.5.2 Screening prioritisation

Despite the fact that screening prioritisation cannot reduce the number of articles that need to be screened, it still has prominent practical benefit by speeding up manual assessment. Typically, screening prioritisation achieve the acceleration of screening process by presenting assessors with an order of the articles, with the articles that are more likely to be relevant to the systematic review at the top of the order. This allows researchers to gain a better understanding of inclusion criteria at the early stage of screening (O'Mara-Eves *et al*., 2015). However, there is a caveat. This could lead to biases and careless assessment when a researcher thinks articles are less likely to be relevant, generating undetected false negatives. Additionally, Cohen (2008) stated that "by reviewing the most likely important documents before other documents, the human reviewers or curators are more likely to be able to "get up to speed" on the current developments within a domain more quickly". In my ACA-LDA method, screening

prioritisation can be achieved by combing the theta matrix (the number of documents by number of topic matrix) to the literature database, creating a new Excel datasheet. As a result, I was able to inspect the topic proportion for each article. Sorting the topic proportions for most relevant and irrelevant topics allow researchers to cluster similar types of articles together (e.g. those in the same sub-field or sharing methodological or conceptual similarities). For example, I can sort the new datasheet by the proportions of topic 22 (physic, exercis, depress, patient, studi). Therefore, the articles related to clinical studies, which are apparently irrelevant to my meta-analysis, can be assessed and quickly eliminated at the same time. Specifically, this approach improves the workflow and efficiency for researchers to assess the relevance of the articles for their meta-analysis.

2.5.3 Limitations and future implications

To reduce the number of articles that needed to be assessed, automatic classification requires a training data set to build an ACA model that can be used for LDA analysis. However, there might be a potential bias if the training data set cannot represent a whole literature database, a phenomenon called "hasty generation" (Wallace, Small, Brodley, & Trikalinos, 2010). My ACA-LDA process might be a suitable method for a review update, which already has a pre-existing, well curated and manually assessed literature database. For a brand systematic review, it is still a question as to how many articles should be used to form a training set and a test set to be able to represent a whole literature corpus. In addition, the inbreeding database was the only dataset I applied with ACA-LDA approach in my thesis. Whether using -0.6 LD1 as a threshold value for LDA is viable option for other datasets? The lack of replication and assessment of this ACA-LDA process on other datasets makes it tough to draw a conclusion on whether -0.6 is a general suitable threshold. Furthermore, the disagreement between my manual assessment and ACA-LDA prediction indicates that some information might be missing in these abstracts. The information required for a systematic review depends on how we define the scope and inclusion criteria. In my case, I excluded any articles using agricultural species or populations experienced artificial selection. However, this information is typically presented in the method section of an articles, making it unlikely to be detected by ACA. Utilising text-mining approach on full-text scale could improve the accuracy of automatic classification (Westergaard, Staerfeldt,

Tonsberg, Jensen, & Brunak, 2018). However, it may still be unrealistic to expect that text mining will perform perfectly, since text mining approaches at full-text level is likely to be limited by the format of text (e.g. in scans of physical copies), biasing for or against older articles, and some articles are likely to be unavailable behind a paywall (Cornford, Millard, Gonzalez-Suarez, Freeman, & Johnson, 2022). Text-mining of full texts would also require higher performance computers and software. At the current stage, I would recognise that the ACA-LDA approach to abstract assessment used here has high potential and is likely to be a reliable tool to support assessment of article relevance.

I suggest that future studies should experiment with the ACA-LDA approach using different data sets with different sizes of training sets, to validate its utility in the article screening process for meta-analysis and to test whether applying LD1 as -0.6 is a universal threshold for this method.

2.6 References

Altaweel, M., Bone, C., & Abrams, J. (2019). Documents as data: A content analysis and topic modeling approach for analyzing responses to ecological disturbances. *Ecological Informatics, 51*, 82-95. doi:10.1016/j.ecoinf.2019.02.014

Ananiadou, S., McNaught, J., & Karamanis, N. (2005). Text Mining for Biology And Biomedicine. *Artech House, London, 33*.

Arnqvist, G., & Wooster, D. (1995). Meta-analysis - synthesizing research findings in ecology and evolution. *Trends in Ecology & Evolution, 10*(6), 236-240. doi:10.1016/s0169-5347(00)89073-4

Blei, D. M. (2012). Probabilistic Topic Models. *Communications of the Acm, 55*(4), 77-84. doi:10.1145/2133806.2133826

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research, 3*(4-5), 993-1022. doi:10.1162/jmlr.2003.3.4-5.993

Cohen, A. M. (2008). Optimizing Feature Representation for Automated Systematic Review Work Prioritization. *AMIA ... Annual Symposium proceedings. AMIA Symposium*, 121-125.

Cohen, J. (1968). Weighted Kappa - Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin, 70*(4), 213-&. doi:10.1037/h0026256

Cornford, R., Millard, J., Gonzalez-Suarez, M., Freeman, R., & Johnson, T. F. (2022). Automated synthesis of biodiversity knowledge requires better tools and standardised research output. *Ecography, 2022*(3). doi:10.1111/ecog.06068

Hearst, M. A. (1999). *Untangling text data mining*. Paper presented at the Proceedings of the 37th annual meeting of the Association for Computational Linguistics on Computational Linguistics, College Park, Maryland. https://doi.org/10.3115/1034678.1034679

James, K. L., Randall, N. P., & Haddaway, N. R. (2016). A methodology for systematic mapping in environmental sciences. *Environmental Evidence, 5*(1). doi:10.1186/s13750-016-0059-6

Maindonald, J., & Braun, W. (2010). Regression on principal component or discriminant scores. In J. Maindonald & W. J. Braun (Eds.), *Data Analysis and Graphics Using R: An Example-Based Approach* (3 ed., pp. 410-426). Cambridge: Cambridge University Press.

McCallen, E., Knott, J., Nunez-Mir, G., Taylor, B., Jo, I., & Fei, S. L. (2019). Trends in ecology: shifts in ecological research themes over the past four decades. *Frontiers in Ecology and the Environment, 17*(2), 109-116. doi:10.1002/fee.1993

Millard, J. W., Freeman, R., & Newbold, T. (2020). Text-analysis reveals taxonomic and geographic disparities in animal pollination literature. *Ecography, 43*(1), 44-59. doi:10.1111/ecog.04532

Nakagawa, S., & Cuthill, I. C. (2007). Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biological Reviews, 82*(4), 591-605. doi:10.1111/j.1469-185X.2007.00027.x

Neaves, L. E., Eales, J., Whitlock, R., Hollingsworth, P. M., Burke, T., & Pullin, A. S. (2015). The fitness consequences of inbreeding in natural populations and their implications for species conservation - a systematic map. *Environmental Evidence, 4*(1). doi:10.1186/s13750-015-0031-x

Nunez-Mir, G. C., Iannone, B. V., Curtis, K., & Fei, S. L. (2015). Evaluating the evolution of forest restoration research in a changing world: a "big literature" review. *New Forests, 46*(5-6), 669-682. doi:10.1007/s11056-015-9503-7

Nunez-Mir, G. C., Iannone, B. V., Pijanowski, B. C., Kong, N. N., & Fei, S. L. (2016). Automated content analysis: addressing the big literature challenge in ecology and evolution. *Methods in Ecology and Evolution, 7*(11), 1262-1272. doi:10.1111/2041-210x.12602

O'Mara-Eves, A., Thomas, J., McNaught, J., Miwa, M., & Ananiadou, S. (2015). Using text mining for study identification in systematic reviews: a systematic review of current approaches. *Systematic Reviews, 4*. doi:10.1186/2046-4053-4-5

Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). stm: An R Package for Structural Topic Models. *Journal of Statistical Software, 91*(2), 1-40. doi:10.18637/jss.v091.i02

Stewart, G. B., Coles, C. F., & Pullin, A. S. (2005). Applying evidence-based practice in conservation management: Lessons from the first systematic review and dissemination projects. *Biological Conservation, 126*(2), 270-278. doi:10.1016/j.biocon.2005.06.003

Stockwell, P., Colomb, R. M., Smith, A. E., & Wiles, J. (2009). Use of an automatic content analysis tool: A technique for seeing both local and global scope. *International Journal of Human-Computer Studies, 67*(5), 424-436. doi:10.1016/j.ijhcs.2008.12.001

R Core Team (2019). R: A language and environment for statistical computing.

Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/.

Thomas, J., McNaught, J., & Ananiadou, S. (2011). Applications of text mining within systematic reviews. *Research Synthesis Methods, 2*(1), 1-14. doi:10.1002/jrsm.27

Wallace, B. C., Small, K., Brodley, C. E., & Trikalinos, T. A. (2010). *Active learning for biomedical citation screening*. Paper presented at the Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining, Washington, DC, USA. https://doi.org/10.1145/1835804.1835829

Westergaard, D., Staerfeldt, H. H., Tonsberg, C., Jensen, L. J., & Brunak, S. (2018). A comprehensive and quantitative comparison of text-mining in 15 million full-text articles versus their corresponding abstracts. *Plos Computational Biology, 14*(2). doi:10.1371/journal.pcbi.1005962

Chapter 3: Sources of heterogeneity in inbreeding depression in populations of animals and plants: a meta-analysis

3.1 Abstract

Inbreeding depression, the reduction in fitness of inbred offspring, is a major concern for the conservation of small and isolated populations. Inbreeding depression is known to reduce mean fitness. However, the magnitude and severity of inbreeding depression can vary depending on the traits measured and species' characteristics. Progress in understanding this heterogeneity in the magnitude of inbreeding depression, and in drawing up appropriate conservation policies that incorporate it has been limited. A prominent reason for this is that the large body of existing evidence that could inform on how and why inbreeding responses vary has not been sufficiently well synthesised. In this study, I used meta-analysis to investigate how inbreeding depression varies with regard to the types of traits, species characteristics and other contextual factors in populations of animals and plants. My results indicate significant variation in inbreeding depression among different types of phenotypic traits. Fecundity and fitness suffered more inbreeding depression than behavioural, physiological and defensive traits. Additionally, phenotypes observed in late life history stages experienced more inbreeding depression than those in early and middle life history stages. However, I did not detect any variation among taxonomic groups. In conclusion, inbreeding depression should be considered as an intrinsic component in conservation management, and it did influence population persistence by decreasing population growth and reproductive output.

3.2 Introduction

Genetic diversity is an essential element that underpins the ability of a populations to respond to environmental variation and change. It is the raw material upon which natural selection can operate, enabling evolutionary adaptation to changing habitat conditions (Hoffmann & Sgro, 2011). Maintaining sufficient genetic diversity is a major concern in conservation biology to ensure population' long term sustainability. However, many wild populations suffer from the erosion and loss of genetic variation due to reduction in population size, restricting the efficiency of selection and consequently limiting opportunities for local adaptation to future environmental changes (Charlesworth, 2003; Frankham *et al*., 1999).

In small and isolated populations, inbreeding becomes inevitable because individuals have no option but to mate with their relatives. In consequence, their offspring may suffer from a reduction of fitness due to inbreeding depression (Crnokrak & Roff, 1999; Frankham, Ballou, & Briscoe, 2010; Hedrick & Garcia-Dorado, 2016; Keller & Waller, 2002). The genetic mechanisms of inbreeding depression arise from the random fixation of detrimental variants, as alternative alleles are lost stochastically, resulting in the reduction in the frequency of heterozygote genotypes. In the overdominance theory, inbreeding depression occurs through the loss of heterozygotic loci, which have superior fitness compared with their homozygotic counterparts (Charlesworth & Charlesworth, 1999; Charlesworth & Charlesworth, 1987; Charlesworth & Willis, 2009; Kristensen, Pedersen, Vermeulen, & Loeschcke, 2010). Alternatively, in the dominance theory, inbreeding depression is caused by the unmasking of recessive deleterious alleles, which confer detrimental phenotypes and the expression of genetic load (Charlesworth & Charlesworth, 1999; Charlesworth & Charlesworth, 1987; Charlesworth & Willis, 2009; Kristensen *et al*., 2010). When such variants are exposed due to homozygosity, the adverse phenotypes are expressed in their offspring, causing the reduction in reproductive output and, potentially, in consequence, population viability (Crnokrak & Roff, 1999; Keller & Waller, 2002).

The reduction in heterozygosity frequency or expression of deleterious alleles due to inbreeding could threaten population sustainability by decreasing reproductive output and survival. Along with inbreeding depression, genetic drift could randomly fix or eliminate alleles in small populations (Hedrick & Kalinowski, 2000). When alleles are lost, it will lead to an increase in the frequency of homozygosity, which has inferior benefits compared with two complimentary alleles in heterozygosity (Reed & Frankham, 2003). In the long run, populations may lose evolutionary potential and the ability to track changing environments. Persistently small population sizes can compound the effects of inbreeding depression by allowing opportunities for non-genetic factors, including environmental variation and demographic stochasticity, to negatively impact on population dynamics (Caro & Laurenson, 1994). Ultimately, this

negative feedback of decline in population size can leads to an “extinction vortex” (Frankham *et al*., 2010).

Fitness reduction caused by inbreeding depression can be alleviated by exposing deleterious alleles to natural selection, a process called purging (Hedrick & Garcia-Dorado, 2016). When individuals express deleterious alleles, they become vulnerable to natural selection due to low reproductive output and survival rate, resulting to the failure to pass their genotypes to future generations, and consequently reducing the frequency of deleterious alleles (Wang, Hill, Charlesworth, & Charlesworth, 1999). The effectiveness of purging has been assumed to cause heterogeneity in the intensity of inbreeding depression. This is because its efficiency is thought to depend on the architecture of the inbreeding load that underlies phenotypic traits and the effects of deleterious mutations (Charlesworth & Charlesworth, 1999; Hedrick & Garcia-Dorado, 2016; Keller & Waller, 2002). For instance, evolutionary theory predicts that purging is less effective against recessive deleterious alleles that are expressed late in life history because there is a lower contribution to reproduction late in the life history (Hamilton, 1966; Williams, 1957). Additionally, predominantly selfing populations are expected to experience intense purging and therefore are less likely to express inbreeding depression compared with predominantly outcrossing species (Husband & Schemske, 1996; Lande & Schemske, 1985). Likewise, large-statured plants or woody perennials have less possibility to evolve towards self-mating breeding systems due to the accumulation of mutations over time. Therefore, in these species, significant inbreeding depression is highly probable when mating occurs between close relatives (Scofield & Schultz, 2006). These examples further fortify the idea that the interaction between natural selection and mating systems is a major source of variation in inbreeding depression could stem from.

A growing number of studies have attempted to synthesise the relationship between inbreeding depression and phenotypic traits (Angeloni, Ouborg, & Leimu, 2011; DeRose & Roff, 1999; Husband & Schemske, 1996), species characteristics (Angeloni *et al*., 2011; Husband & Schemske, 1996) and mating systems (Angeloni *et al*., 2011; Duminil, Hardy, & Petit, 2009; Husband & Schemske, 1996). These studies have

provided a fundamental background for understanding heterogeneity in inbreeding depression. In summary, their results showed inbreeding depression was more severe in (1) traits were closely related to fitness; (2) traits expressed at late life history stage; (3) long-lived species; (4) self-incompatible or outcrossing species. However, all of them have been limited in their scope of study. Moreover, several of these studies synthesised evidence without taking into account within-study measurement error due to variation in sample sizes or experimental designs. The studies by Angeloni *et al*., Duminial, Hardy & Petit and Husband and Schemske focused on plant species only whereas DeRose & Roff only complied their data from animals. Moreover, the study by Duminil, Hardy and Petit (2009) utilised regression to analyse the effects of mating systems on inbreeding depression. Mann-Whitney U-test and Spearman's rank correlation were applied to test the impacts of life-history traits and mating systems on inbreeding depression. DeRose and Roff (1999) employed Kruskal-Wallis and paired t-tests to evaluate the relationship between phenotypic traits and inbreeding depression. None of these statistical methods are formal meta-analysis. Therefore, our understanding of heterogeneity in inbreeding depression is potentially biased, is incomplete taxonomically, and fails to make use of the large body of evidence currently available. Given the impacts of inbreeding depression on wild populations, there is an urgent need to carry out a comprehensive synthesis and meta-analysis of the relationship between inbreeding depression and sources of heterogeneity, incorporating plant and animal diversity and leveraging the large amount of information that now exists on this topic.

Meta-analysis provides us a powerful tool that allows us to rigorously synthesise evidence-based studies within a linear modelling framework (Borenstein, Hedges, Higgins, & Rothstein, 2009; Koricheva, Gurevitch, & Mengersen, 2013). By conducting Bayesian meta-analysis, I aim to disentangle the causes and consequences of inbreeding depression and relate the intensity of inbreeding depression to sources of heterogeneity, including types of phenotypic traits and species characteristics.

Here, I exclude extrinsic environmental sources of heterogeneity, which are the major focus of Chapter 4. My goal is to provide a foundation establishing the factors that are

strongly associated with inbreeding depression, and therefore need to be controlled in the future studies, and considered in conservation programmes. I predicted that fitness component trait, traits at late life history stage, species with longer life expectancy and outcrossing species experienced more inbreeding depression. Quantification of the key sources of heterogeneity influencing inbreeding depression is also a pre-requisite to evaluating the role of environmental stress in controlling the consequences of inbreeding.

3.3 Methods

3.3.1 Literature search

Literature searches were performed to retrieve articles investigating inbreeding depression empirically. A comprehensive, subject-based inbreeding database was compiled by Neaves and her colleagues in 2015, to construct a systematic map on the fitness consequences of inbreeding depression. This database contains inbreeding depression related literature from three online databases: Web of sciences core collection, Scopus and JSTOR, documenting the scope and quality of the evidence on the phenotypic outcomes of inbreeding in natural populations. To utilise this resource to facilitate a comprehensive meta-analysis of the factors generating variation in the intensity of inbreeding depression, it was necessary to update this database since it only contained articles before 12th August 2013. On 4 May 2018, I conducted a literature search with identical search strings and methods as used by Neaves *et al*. (2015; Chapter 2) to retrieve the relevant articles from 2015 to 2018.

I also conducted a supplementary literature search to address the difference in the number of search returns between my initial search and Neaves' database (Chapter2, Fig. 3). Particularly, there were some missing articles retrieved by Neaves *et al*. (2015), which were known to be relevant to my study, but which hadn't been returned in my first updated search using the same search strings. The additional supplementary literature search used refined search strings to capture the titles and abstracts of relevant articles that were missing in my initial search. Search strings were designed to improve the scope of the search without losing specificity. The final search strings of the supplementary search are listed in Table 1.

Ultimately, database-specific literature records were combined in a master library by endnote software (version X 9.2). Duplicates, book chapters, non-English and conference papers were removed. The final spreadsheet was created by exporting the endnote library to excel for assessment of article relevance.

Table 1. Search strings for the updated literature search to retrieve inbreeding depression relevant articles from online databases.

Group	Search string
Inbreeding related strings	"In*breeding coefficient$"
	"Cost$ NEAR/5 in*breeding"
	(inbred SAME mating*) NOT (("Quantitative trait loc*")OR(QTL*))
	(inbred SAME (offspring OR progeny))
	Selfed SAME out*
	"Optimal outcrossing" OR "Outcrossing distance"
	"Benefit* NEAR/5 dispersal"
	"Cost* NEAR/5 dispersal"
	("Natal dispersal" AND (inbred OR in*breeding OR heterosis OR self* OR fitness))
	(Philopatr* AND (inbred OR in*breeding OR heterosis OR self* OR fitness))
Fitness related strings	(Depression SAME in*bre*)
	(Depression SAME fitness)
	(Heterosis AND in*breeding)
	"Genetic load"
Effects of inbreeding related strings	Effect$ NEAR/5 in*breeding
	consequence$ NEAR/5 Inbreeding
	influence$ NEAR/5 inbreeding
	Outcome$ NEAR/5 inbreeding

Boolean syntax uses the Web of Science format, and changes were made, as appropriate, to adapt the search to the other two literature databases used (Scopus and JSTOR).

3.3.2 Inclusion and exclusion criteria

The inclusion and exclusion criteria were described in the chapter 2. However, there were additional criteria I applied in this meta-analysis, which are listed below.

- **Types of subjects**

Experimental populations derived from natural populations and kept in human-maintained environments without any artificial selection or bottlenecks were also included.

- **Types of phenotypic outcome**

  Articles were eliminated if an increase in trait value could not be reliably linked with a specific directional change in fitness. For example, there is no a priori basis for assigning different values of some traits, such as the concentration of chemical compounds measured in an individual as representing higher or lower fitness levels.

- **Sources of heterogeneity**

  In this study, I investigated the following sources of heterogeneity and their relationship with inbreeding depression effect sizes:

  - Taxonomic group: High-level taxonomic classification was recorded for species: invertebrate, plant and vertebrate.
  - Inbreeding coefficient *(F)*: Wright's *F,* describing the probability that two alleles are identical by descent was recorded for each inbred offspring (Wright, 1922).
  - Trait type: Categorial description of traits, one of behaviour, defence, development, fecundity, growth rate, physiology, size, survival, or viability.
  - Life history stage: Categorial description of when traits measured were timed relative to the study species life history. *Early stage* was defined when the traits were measured before morphological maturity or establishment. *Late stage* was defined when the traits were measured after sexual maturity was reached. Any trait recorded between early stage and late stage was categorised as *middle stage*.
  - Fitness class: Categorial description of the traits whether they are directly related to fitness or not. Fitness component traits were defined to include fitness, fecundity, survival, and viability. Alternatively, non-fitness components included behaviour, defence, development, growth rate, physiology, and size.
  - Types of population: Categorial description of the populations whether they represent natural populations or not based on the extent of human

intervention. If a population was obtained from the wild but was maintained in lab for less than 5 generations, we defined it as a natural population. Otherwise, it was defined as a lab population.

- Life expectancy: Description of the life expectancy of an organism. For animals, life expectancy was recorded in the unit of day; For plants, life expectancy was categorised as annual, biennial, perennial. If life expectancy was not recorded in articles or the relevant information could not be retrieved from literature search, I marked as unknown.
- Mating system 1: Two mating system schemes were recorded. Animals and plants were recorded separately in the first scheme. For animals, I recorded whether they mate with more than one partner during their life histories (polygamy), or only one (monogamy). For plants, we categorized their mating systems based on outcrossing rates measured from molecular markers. Mix-mating was defined when values were between 0.2 to 0.8 (Schemske & Lande, 1985; Winn *et al*., 2011). Any value larger or smaller than that was classified as outcrossing or inbreeding, respectively. If the mating system was not recorded in articles or the relevant information could not be retrieved from literature search, I marked as unknown.
- Mating system 2: For the second scheme of mating system, I categorized both plants and animals' sexual reproductive modes, one of mix-mating or outbreeding.
- Types of study: Categorial description of how natural the environmental conditions under which traits were measured. One of lab, greenhouse, common garden, or field.

3.3.3 Article screening strategy

The relevance of all articles in the literature database was assessed in a hierarchical way by first scanning titles, followed by the abstracts identified to have relevant titles and finally reading full-text. The article screening process was accelerated by the results of ACA at abstract level. Particularly, I utilised screening prioritisation techniques achieved by sorting the articles based on the topics identified by ACA, allowing researchers to evaluate the articles in the same field at the same time. I sorted the articles with high proportions of irrelevant topics to the top of the list and quickly

eliminated irrelevant articles by scanning their abstracts. (Chapter2, method section) The relevance of articles was assessed based on the inclusion criteria (types of population, types of phenotypic outcome, types of comparators and types of study) described in Chapter 2.

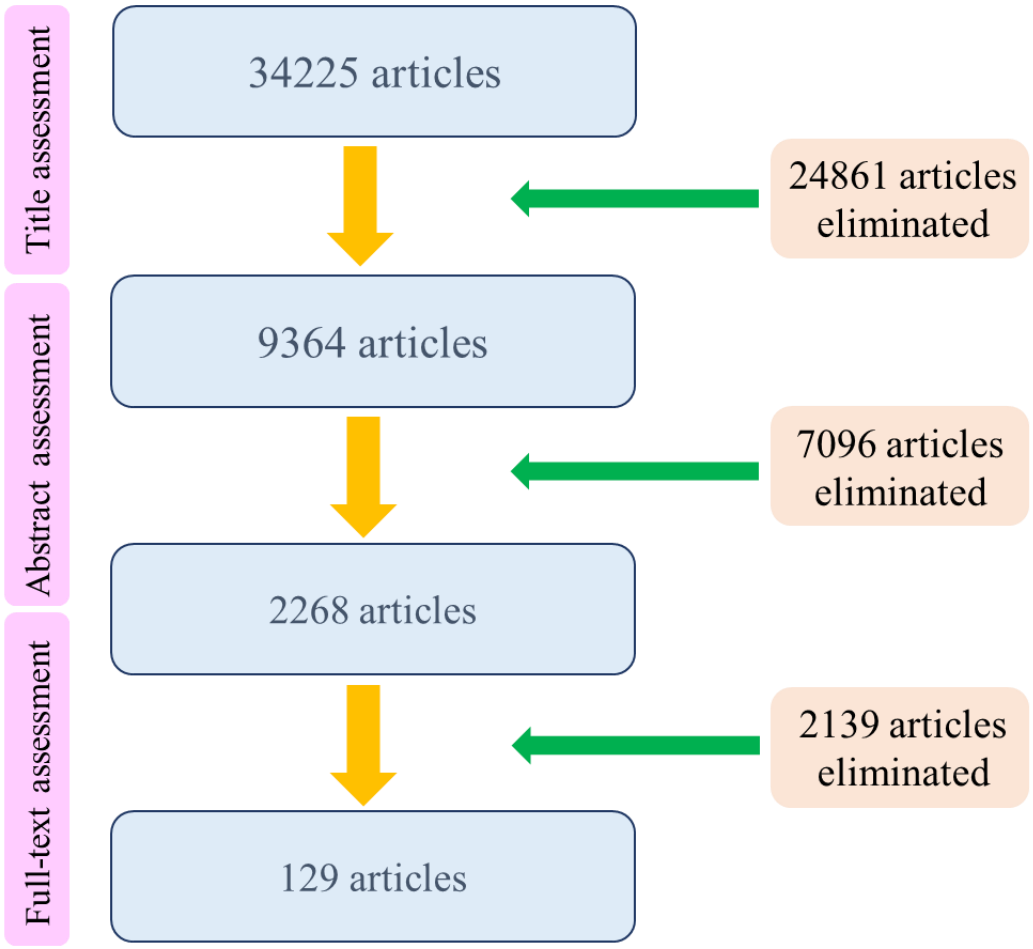


Fig. 1 Prisma diagram of article screening procedure showing the numbers of articles passed through title assessment, abstract assessment and full-text assessment.

3.3.4 Data extraction

I extracted data for calculation of effect sizes and for recording the sources of heterogeneity from relevant articles identified after the full-text assessment stage. I retrieved all values required to calculate inbreeding effect sizes for inbred and outbred offspring and their measurement error variance (*mev*) from relevant articles. Specifically, data required for effect size and *mev* calculation were phenotypic means of inbred and

outbred offspring, their corresponding sample sizes, and standard deviations. Details are described in the following paragraph.

Data for different traits or where inbred progeny had a differing inbreeding coefficient were retrieved respectively. Phenotypic means between inbred and outbred offspring for each trait were calculated for inbreeding effect sizes separately. Likewise, phenotypic means of inbred offspring with different inbreeding coefficient were compared with outbred offspring to calculate their inbreeding effect sizes. Each effect sizes derived above was considered as an individual study within an article. For continuously distributed traits, the mean phenotypic values measured for each trait were extracted for inbred and non-inbred groups separately. The corresponding sample sizes for phenotypic values were also recorded for each group. I retrieved associated standard deviation values directly from tables and main texts or calculated it from the standard error and sample size where necessary. For binomially distributed phenotypic values (e.g., counts) were recorded as a proportion or percentage. The accompanying standard deviations were retrieved directly or calculated with the equation below when its corresponding sample sizes were presented in the articles (Koricheva *et al*., 2013):

$$S(p) = \sqrt{Var(p)} = p(1-p)/n$$

Where p is the probability of success in n observations.

If the values required for the inbreeding effect sizes and error variance were presented in plots, WebPlotDigitizer (version 4.4) was used to estimate the raw values (Rohatgi, 2021). This software allows researchers to extract the raw values from figures (e.g. bar plots or scatter plots), by aligning the data points to X and Y axes. Phenotypic means can be extracted directly. Standard deviation can be calculated from the standard error by estimating the length of an error bar with WebPlotDigitizer.

The factors listed above that cause the sources of heterogeneity were also extracted either directly or via literature search. Life expectancy, mating system were sources of heterogeneity frequently required to retrieved from literature searches. Meanwhile, I approached the authors of the articles to request raw data when the values to calculate

inbreeding effect sizes and error variance were not presented in an article in the required form.

3.3.5 Calculation of effect sizes and measurement error variance

I calculated meta-analysis effect sizes for each study with log response ratios, using Bayesian mixed-effects models from MCMCglmm package (version 2.29). Study was defined as an individual inbreeding effect size for each comparison calculated from phenotypic means of inbred and outbred groups. The log response ratio was calculated as the effect size for each phenotypic trait with the following equation:

$$ES = -\ln(\mu_I/\mu_O) \quad (1)$$

where μ_I and μ_O are trait means for the inbred group and non-inbred group (outbred) respectively (Borenstein *et al*., 2009). The study measurement error variance was calculated as

$$mev = S_I^2\left(\frac{1}{n_I\mu_I^2}\right) + S_O^2\left(\frac{1}{n_O\mu_O^2}\right) \quad (2)$$

where n_I and n_o are the sample sizes of inbred offspring and non-inbred offspring and S_I and S_O are standard deviations of inbred group and non-inbred group respectively (Borenstein *et al*., 2009). The effect size measures inbreeding response, the deviation of inbred offspring's phenotypic values from the non-inbred comparators in each trait. Positive effect sizes indicate inbreeding depression whereas negative effect sizes indicate inbred groups have higher fitness values than outbred groups.

3.3.6 Meta-Analysis

The log response ratio effect sizes were analysed with the MCMCglmm package (version 2.29) in R, which allows researchers to fit effect size data within Bayesian generalised mixed modelling framework, via Markov chain Monte Carlo sampling (Hadfield, 2010; R Core Team, 2019). To control for within-study variation in effect size precision, *mev* was fitted in the models as a set of within study variance components using the *mev* argument. Variables fitted in meta-analyses were either fitted as random effects or fixed effects. The sources of heterogeneity listed above were fitted as fixed effects, representing the explanatory variables of primary interest in this study.

Article identity was fitted as a set of random effects in order to capture variation in effect sizes between articles and account for the nested structure of the data.

A total of 1.1×10^5 iterations were run for each meta-analysis model, with a burn-in of 1×10^4 iterations and then a following thin interval of 100 iterations where parameters were extracted. The sample size for the posterior distribution was 1000 for each fitted model. Priors for random effects were non-informative uniform improper distributions on the standard deviation of the random effects (Gelman, 2006). Pooled effects for each model were estimated by retrieving posterior mean estimates of intercepts or other relevant parameters along with their 95% credible intervals. Credible intervals indicate a posterior distribution or a predictive distribution that an unobserved data might fall with a particular probability. The posterior means estimates for parameters were considered as statistically significant when their corresponding credible intervals did not overlap with zero. I estimated Bayesian p-values, representing the probability of the parameter's location, by inspecting the proportion of the posterior distribution overlapping with zero (or with the point location for another parameter). Results were plotted as forest plots with their corresponding effect sizes and 95% credible intervals.

The relevance of predictor variables to explain inbreeding responses was examined in a stepwise manner. Specifically, the model was constructed first with a minimal model:

Inbreeding effect size ~ inbreeding coefficient

Inbreeding coefficients for inbred progeny are expected to be positively correlated with inbreeding effect sizes since the inbreeding coefficient quantifies the degree of inbreeding as the probability of identical by descent (Wright, 1922). All models retained the inbreeding coefficient as a fixed effects predictor. Subsequently, the sources of heterogeneity of interest were fitted as additional fixed effects. Model goodness of fit for these explanatory variables was assessed via the Deviance Information Criteria (DIC) (Spiegelhalter, Best, Carlin, & van der Linde, 2002). I considered the predictor variables were able to significantly explain the variation of inbreeding effect sizes when adding that variable caused a consistent decrease in the values of DIC over three replicate model runs.

3.4 Results

3.4.1 Descriptive statistics

The two phases of my literature searches in combination of Neaves' literature database yielded 29836 articles, published between 1941 and 2018. During this period, the numbers of articles related to phenotypic responses to inbreeding increased quickly until 2014. After that the numbers of articles decreased slightly but sustained over years (Fig. 2). The articles assessed as relevant for my meta-analysis at the full text level consisted of 129 research papers (containing 929 effect sizes), including 11 categories of phenotypic traits in both fitness component and non-fitness component traits. Fecundity and size were the most studied traits, which accounted for 27.9% and 26.1% of effect sizes respectively, followed by survival (15.9%). Conversely, defence and fitness were represented by a small proportion of effect sizes (2.2% and 3.1 % of effect sizes, respectively). The majority of studies in my meta-analysis used plants (66.5%) as study taxa, followed by invertebrates (28.9%). Vertebrates only accounted for 4.6% of total studies (Appendix 3).

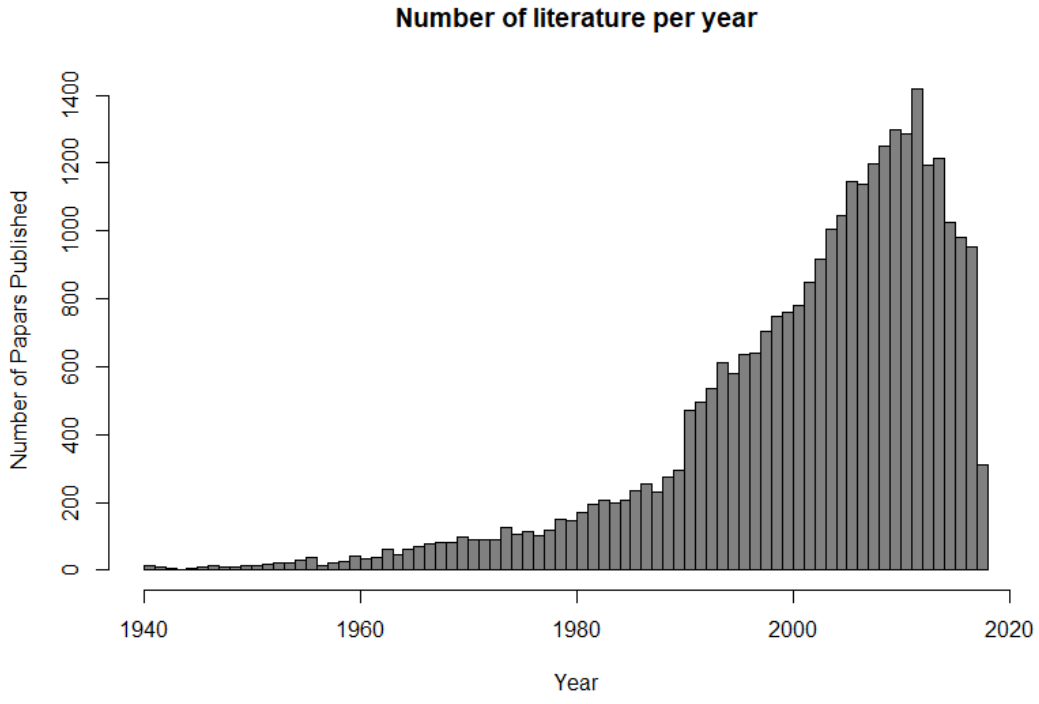


Fig. 2 The numbers of inbreeding-related articles retrieved by my literature search. The literature search was conducted in May, 2018 and therefore only a part of papers published in 2018 were retrieved.

3.4.2 Inbreeding coefficient and inbreeding effect sizes

In general, inbreeding depression effect sizes were significantly positive, indicating significant phenotypic or fitness costs during inbreeding and were linearly related to the inbreeding coefficient (posterior mean effect size = 0.244, pMCMC < 0.001, Fig. 3).

Predicted inbreeding effect sizes were consistently positive regardless of the level of the inbreeding coefficient, with credible intervals that did not overlap with zero at any point along the full range of the inbreeding coefficient (Fig. 3). Even when inbreeding coefficient is zero, I still detected significant inbreeding effect size (posterior mean effect size = 0.1, pMCMC = 0.015, Fig. 3). The estimated effect sizes were 0.332 and 0.216 when inbreeding coefficient was 1 and 0.5 respectively. 0.332 and 0.216 inbreeding effect sizes indicated 28 % and 19 % of fitness reduction of inbred offspring comparing to outbred offspring respectively.

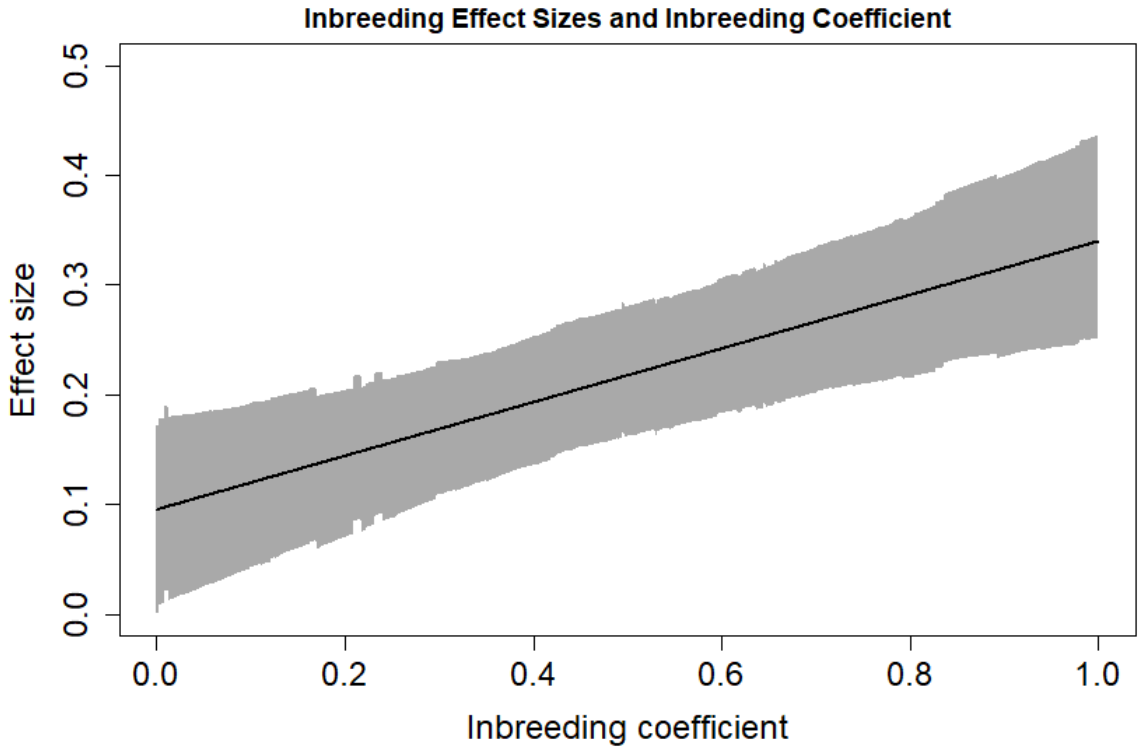


Fig. 3 The positive relationship between inbreeding effect sizes and inbreeding coefficient. The inbreeding effect sizes were calculated by the log response ratio of the phenotypic values of inbred offspring divided by the phenotypic values of outbred offspring multiplied by -1. The greater effect sizes indicate higher intensity of inbreeding depression. The shaded area represents the 95% credible zone for the regression line, indicating the zone within which the regression line is located with 95% probability. The relationship between inbreeding effect sizes and inbreeding coefficient can be expressed as follows: Inbreeding effect size = 0.099+0.233 x inbreeding coefficient.

Table 2. Summary of fitted meta-analyses containing fix-effect explanatory variables.

Model	Fixed effects	DIC	Among-article variance	Within-article variance	Within-study variance	Species
A1	~Intercept	-243.78	0.0864	0.0304	0.0722	All
A2	~*F*	-263.97	0.0949	0.0293	0.0722	All
A3	~*F* + Trait type	-269.47	0.0900	0.0291	0.0722	All
A4	~*F* * Trait type	-270.64	0.0912	0.0288	0.0722	All
B1	~Intercept	-243.78	0.0864	0.0304	0.0722	All
B2	~*F*	-263.97	0.0949	0.0293	0.0722	All
B3	~*F* + Fitness class	-258.43	0.0925	0.0296	0.0722	All
B4	~*F* * Fitness class	-256.27	0.0921	0.0298	0.0722	All
C1	~Intercept	-473.56	0.0968	0.0223	0.0672	All
C2	~*F*	-498.72	0.1044	0.0216	0.0672	All
C3	~*F* + Life history stage	-524.44	0.1064	0.0208	0.0672	All
C4	~*F* * Life history stage	-522.75	0.107	0.0208	0.0672	All
C5	~*F* + Life history stage * Mating system2	-526.40	0.105	0.0207	0.0672	All
D1	~Intercept	-337.69	0.0202	0.0220	0.0078	Plant
D2	~*F*	-359.34	0.0216	0.0211	0.0078	Plant
D3	~*F* + Life history stage	-385.14	0.0242	0.0201	0.0078	Plant
D4	~*F* + Life history stage + Mating system2	-381.98	0.0250	0.0201	0.0078	Plant
D5	~*F* + Life history stage * Mating system2	-385.15	0.0208	0.0200	0.0078	Plant
E1	~Intercept	-160.21	0.1982	0.0216	0.0041	Animal
E2	~*F*	-157.68	0.1952	0.0217	0.0041	Animal
E3	~*F* + Life history stage	-152.43	0.1969	0.0222	0.0041	Animal
F1	~Intercept	-243.78	0.0864	0.0304	0.0722	All
F2	~*F*	-263.97	0.0949	0.0293	0.0722	All
F3	~*F* + Taxonomic group	-265.01	0.0968	0.0295	0.0722	All
F4	~*F* * Taxonomic group	-259.26	0.0869	0.0296	0.0722	All
G1	~Intercept	-350.83	0.0189	0.0218	0.0083	Plant
G2	~*F*	-372.15	0.0207	0.0209	0.0083	Plant
G3	~*F* + Life span	-370.00	0.0210	0.0209	0.0083	Plant
H1	~Intercept	28.10	0.0481	0.0465	0.0054	Animal
H2	~*F*	29.03	0.0487	0.0465	0.0054	Animal

H3	~*F* + Life span	30.15	0.0501	0.0466	0.0054	Animal
I1	~Intercept	-315.16	0.0212	0.0189	0.0103	Plant
I2	~*F*	-326.29	0.0224	0.0183	0.0103	Plant
I3	~*F* + Mating system 1	-322.57	0.0224	0.0185	0.0103	Plant
J1	~Intercept	39.53	0.0495	0.0490	0.0051	Animal
J2	~*F*	40.89	0.0490	0.0492	0.0051	Animal
J3	~*F* + Mating system 1	41.71	0.0516	0.0489	0.0051	Animal
K1	~Intercept	-243.78	0.0864	0.0304	0.0722	All
K2	~*F*	-263.97	0.0949	0.0293	0.0722	All
K3	~*F* + Mating system 2	-265.92	0.0958	0.0295	0.0722	All

F: inbreeding coefficient; *Mating system 1* and *mating system 2* are the categorial description of the two types of mating system described in 3.3.2 Typical measurement error variance was estimated by the median of *mev (*Whitlock *et al*., 2013*).*

3.4.3 Phenotypic traits and inbreeding effect sizes

The effects of phenotypic trait types on inbreeding depression were analysed with 929 effect sizes. All types of phenotypic traits experienced significant inbreeding depression apart from behaviour (Fig. 4; posterior mean effect size for behaviour = -0.156, credible interval = -0.495 – 0.175, pMCMC = 0.346), defence (posterior mean effect sizes = 0.217, credible interval = 0.051 – 0.369, pMCMC = 0.004) and physiology (posterior mean effect sizes = 0.178, credible interval = 0.06 – 0.285, pMCMC = 0.002; Fig. 4). Fitness and fecundity were two phenotypic traits suffered the most from inbreeding depression (Fitness: posterior mean effect size = 0.339, credible interval = 0.186 – 0.477, pMCMC < 0.001; Fecundity: posterior mean effect size = 0.279, credible interval = 0.215 – 0.333, pMCMC < 0.001, Fig. 4). Adding phenotypic trait type as a predictor improved model goodness of fit (change in DIC = -5.4, Table. 2). Orthogonal contrasts indicated that fecundity experienced relatively severe of inbreeding depression comparing to other trait types except for defence, fitness and growth rate. Orthogonal contrasts also revealed that behaviour had significantly smaller inbreeding effect sizes compared with all other phenotypic traits.

Both fitness-component and non-fitness component traits suffered significant inbreeding depression. The estimated effect sizes for fitness component and non-fitness component traits were 0.225 and 0.205, respectively. However, I did not detect any difference between these two fitness class categories (pMCMC = 0.264). Adding this variable did not improve model goodness of fit (change in DIC = +5.54, Table. 2) Even

though, fitness component traits were defined as a combination of 4 trait types (fitness, fecundity, survival, and viability). The trait types with highest inbreeding effect sizes were fitness (0.339) and fecundity (0.281). The non-fitness related traits that are less susceptible to inbreeding depression compared to fecundity are only observed in behavioural (pMCMC = 0.436, for contrasts of fecundity with behavioural traits), defensive (pMCMC =0.004, for contrasts of fecundity with defensive traits) and physiological traits (pMCMC = 0.002, for contrasts of fecundity with physiological traits).

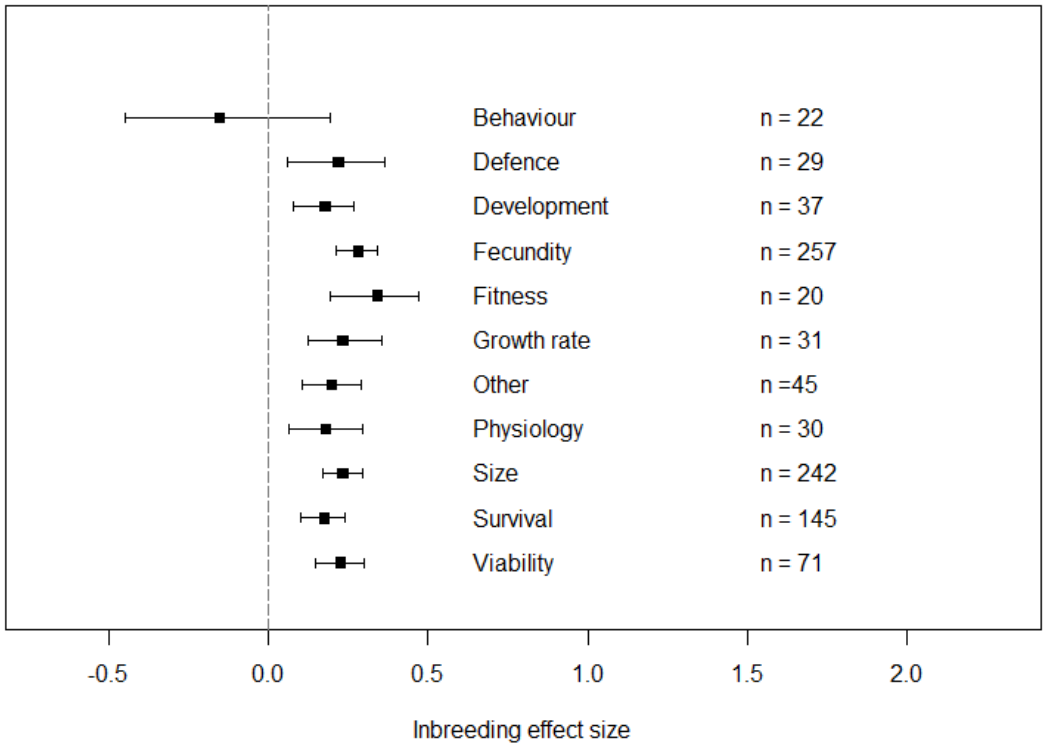


Fig. 4 Pooled effects summarising the variation of inbreeding depression in response to phenotypic trait types in animals and plants. Positive values indicate the expression of inbreeding depression. Error bars show 95 % of credible intervals. The effect size was considered significant if credible intervals did not overlap with zero. n indicates the numbers of studies.

3.4.4 Life history stages and inbreeding effect sizes

Variation in inbreeding depression with different life history stages was analysed using 875 effect sizes. Significant inbreeding depression (significantly positive pooled effect sizes) was detected in all investigated life history stages (Fig. 5). *Lat stage* suffered greatest inbreeding depression, with estimated effect sizes 0.253 (credible interval =

0.189 – 0.316, Fig. 5.). The level of inbreeding depression in *late stage* was significantly greater (29%) than the level of inbreeding depression in *early stage* and *middle stage* (both pMCMC < 0.001, Fig. 5). However, there was no significant difference of inbreeding depression between early stage and middle stage (pMCMC = 0.358). Adding the life history stage predictor improved the model goodness of fit (change in DIC = -25.72, Table. 5). Further analysis was conducted to investigate variation in inbreeding depression with life history stage in animals and plants separately. The results implied that the majority of variation in inbreeding effect sizes with life history stage stemmed from effects occurring in plants (Fig. 6; Fig. 7). Adding life history stage to the plant model improved model goodness of fit (change in DIC = -25.8, Table. 2). Conversely, the improvement in model goodness of fit was not detected in the animal dataset when the life history stage predictor was added.

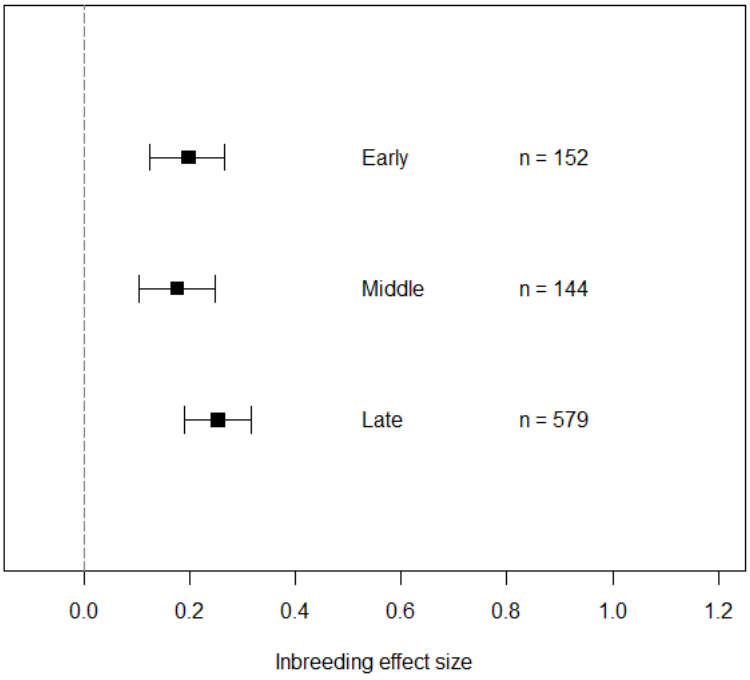


Fig. 5 Pooled effects summarising the variation of inbreeding depression with different life history stages in animals and plants. The *early stage* was defined when traits were measured before morphological maturity or establishment; the *late stage* was defined when the traits were measured after sexual maturity had been reached. Any trait measured between these two stages was classified as the *middle stage*. The interpretations of pooled effect estimates, credible intervals and n follow Fig. 4.

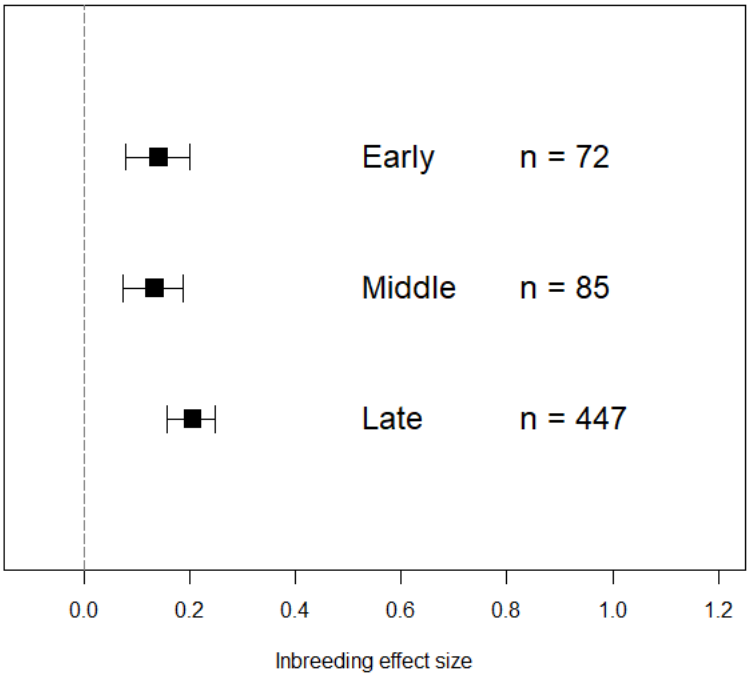


Fig. 6 Pooled effects summarising the variation of inbreeding depression with life history stages in plants. The definitions of each stage were set out in Fig. 5. The interpretations of pooled effect estimates, credible intervals and n follow Fig. 4.

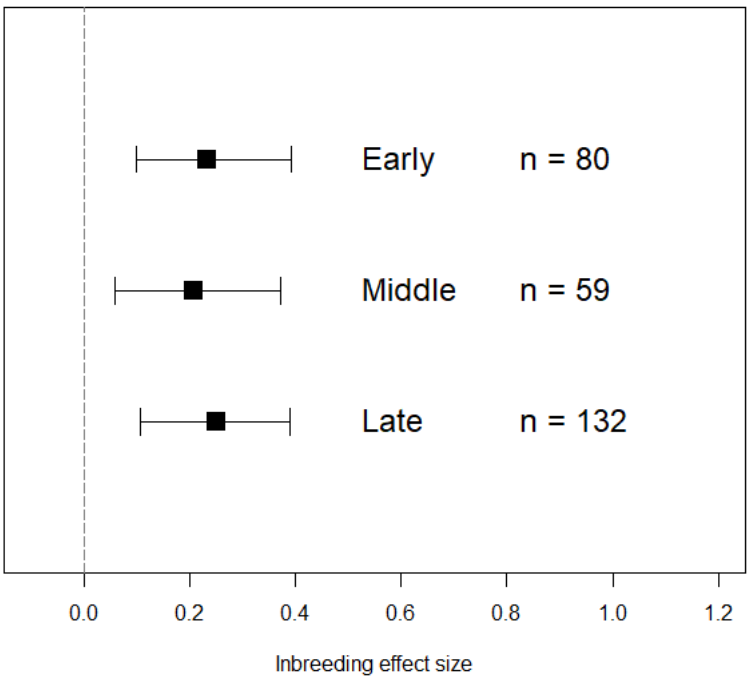


Fig. 7 Pooled effects summarising the variation of inbreeding depression with life history stages in animals. The definitions of each stage were set out in Fig. 5. The interpretations of pooled effect estimates, credible intervals and n follow Fig. 4.

3.4.5 Mating system

Variation in inbreeding depression with mating systems was analysed with 889 and 929 datapoints for definition 1 and definition 2 of mating system respectively. The definitions of *mating system 1* and *mating system 2* were described in the method section (sources of heterogeneity). Due to the information of *mating system 1* could not be retrieved by literature search for some articles (recorded as unknown), the number of the effect sizes for mating system 1 was smaller. Overall, the two definitions of mating system were not significant predictors of the inbreeding effect sizes. Neither predictor improved model goodness of fit (Table 2). Nevertheless, when the dataset was analysed separately for animals and plants, I detected the marginal variation of inbreeding depression among *mating system 2* in plants. Mix-mating plants and outcrossing plants differed marginally non-significantly in their response to inbreeding (pMCMC = 0.064) despite the fact that adding *mating system 2* did not improve model goodness of fit. Specifically, outcrossing plants had higher pooled estimated inbreeding effect size (0.308) than mix-mating plants (0.166). However, when *mating system 2* and life history stage were investigated simultaneously in plants, I discovered that not only was there significant variation in inbreeding depression among different life history stages, but that the pattern of this variation differed significantly between mix-mating and outcrossing species (Fig. 8). Adding an interaction term between *mating system 2* and life history stage improved the model goodness of fit (change in DIC = -25.81, Table. 2). Inbreeding depression was identified in mix-mating plant species and the magnitude of inbreeding depression in *late stage* was significantly larger than in *early* and *middle stages* (pMCMC < 0.001 and pMCMC = 0.002 for contrasts of *late* with *early* and *middle stages*, respectively; Fig. 8). This pattern was not observed in the strictly outcrossing plants (Fig. 8). Here, inbreeding depression was only detected in the *early stage* and its magnitude was significantly larger than the inbreeding depression observed in the *middle* and *late stages* even though the sample sizes were not large (pMCMC = 0.004 and pMCMC < 0.001 for contrasts of *early* with *middle stages*, respectively; Fig. 8).

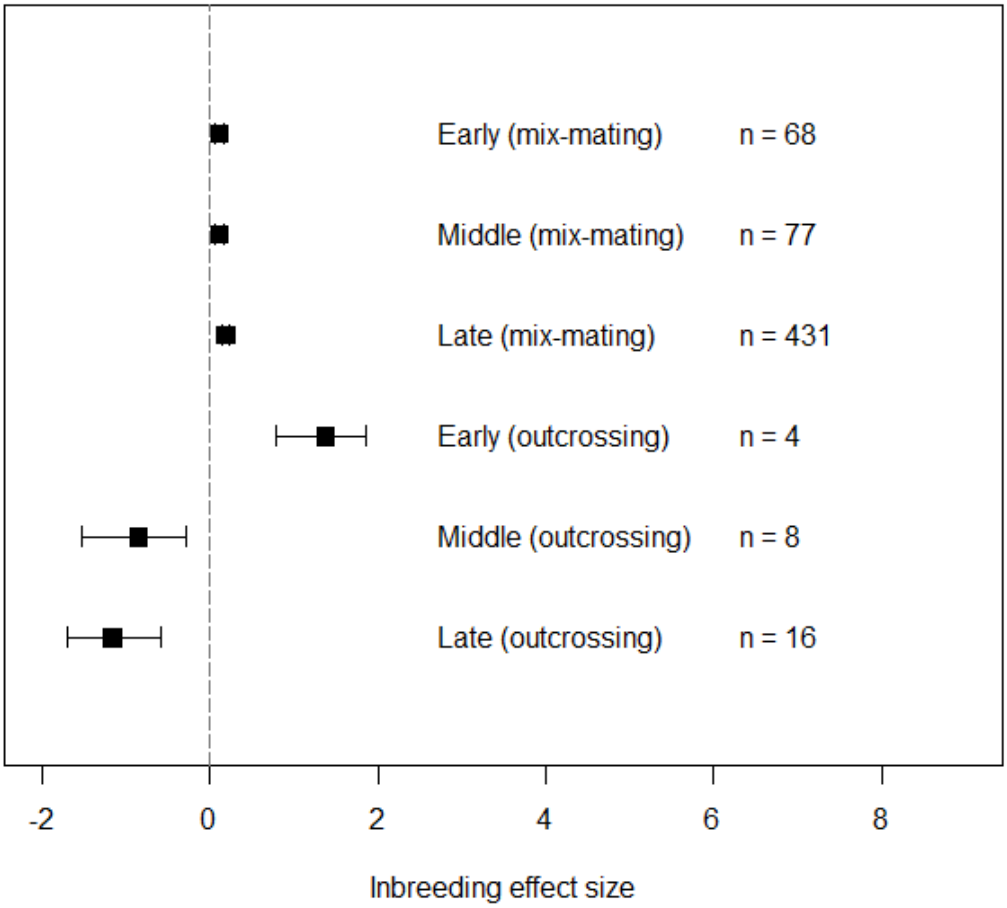


Fig. 8 Pooled effects summarising the variation of inbreeding depression in response to life history stage and *mating system 2* in plants. The definitions of each stage were set out in Fig. 5. The interpretations of pooled effect estimates, credible intervals and n follow Fig. 4.

3.4.6 Species characteristics and inbreeding effect sizes

Overall, neither taxonomic group nor organism life expectancy were significant predictors of inbreeding effect sizes. Including taxonomic group improved model goodness of fit marginally (change in DIC = -1.04, Table. 2). Invertebrate, plant, and vertebrate all suffered from inbreeding depression since the posterior mean of effect sizes for these group differ significantly from zero (posterior mean effect size: 0.2455 for invertebrate, 0.1929 for plant and 0.224 for vertebrate; pMCMC: <0.001 for invertebrate and plant, 0.056 for vertebrate; Fig. 9).

Life expectancy was analysed in animals and plants separately and was not a significant predictor of inbreeding effect sizes in either plants or animals. Including life expectancy as a predictor did not improve model goodness of fit (change in DIC = +2.1 and +1.47 for plants and animals, respectively, Table. 2).

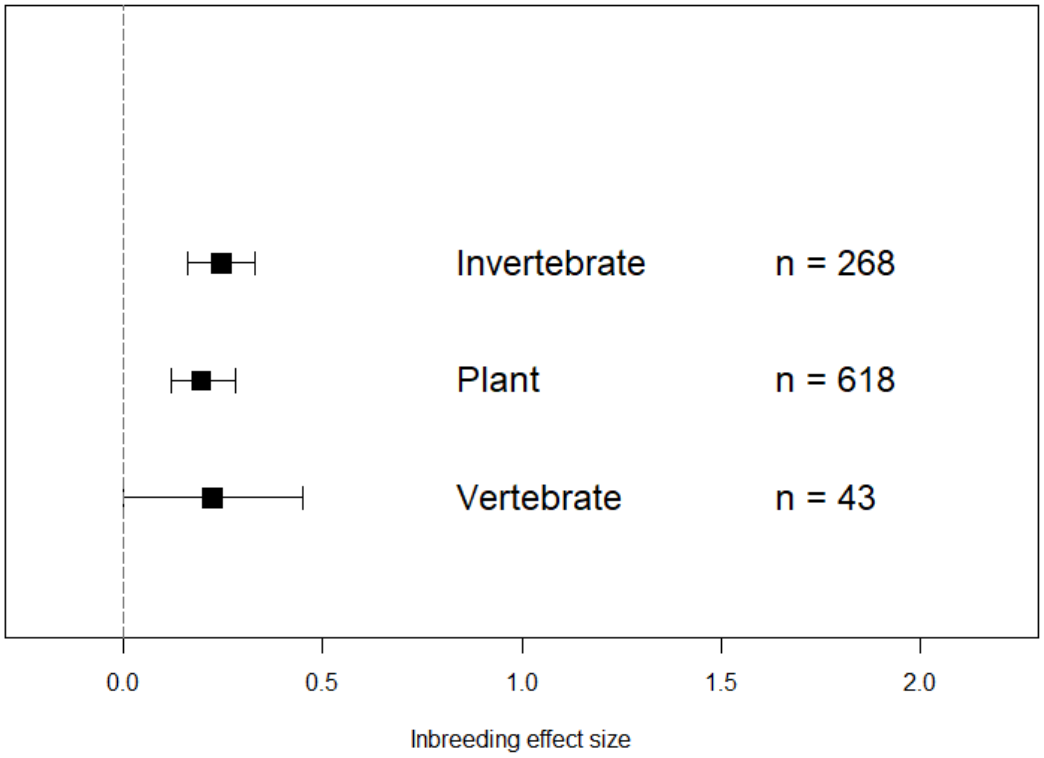


Fig. 9 Pooled effects summarising the variation of inbreeding depression in response to taxonomic groups. The interpretations of pooled effect estimates, credible intervals and n follow Fig. 4.

3.4.7 Heterogeneity in inbreeding response

The between-article variance was 0.0864 in the minimal model and accounted for 48.71% of total heterogeneity in inbreeding response (Table. 2). The between-study variance was 0.0304, which accounted for 16.08% of total heterogeneity. The remaining heterogeneity (38.2%) was attributed to within-study variance or *mev*.

3.5 Discussion

In this meta-analysis, I have investigated the variation in inbreeding depression with different sources of heterogeneity, including types of phenotypic traits, life history

stages and mating systems. However, while my results show that inbreeding depression is ubiquitous, the sources of heterogeneity that I investigated explain variation in responses to inbreeding in only a small number of cases. Only phenotypic traits and life history stages were significant predictors of inbreeding depression. The interaction between life history stage and *mating system 2* was detected in plants.

3.5.1 Inbreeding coefficient and inbreeding effect sizes

Inbreeding depression is expected to be absent in outbred offspring (F=0). Though, I still detected significant inbreeding effect sizes when the inbreeding coefficient was zero. A consequence of fitting a simple linear model was that the regression line was not constrained to go pass through the origin, where inbreeding depression is zero. As a result, my analysis implies significant inbreeding depression when inbreeding coefficient is zero. However, this does not mean that there necessarily is inbreeding depression when there is no inbreeding. The possible reason might be a non-linear relationship between the inbreeding coefficient and inbreeding effect sizes when the inbreeding coefficient is below 0.25. However, I did not test the non-linear relationship within this region since the paucity of data points of inbreeding coefficient below 0.125 (Fig. 3). The lowest inbreeding coefficient in my dataset is 0.125, making it challenging to capture the relationship when the inbreeding coefficient is close to zero.

3.5.2 The relationship between phenotypic traits and inbreeding depression

In 1989, an observation made by Falconer among multiple studies showed that inbreeding depression was much more severe in fitness-related traits (e.g. reproduction, survival) compared to the traits that were less related to fitness, such as morphological traits (Falconer, 1989). The mechanisms behind this assumption involve the interaction between natural selection and the genetic architecture within traits (DeRose & Roff, 1999; Husband & Schemske, 1996). Specifically, life-history traits (those closely related to fitness) are more vulnerable to inbreeding depression since these traits are predominantly determined by dominance genetic variance, whereas morphological traits are largely underpinned by additive genetic variance (Roff, 1998). Traits determined by additive genetic variance are typically under weak directional or stabilising selection (Lynch & Walsh, 1998). Changing phenotypic means in these traits results

in selectively neutral, making them less susceptible to inbreeding depression. However, my meta-analysis showed that fitness-component traits—defined to include the trait types of fitness, fecundity, survival, and viability—did not suffer more inbreeding depression, on average, than other traits that are less directly related to fitness. Two of the fitness component traits (fitness and fecundity) experienced the greatest inbreeding depression compared with any other trait type and suffered significantly greater costs on inbreeding than the following non-fitness component traits: behavioural, developmental, and physiological traits. This result suggested that fitness component traits harbour higher dominance variance than non-fitness component traits might not be a universal pattern. This point of view is further supported by a narrative review stating that fitness component traits seemed to have higher additive genetic variation than we expected before (Merila & Sheldon, 1999).

I found that the inbreeding response for behaviour was significantly lower than for other traits. The possible explanation for this result might derive from the genetic architecture underlying behavioural traits. The majority of behavioural traits measured in my meta-analysis are related to mating and courtship. These types of traits were reported to consist of multiple behavioural components (Arbuthnott, 2009; Mackay *et al*., 2005). For example, courtship in Drosophila involve several distinct behavioural traits, such as wing display, body shaking and circling, which are mediated by different loci. A deleterious recessive allele in one of these loci might have relatively less impact on the phenotype if other behavioural components still function properly, consequently limiting the expression of inbreeding depression. Additionally, behaviour is believed to have a large degree of plasticity, enabling it to flexibly respond to environmental variation independently of genetic architecture (Reale, Dingemanse, Kazem, & Wright, 2010; Reale & Roff, 2002). As a result, the lack an overall signal of inbreeding depression in behavioural traits could be due to the behavioural compensation that mitigates genetically suboptimal phenotypes (Reale & Roff, 2002). Furthermore, behavioural traits are proposed to be maintained by balancing selection (Boon, Reale, & Boutin, 2007; Dingemanse, Kazem, Reale, & Wright, 2010; Wolf, van Doorn, Leimar, & Weissing, 2007). Traits undergo stabilising selection for an intermediate optimum, resulting in weakly directional dominance will be less vulnerable to inbreeding

depression because recessive alleles that cause lead to shifts in trait mean in either direction would be selectively neutral (Lynch & Walsh, 1998).

Whilst the results presented here do not allow the mechanisms cause the overall lack of inbreeding depression in behavioural traits to be identified, they do highlight possible differences in genetic architecture and plasticity in this trait that may underpin the observed variation of inbreeding depression. Further research could utilise genetic sequencing, quantitative genetic approaches and modern molecular approaches to detect the existence of balancing selection or deleterious alleles (Charlesworth & Willis, 2009). This could improve our understanding the genetic mechanisms underlying the heterogeneity of inbreeding depression among phenotypic traits.

3.5.3 The relationship among life history stages, mating system and inbreeding depression

My meta-analysis revealed that on average, there was more severe inbreeding depression in late life history stages than in *early* and *middle* life history stages. Further investigation indicated there was an interaction between life history stage and *mating system 2* and this pattern was more prominent in mixed-mating plants (defined by my *mating system 2*). This result is partially consistent with the results presented in studies by Angeloni *et al*. (2011) and Husband & Schemske (1996) despite different classifications of mating systems among these studies and mine. In Angeloni and co-workers' (2011) paper, they classified the mating system of plants as either self-compatible or self-incompatible, while Husband & Schemske's paper (1996) categorised the mating system of plants based on estimation from the segregation of molecular marker. In Angeloni's meta-analysis, they found lowest inbreeding depression in early life history stages (germination) in self-compatible species. Likewise, Husband & Schemske's studies showed high levels of inbreeding depression in the late life history stage in predominantly selfing species. This phenomenon can be explained by evolutionary theory, which predicts that purging is less efficient against recessive deleterious alleles that are expressed late in life history due to a lower contribution to reproduction late in the life history (Hamilton, 1966; Williams, 1957). As a result, purging against late-life history mutations should be inefficient, making these traits more vulnerable to inbreeding depression.

However, substantial differences were observed between the results of my study and previous studies in the pattern of inbreeding depression in different life history stages in strictly outcrossing plants. Husband & Schemske's studies (1996) found high levels of inbreeding depression in late life history stages (seed production and reproduction) compared to early life history stages (germination) in predominately outcrossing plants whereas Angeloni's meta-analysis (2011) shows similar levels of inbreeding depression among different life history stages in self-incompatible species. My result implied that *early stages* suffered significantly high inbreeding depression while *middle* and *late stages* experienced inbreeding benefits. These inconsistencies might stem from the difference in definitions of life history stages between my meta-analysis and other's studies. Both Angeloni and Husband & Schemske's studies classified growth and survival as individual categories However, there two traits could be measured at early, middle, or late life history stages, resulting in the imprecise categories in their analysis. Additionally, inconsistencies could arise from low statistical power due to the statistical methods that researchers employed or low sample sizes. Both Angeloni's and my meta-analysis suffered from low sample sizes for outcrossing plants compared to the sample sizes we retrieved from mix-mating plants. Furthermore, Husband & Schemske's study utilised Mann-Whitney U-tests, which did not control for within study variation in sample size and effect size precision. Expanding the meta-analysis to include more outcrossing species might help to improve our understanding on the interaction between mating systems and life history stages. Because of the lack of opportunity to expose deleterious alleles to selection, the absence of variation in inbreeding depression in outcrossing plants was expected.

In comparison to Angeloni's and Husband & Schemske's studies, which focussed exclusively on plants, my meta-analysis provides a further insight into the impacts of inbreeding depression at different life history stages in animals. My results showed significant effects of inbreeding depression on all life history stages but there was no observable variation among in the intensity of inbreeding depression between them. The absence of differences between *early* and *late stages* in animals might indicate that purging acted to the same degree in early and late stages of animals. Purging requires the exposure of deleterious alleles to natural selection (Byers & Waller, 1999;

Charlesworth & Charlesworth, 1999). However, the majority of animals in my dataset are naturally out-mating (79.7 %), implying the potential deficiency in inbreeding events before experiments started. Therefore, the opportunity to expose deleterious alleles to natural selection might be rare in animals that were studies in my meta-analysis, resulting in the lack of difference in inbreeding depression.

3.5.4 The relationship between species characteristics (life expectancy and taxonomic group) and inbreeding depression4

Majority of studies that investigated the effects of species characteristics on inbreeding depression were focused on plants (Angeloni *et al*., 2011; Duminil *et al*., 2009; Morgan, 2001; Scofield & Schultz, 2006). Little attention has been paid to sources of heterogeneity for inbreeding depression in animals. My meta-analysis is the first study to explore the relationship between species characteristics and inbreeding depression in wide range of taxonomic groups. In plants, it is widely acknowledged that small-statured (short-lived) species have lower inbreeding depression than large-statured (typically longer-lived) species (Duminil *et al*., 2009; Scofield & Schultz, 2006). The underlying hypothesis explaining this difference may be the increase in the opportunity for the accumulation of deleterious alleles during the lifetime of long-lived, large-statured plant species through somatic mutations in meristems that give rise to reproductive tissues without a concurrent increase in the effectiveness of selection (Duminil *et al*., 2009; Scofield & Schultz, 2006). Thus, their offspring are more likely to display inbreeding depression than small-statured plants since large-statured plants have more opportunity to accumulate genetic load and a lower efficiency at purging it, due to long generation times. If this is the general explanation behind the variation of inbreeding depression in response to different life-forms, we could also assume that long-lived plants might suffered more inbreeding depression than short-lived plants.

However, my result didn't reveal any heterogeneity in inbreeding depression among different life expectancy or taxonomic groups, which is contradictory to our previous understanding. The lack of the variation between life-forms or life expectancy between inbreeding depression in my result might reflect the difference in mutation rates among different life-forms. Analysing the DNA sequence between woody perennial and

annual plants, Xie and co-workers (2016) detected lower mutation rate per unit time in woody perennials. As a result, with lower mutation rates in long-statured plants, it's possible to expect that inbreeding depression in long-lived species might not be more severe than that in shorter-lived species. Additionally, a model conducted by Lesaffre and Billiard (2021) found that inbreeding depression varied weakly with respect to life expectancy. Their model incorporated a physiological growth model and multilocus genetic approaches, considering how age-structure influenced fitness. For example, fecundity typically scales positive with age in perennial plants (Franco & Silvertown, 1996). Mutations deterring plants' growth could therefore deteriorate inbreeding depression due to the growth delays could impact their fecundity negatively. As a result, the theory which predicts long-lived species suffers more inbreeding depression only occurs within a narrow range of conditions in their model. They attributed this outcome to what traits were measured and how strong is the selection.

The analyses presented here have larger data sizes and more rigorous statistical methods compared to any of the previous meta-analysis studies. Thus, I argue that the lack of variance in inbreeding depression in my results might better reflect the general relationship between inbreeding depression and sources of heterogeneity in natural populations. The failure to find variation could be attributed to the magnitude of inbreeding depression was compound by the intensity of selection and what trait it was measured. For example, long life expectancy species might suffer less fitness reduction if inbreeding depression is measured in behavioural traits or traits in early life history stage. This perspective is further reinforced by the models demonstrating that inbreeding depression is independent of the strength of selection when selection is ineffective or populations are small (Bataillon & Kirkpatrick, 2000; Roze, 2015). In small populations, the effects of drift outweigh the effects of natural selection. When selection is ineffective, we might fail to detect the variation of inbreeding depression since all genetic loads are retain across populations.

3.5.5 Conclusion

Inbreeding depression has been recognised as a major threat for population persistence in small and fragmented populations. Understanding the heterogeneity of inbreeding

depression with respect to trait-specific, population-specific, and species-specific effects is of fundamental importance in conservation management when scientists try to translate the consequences of inbreeding into practical conservation management strategies to maximise population persistence. The inbreeding responses observed in my meta-analysis do not relate directly to demographic sustainability in natural populations. However, this review provides insights into this understanding by quantifying the phenotypic and fitness consequences in inbreeding depression related to a wide range of potential sources of heterogeneity. The inbreeding effects on the trait types fitness and fecundity, which experienced the most severe costs on inbreeding, are predicted to threaten population persistence by decreasing population growth and reproductive output (Crnokrak & Roff, 1999; Keller & Waller, 2002).

The available evidence synthesised by my meta-analysis has several implications for conservation. First, the phenotypic costs of inbreeding depression are a ubiquitous phenomenon. Inbreeding effect sizes were significant in most of the cases irrespective of traits, life history stages, taxonomic groups, and other predictors that I investigated. As a result, inbreeding depression should be considered as an intrinsic component of population management in conservation, regardless of taxonomic group and other contextual factors. Second, despite the fact that inbreeding depression was found to be ubiquitous, its magnitude could vary considerably depending on the trait types or life history stages that were measured. Therefore, evaluating inbreeding consequences for population sustainability with single traits could lead to biases and an over- or underestimation of demographic consequences. Studies which assess the extinction risks posed by inbreeding depression should always incorporate traits and life history stages in quantitative way. Lastly, the general paucity of variation in inbreeding depression in relation to species characteristics doesn't mean conservation practitioners should ignore these factors when they design conservation measurements. The failure to find differences in inbreeding depression regarding species characteristics is likely to indicate that the magnitude of inbreeding depression is complicated by the strength of natural selection and the genetic architecture underpinning the traits that are in focus.

Despite the efforts of my work to generalise heterogeneity in inbreeding responses, the scope of my systematic review is still limited to the specific factors I investigated.

Specifically, getting the best predictors for sources of heterogeneity is limited by the way articles report information on these factors. For some species the necessary data were absent. This limited sample sizes and limited the resolution of the predictors that represented sources of heterogeneity, making it less difficult to detect hypothesised differences. In addition, there are still some factors that were not examined in my meta-analysis, which could potentially influence the magnitude of inbreeding depression. For example, population size and demographic history were reported to influence the phenotypic outcomes of inbreeding depression (Frankham, 1998; Frankham *et al*., 1999; Michaels, Shi, & Mitchell, 2008; Ouborg & Vantreuren, 1994). However, the studies with this information are quite rare partly due to the required time and resources to document these variables is generally large (for each effect size a series of other papers may need to be examined). Therefore, I would suggest that where resources allow, future studies should strive to record in as much detail as possible the factors that may lead to the variation in inbreeding response. The key sources of variation include phenotypic trait, life history stage when traits are measured, mating system and perhaps population size. This would allow future researchers to model this heterogeneity in meta-analyses and provide a clearer understanding on the phenotypic and fitness outcomes of inbreeding to improve conservation.

3.6 Reference

Angeloni, F., Ouborg, N. J., & Leimu, R. (2011). Meta-analysis on the association of population size and life history with inbreeding depression in plants. *Biological Conservation, 144*(1), 35-43. doi:10.1016/j.biocon.2010.08.016

Arbuthnott, D. (2009). The genetic architecture of insect courtship behavior and premating isolation. *Heredity, 103*(1), 15-22. doi:10.1038/hdy.2009.22

Bataillon, T., & Kirkpatrick, M. (2000). Inbreeding depression due to mildly deleterious mutations in finite populations: size does matter. *Genetical Research, 75*(1), 75-81. doi:10.1017/s0016672399004048

Boon, A. K., Reale, D., & Boutin, S. (2007). The interaction between personality, offspring fitness and food abundance in North American red squirrels. *Ecology Letters, 10*(11), 1094-1104. doi:10.1111/j.1461-0248.2007.01106.x

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*: John Wiley & Sons, Ltd.

Byers, D. L., & Waller, D. M. (1999). Do plant populations purge their genetic load? Effects of population size and mating history on inbreeding depression. *Annual Review of Ecology and Systematics, 30*, 479-513. doi:10.1146/annurev.ecolsys.30.1.479

Caro, T. M., & Laurenson, M. K. (1994). Ecological and genetic-factors in

conservation - a cautionary tale. *Science, 263*(5146), 485-486. doi:10.1126/science.8290956
Charlesworth, B., & Charlesworth, D. (1999). The genetic basis of inbreeding depression. *Genetical Research, 74*(3), 329-340. doi:10.1017/s0016672399004152
Charlesworth, D. (2003). Effects of inbreeding on the genetic diversity of populations. *Philosophical Transactions of the Royal Society B-Biological Sciences, 358*(1434), 1051-1070. doi:10.1098/rstb.2003.1296
Charlesworth, D., & Charlesworth, B. (1987). Inbreeding depression and its evolutionary consequences. *Annual Review of Ecology and Systematics, 18*, 237-268. doi:10.1146/annurev.ecolsys.18.1.237
Charlesworth, D., & Willis, J. H. (2009). Fundamenal concepts in genetics the genetics of inbreeding depression. *Nature Reviews Genetics, 10*(11), 783-796. doi:10.1038/nrg2664
Crnokrak, P., & Roff, D. A. (1999). Inbreeding depression in the wild. *Heredity, 83*, 260-270. doi:10.1038/sj.hdy.6885530
DeRose, M. A., & Roff, D. A. (1999). A comparison of inbreeding depression in life-history and morphological traits in animals. *Evolution, 53*(4), 1288-1292. doi:10.2307/2640831
Dingemanse, N. J., Kazem, A. J. N., Reale, D., & Wright, J. (2010). Behavioural reaction norms: animal personality meets individual plasticity. *Trends in Ecology & Evolution, 25*(2), 81-89. doi:10.1016/j.tree.2009.07.013
Duminil, J., Hardy, O. J., & Petit, R. J. (2009). Plant traits correlated with generation time directly affect inbreeding depression and mating system and indirectly genetic structure. *Bmc Evolutionary Biology, 9*. doi:10.1186/1471-2148-9-177
Falconer, D. S. (1989). *An introduction to quantitative genetics* (3 rd ed.): Wiley.
Franco, M., & Silvertown, J. (1996). Life history variation in plants: An exploration of the fast-slow continuum hypothesis. *Philosophical Transactions of the Royal Society B-Biological Sciences, 351*(1345), 1341-1348. doi:10.1098/rstb.1996.0117
Frankham, R. (1998). Inbreeding and extinction: Island populations. *Conservation Biology, 12*(3), 665-675. doi:10.1046/j.1523-1739.1998.96456.x
Frankham, R., Ballou, J. D., & Briscoe, D. A. (2010). *Introduction to conservation genetics* (2 nd ed.): Cambridge University press.
Frankham, R., Lees, K., Montgomery, M. E., England, P. R., Lowe, E. H., & Briscoe, D. A. (1999). Do population size bottlenecks reduce evolutionary potential? *Animal Conservation, 2*(4), 255-260. doi:10.1111/j.1469-1795.1999.tb00071.x
Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models(Comment on an Article by Browne and Draper). *Bayesian Analysis, 1*(3), 515-533. doi:10.1214/06-ba117a
Hadfield, J. D. (2010). MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. *Journal of Statistical Software, 33*(2), 1-22. doi:10.18637/jss.v033.i02
Hamilton, W. D. (1966). The moulding of senescence by natural selection. *Journal of Theoretical Biology, 12*(1), 12-45. doi:https://doi.org/10.1016/0022-

5193(66)90184-6
Hedrick, P. W., & Garcia-Dorado, A. (2016). Understanding Inbreeding Depression, Purging, and Genetic Rescue. *Trends in Ecology & Evolution, 31*(12), 940-952. doi:10.1016/j.tree.2016.09.005
Hedrick, P. W., & Kalinowski, S. T. (2000). Inbreeding depression in conservation biology. *Annual Review of Ecology and Systematics, 31*, 139-162. doi:10.1146/annurev.ecolsys.31.1.139
Hoffmann, A. A., & Sgro, C. M. (2011). Climate change and evolutionary adaptation. *Nature, 470*(7335), 479-485. doi:10.1038/nature09670
Husband, B. C., & Schemske, D. W. (1996). Evolution of the magnitude and timing of inbreeding depression in plants. *Evolution, 50*(1), 54-70. doi:10.2307/2410780
Keller, L. F., & Waller, D. M. (2002). Inbreeding effects in wild populations. *Trends in Ecology & Evolution, 17*(5), 230-241. doi:10.1016/s0169-5347(02)02489-8
Koricheva, K., Gurevitch, J., & Mengersen, K. (2013). *Handbook of Meta-analysis in Ecology and Evolution* Princeton University Press.
Kristensen, T. N., Pedersen, K. S., Vermeulen, C. J., & Loeschcke, V. (2010). Research on inbreeding in the 'omic' era. *Trends in Ecology & Evolution, 25*(1), 44-52. doi:10.1016/j.tree.2009.06.014
Lande, R., & Schemske, D. W. (1985). The evolution of self-fertilization and inbreeding depression in plants .1. genetics models. *Evolution, 39*(1), 24-40. doi:10.2307/2408514
Lesaffre, T., & Billiard, S. (2021). On Deleterious Mutations in Perennials: Inbreeding Depression, Mutation Load, and Life-History Evolution. *American Naturalist, 197*(5), E143-E155. doi:10.1086/713499
Lynch, M., & Walsh, B. (1998). *Genetics and Analysis of Quantitative Traits*: Sinauer.
Mackay, T. F. C., Heinsohn, S. L., Lyman, R. F., Moehring, A. J., Morgan, T. J., & Rollmann, S. M. (2005). Genetics and genomics of Drosophila mating behavior. *Proceedings of the National Academy of Sciences of the United States of America, 102*, 6622-6629. doi:10.1073/pnas.0501986102
Merila, J., & Sheldon, B. C. (1999). Genetic architecture of fitness and nonfitness traits: empirical patterns and development of ideas. *Heredity, 83*, 103-109. doi:10.1046/j.1365-2540.1999.00585.x
Michaels, H. J., Shi, X. J., & Mitchell, R. J. (2008). Effects of population size on performance and inbreeding depression in Lupinus perennis. *Oecologia, 154*(4), 651-661. doi:10.1007/s00442-007-0861-5
Morgan, M. T. (2001). Consequences of life history for inbreeding depression and mating system evolution in plants. *Proceedings of the Royal Society B-Biological Sciences, 268*(1478), 1817-1824. doi:10.1098/rspb.2001.1741
Neaves, L. E., Eales, J., Whitlock, R., Hollingsworth, P. M., Burke, T., & Pullin, A. S. (2015). The fitness consequences of inbreeding in natural populations and their implications for species conservation - a systematic map. *Environmental Evidence, 4*(1). doi:10.1186/s13750-015-0031-x
Ouborg, N. J., & Vantreuren, R. (1994). The significance of genetic erosion in the process of extinction .4. Inbreeding load and heterosis in relation to population-size in the mint Salvia-pratensis. *Evolution, 48*(4), 996-1008. doi:10.2307/2410361

Reale, D., Dingemanse, N. J., Kazem, A. J. N., & Wright, J. (2010). Evolutionary and ecological approaches to the study of personality. *Philosophical Transactions of the Royal Society B-Biological Sciences, 365*(1560), 3937-3946. doi:10.1098/rstb.2010.0222

Reale, D., & Roff, D. A. (2002). Quantitative genetics of oviposition behaviour and interactions among oviposition traits in the sand cricket. *Animal Behaviour, 64*, 397-406. doi:10.1006/anbe.2002.3084

Reed, D. H., & Frankham, R. (2003). Correlation between fitness and genetic diversity. *Conservation Biology, 17*(1), 230-237. doi:10.1046/j.1523-1739.2003.01236.x

Roff, D. A. (1998). Effects of inbreeding on morphological and life history traits of the sand cricket, Gryllus firmus. *Heredity, 81*, 28-37. doi:10.1038/sj.hdy.6883630

Rohatgi, A. (2021). WebPlotDigitizer. Retrieved from https://automeris.io/WebPlotDigitizer

Roze, D. (2015). Effects of Interference Between Selected Loci on the Mutation Load, Inbreeding Depression, and Heterosis. *Genetics, 201*(2), 745-+. doi:10.1534/genetics.115.178533

Schemske, D. W., & Lande, R. (1985). The evolution of self-fertilization and inbreeding depression in plants .2. Empirical observations. *Evolution, 39*(1), 41-52. doi:10.1111/j.1558-5646.1985.tb04078.x

Scofield, D. G., & Schultz, S. T. (2006). Mitosis, stature and evolution of plant mating systems: low-Phi and high-Phi plants. *Proceedings of the Royal Society B-Biological Sciences, 273*(1584), 275-282. doi:10.1098/rspb.2005.3304

Spiegelhalter, D. J., Best, N. G., Carlin, B. R., & van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society Series B-Statistical Methodology, 64*, 583-616. doi:10.1111/1467-9868.00353

Team, R. C. (2019). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/.

Wang, J. L., Hill, W. G., Charlesworth, D., & Charlesworth, B. (1999). Dynamics of inbreeding depression due to deleterious mutations in small populations: mutation parameters and inbreeding rate. *Genetics Research, 74*(2), 165-178. doi:10.1017/s0016672399003900

Whitlock, R., Stewart, G. B., Goodman, S. J., Piertney, S. B., Butlin, R. K., Pullin, A. S., & Burke, T. (2013). A systematic review of phenotypic responses to between-population outbreeding. *Environmental Evidence, 2*(1), 13. doi:10.1186/2047-2382-2-13

Williams, G. C. (1957). Pleiotropy, natural-selection, and the evolution of senescence. *Evolution, 11*(4), 398-411. doi:10.1111/j.1558-5646.1957.tb02911.x

Winn, A. A., Elle, E., Kalisz, S., Cheptou, P. O., Eckert, C. G., Goodwillie, C., . . . Vallejo-Marin, M. (2011). Analysis of inbreeding depression in mixed-mating plants provides evidence for selective interference and stable mixed mating. *Evolution, 65*(12), 3339-3359. doi:10.1111/j.1558-5646.2011.01462.x

Wolf, M., van Doorn, G. S., Leimar, O., & Weissing, F. J. (2007). Life-history trade-offs favour the evolution of animal personalities. *Nature, 447*(7144), 581-584. doi:10.1038/nature05835

Wright, S. (1922). Coefficients of inbreeding and relationship. *American Naturalist, 56*, 330-338. doi:10.1086/279872

Xie, Z. Q., Wang, L., Wang, L. R., Wang, Z. Q., Lu, Z. H., Tian, D. C., . . . Hurst, L. D. (2016). Mutation rate analysis via parent-progeny sequencing of the perennial peach. I. A low rate in woody perennials and a higher mutagenicity in hybrids. *Proceedings of the Royal Society B-Biological Sciences, 283*(1841). doi:10.1098/rspb.2016.1016

Chapter 4: The relationship between environmental factors and inbreeding depression in populations of animals and plants: a meta-analysis

4.1 Abstract

Inbreeding depression, the reduction in fitness of inbred offspring, is a major concern in the conservation of small and isolated populations. However, the severity of inbreeding depression varies with environmental conditions, but what extent do environmental changes interact with inbreeding depression across species and populations is poorly understood. In this study, I used meta-analysis to synthesize the evidence and investigate how inbreeding depression varies with environmental conditions and stress in animals and plants. Investigation of 1362 inbreeding depression effect sizes indicated that environmental stress levels, but not environmental types were a significant predictor of inbreeding responses. Specifically, analyses of inbreeding effect sizes in stressful environments across populations showed a unimodal relationship between inbreeding depression and stress intensity in fitness component traits. Additionally, when examining the changes in inbreeding depression after environmental treatment (environmental response effect sizes), the magnitude of inbreeding depression for fitness component traits did not change when populations experienced environmental stress whereas a negative relationship was observed in non-fitness components. Finally, when two or more environmental conditions changed simultaneously, the impacts of inbreeding depression were more severe. Overall, my finding suggested inbreeding depression occurred regardless of environmental stress. However, the magnitude of inbreeding depression in fitness component traits did not change after populations experienced environmental stress. Therefore, environmental stress or inbreeding depression independently influence demography changes in natural populations via fitness reduction.

4.2 Introduction

The expansion of anthropogenic activities and human population have increased the exploitation of natural resources and degraded natural habitats for animals and plants, consequently leading to declines in population size in natural populations (Vitousek,

Mooney, Lubchenco, & Melillo, 1997). When population sizes become small, they are more vulnerable to genetic problems such as genetic drift and inbreeding (Frankham, 1995; Hedrick, 2001; Keller & Waller, 2002). These processes lead to the erosion of genetic diversity and expression of deleterious alleles, consequently limiting evolutionary potential and reducing fitness in natural populations (Charlesworth, 2003; Frankham *et al*., 1999). These genetic problems may be complicated by environmental stress arising from anthropogenic climate changes and other anthropogenic environmental changes, which tend to reduce population sustainability, fragment geographical ranges and isolated species (Bijlsma, Bundgaard, & Boerema, 2000; Frankham, 1995).

Inbreeding depression is defined as the reduction in fitness of offspring as a result of mating among related individuals (Crnokrak & Roff, 1999; Hedrick & Garcia-Dorado, 2016; Keller & Waller, 2002). Fitness costs through inbreeding depression are primarily caused by the unmasking of recessives alleles that confer detrimental phenotypes in homozygous individuals. This is known as the dominance model of inbreeding depression (Charlesworth. & Charlesworth., 1987, 1999; Charlesworth. & Willis, 2009; Kristensen, Pedersen, Vermeulen, & Loeschcke, 2010). Alternatively, however, loci are maintained by heterozygote advantage (the overdominance model), the decline of heterozygote frequency upon inbreeding will also reduce population fitness (Charlesworth & Charlesworth, 1987, 1999; Charlesworth & Willis, 2009; Kristensen *et al*., 2010).

The extent and expression of inbreeding depression can be highly dependent on environmental conditions under which it is being measured, due to the interaction between genotypes and environments, a phenomenon referred to environment-dependent inbreeding depression (Cheptou & Donohue, 2011; Hedrick & Kalinowski, 2000; Henry, Pradel, & Jarne, 2003; Kristensen, Sorensen, Pedersen, Kruhoffer, & Loeschcke, 2006). The expression of inbreeding by environment interaction, arising from the phenotypic difference between inbred and outbred offspring across environmental gradient, provides evidence of environment-dependent inbreeding depression. Inbreeding-environment interaction, a particular type of genotype-environment interaction, is thought to shape the expression of inbreeding depression and to alter patterns of natural selection operates in natural populations (Waller, Dole, & Bersch, 2008).

Consistent relationships between inbreeding depression and environmental factors are not universal. Numerous empirical and theoretical studies have investigated the effects of environmental factors on inbreeding depression. In 2010, Fox and Reed reported that the magnitude of inbreeding depression increased under stressful environment in the seed-feeding beetle *Callosobruchus maculatus*, with the levels of stress explaining 66% of the variation of inbreeding depression. Another study carried out by Kristensen and co-workers (2008), investigated the effects of different temperature regimes on inbreeding depression within different traits in *Drosophila melanogaster*. The results demonstrated that inbreeding depression in egg-to-adult viability only occurred within stressful (low or high) temperature regimes, suggesting that inbreeding depression was environment-dependent and was more severe under stressful environments (Kristensen *et al.*, 2008).

A conflicting result was observed by Henry and his colleagues (2003), who found that inbreeding depression in survival was lower in the field (a more stressful environment) in the freshwater snail, *Physa acuta* (Henry *et al*., 2003). In addition, a greenhouse experiment carried out by Sandner and Matthies (2016) showed that inbreeding depression was less severe when *Silene vulgaris* was subjected to stressful environments (nutrient deficiency and drought). Both an increase, and a decrease in inbreeding depression are plausible consequences of increasing environmental stress. The first two could be explained by inbreeding depression only being prominent in stressful condition since both inbred and outbred populations perform well in benign environments (Fig. 1a); the last two might indicate that both outbred and inbred populations suffer fitness reduction under stress, and therefore inbreeding depression would be only observed in benign environments (Fig. 1b) (Cheptou & Donohue, 2011).

Two key mechanisms are thought to underpin environment-dependent inbreeding depression: environment-dependent phenotypic expression and environment-dependent selection (Cheptou & Donohue, 2011; Kristensen *et al*., 2010). During environment-dependent phenotypic expression (Fig. 2a), the environment shapes the expression of phenotypes that are under constant natural selection (Cheptou & Donohue, 2011). The selection regime in this case is consistent across different environments, meaning that

traits have a uniform relationship to fitness in different environments. Inbred and outbred progeny respond to environmental factors differently via plastic responses in phenotype that alter fitness differences between them. In the second situation (Fig. 2b), the environment does not modify the expression of phenotypes but alters natural selection regimes such that fitness differences between inbred and outbred progeny are environmentally contingent (Cheptou & Donohue, 2011). In either case, environmental change alters either phenotypic response or adaptive values of phenotypes between inbred and outbred progeny leading to shifts in the magnitude of inbreeding depression. In summary, genetic constitution and ecological adaptation underpins environment-dependent inbreeding depression.

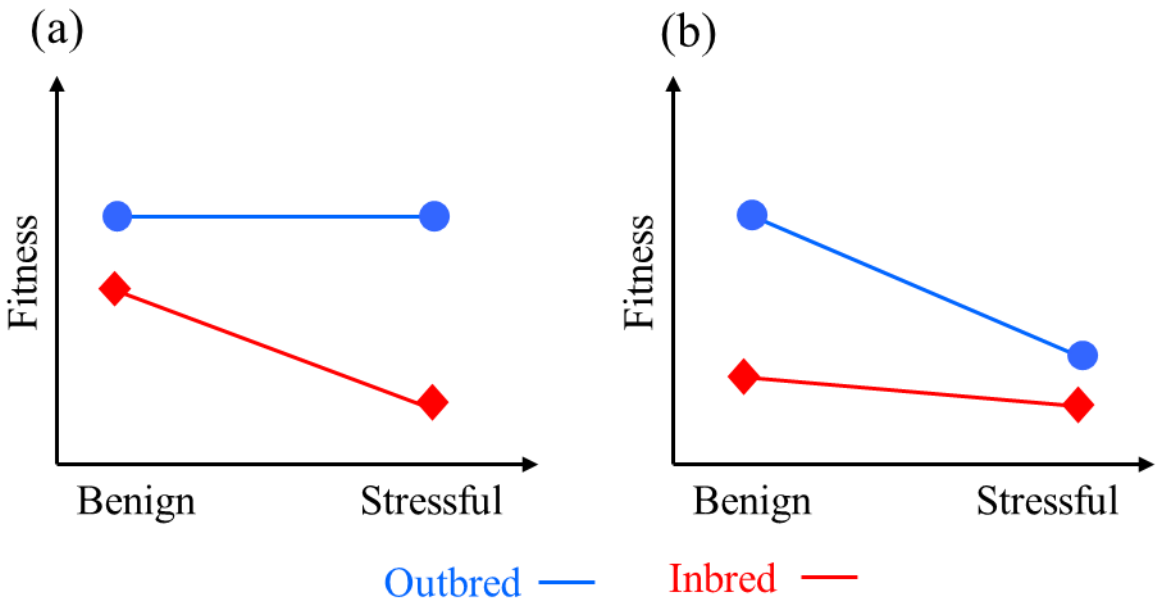


Fig. 1 Potential consequences of stressful environments for inbreeding depression, indicated by the fitness difference between outbred offspring (blue lines) and inbred offspring (red lines). (a) Inbred offspring are more vulnerable to environmental stress than outbred offspring, leading to the increase in inbreeding depression in stressful environments. (b) Outbred offspring are better than inbred offspring at exploiting the opportunities offered by benign environments whereas both groups show similar fitness in stressful environments.

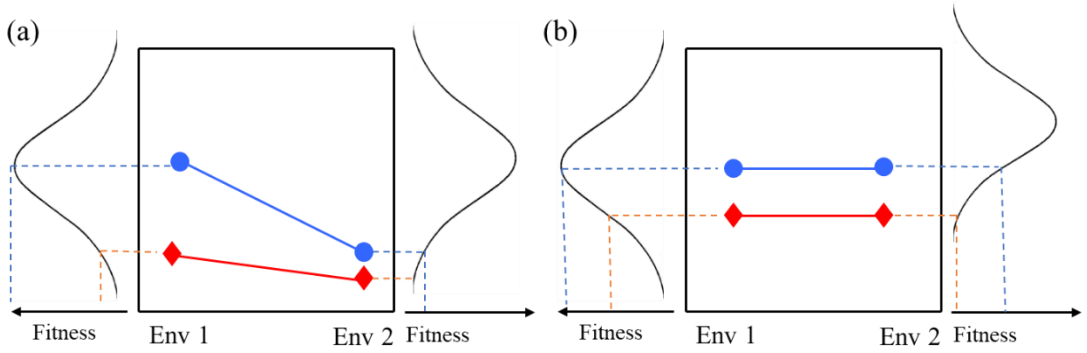


Fig. 2 The underlying mechanisms of environment-dependent inbreeding depression. (a) environment-dependent phenotypic expression: in this scenario, selection on phenotypes remains constant across environments but inbred and outbred phenotypes respond to environments differently, leading to the variation of inbreeding depression between environments. (b) environment-dependent selection: in this scenario, phenotypic expression remains unchanged while selection shifts across environments, resulting in the difference between fitness values of inbred and outbred offspring.

Rather than the contradictory results described above, some studies have attempted to synthesise the evidence, to understand if there is a general direction or pattern to inbreeding depression-environment interactions. Fox and Reed (2011) performed a meta-analysis to investigate the impacts of environmental stress on inbreeding by calculating the difference in the number of lethal equivalents between stressful and benign environments. Lethal equivalent indicates the rate of fitness declines with inbreeding (Hedrick & Garcia-Dorado, 2016). Their result showed that the number of lethal equivalents scaled positively with the intensity of the stress. Likewise, Armbruster and Reed (2005) compared the lethal equivalents under benign and stressful environments across 34 studies and found that the lethal equivalents under stressful environments are significantly greater than benign environments . Although both studies indicated that inbreeding depression is more severe under stressful environments, the analyses were quite different. Fox and Reed's studies compared inbreeding depression at the population level across different intensities of environmental stress, by calculating the changes in the number of lethal equivalents before and after environmental treatments. However, Armbruster and Reed's studies examined this variation by averaging lethal equivalents in benign and stressful environments separately across

individual studies. The variation of stress intensity was not considered in their study. Additionally, Fox and Reed limited the scope of their study to only one phenotypic trait, survival, and did not consider the effects of other fitness components, including reproduction and viability, on inbreeding outcomes. Furthermore, Armbruster and Reed's study suffered from low statistical power, where their study failed to control within studies variance by applying t-test to compare the difference of lethal equivalents between benign and stressful environments. This raised an issue of whether their result was biased due to small sample seize studies. Lastly, different types of stressful treatments were grouped together. None of these studies tried to investigate the impact of different types of environmental factors on inbreeding depression.

A growing number of inbreeding-related papers and limited scopes of the previous meta-analysis indicate that there is a need to expand an empirical synthesis on the evidence of inbreeding responses. An understanding of environmental effects on inbreeding depression will allow conservation practitioners to predict the alteration in fitness outcomes and population demography in response to natural selection exerted by environmental factors. Understanding the impacts of environmental changes on inbreeding depression in the conservation context requires investigation of different environmental factors separately. For example, stress from global warming or invasion of exotic species should not be regarded as the same when considering an appropriate conservation management.

In this study, Bayesian meta-analysis was carried out to understand how inbreeding depression varies with environmental stress. Changes in inbreeding depression were investigated in relation to both changes in environmental stress within populations and overall shifts in environmental stress between populations of different species. Following that, the impacts of different types of environmental factors on inbreeding responses were investigated. Based on the previous studies, I predicted inbreeding depression will scale positively with environmental stress. I interpret the results in the context of their implications for conservation practice.

4.3 Methods

4.3.1 Literature search

A comprehensive, subject-mapped literature database of inbreeding depression has been constructed by Neaves and co-workers (2015). In this study, I updated this resource to 4th May 2018 and utilized it to facilitate a comprehensive meta-analysis of the relationship between inbreeding depression and environmental stress. The literature search initially followed that used by Neaves *et al*. (2015), using identical search strings, using three online databases: Web of sciences core collection, Scopus and JSTOR. Literature searches were completed on 4th May 2018. The search records from the three online databases were exported to endnote (version X 9.2) libraries separately. A master library was created by combining these endnote databases following the removal of conference papers and book chapters. Identical articles were eliminated by applying endnote filters to identify duplicates. The final literature database was exported as an excel spreadsheet for examination of article relevance.

I also carried out a supplementary literature search to address gaps in the literature returned by the initial search. Specifically, I found that a number of articles on inbreeding depression that had been detected by Neaves *et al*. (2015), and which were known to be relevant to fitness consequences of inbreeding depression in animals and plants, had been missed by my more recent search using the same search terms. The literature search strings were therefore refined by identifying the missing articles and examining the key words in the abstracts of missing but potentially relevant articles. The updated search strings were applied to enhance the scope of the search without losing specificity. The final search strings are listed in Table 1.

Table 1. The final search strings used by in this meta-analysis to retrieve articles from online database.

Group	Search string
Inbreeding related strings	"In*breeding coefficient$"
	"Cost$ NEAR/5 in*breeding"
	(inbred SAME mating*) NOT (("Quantitative trait loc*")OR(QTL*))
	(inbred SAME (offspring OR progeny))
	Selfed SAME out*

	"Optimal outcrossing" OR "Outcrossing distance"
	"Benefit* NEAR/5 dispersal"
	"Cost* NEAR/5 dispersal"
	("Natal dispersal" AND (inbred OR in*breeding OR heterosis OR self* OR fitness))
	(Philopatr* AND (inbred OR in*breeding OR heterosis OR self* OR fitness))
Fitness related strings	(Depression SAME in*bre*)
	(Depression SAME fitness)
	(Heterosis AND in*breeding)
	"Genetic load"
Effects of inbreeding related strings	Effect$ NEAR/5 in*breeding
	consequence$ NEAR/5 Inbreeding
	influence$ NEAR/5 inbreeding
	Outcome$ NEAR/5 inbreeding

Boolean syntax above is based on Web of sciences template, and changes were made to adapt the search to the other two literature databases used here.

4.3.2 Inclusion and exclusion criteria

The inclusion and exclusion criteria were described in the chapter 2. However, there were additional criteria I applied in this meta-analysis, which are listed below.

- **Type of study**

  Study was defined as a separate inbreeding outcome calculated from individual compared groups within an article. In this meta-analysis, I included articles reported inbreeding effects under different biotic or abiotic conditions.

- **Types of subjects**

  Experimental populations derived from a natural population and maintained at a large population size with minimal artificial selection were also included.

- **Types of phenotypic outcome**

  I considered relevant phenotypic data from the offspring of inbred and outbred crosses under different environmental treatments. Studies were eliminated if the probable direction of natural selection acting on a trait could not be established. An example of a trait in this category could be the concentration of chemical compounds, where the relationship of fitness to the chemical's concentration is unknown.

- **Sources of heterogeneity**

  The sources of heterogeneity were described in the chapter 3

4.3.3 Article screening strategy

The relevance of all articles in the literature database was assessed in a hierarchical way. First, I screened the article abstracts with the results of Automated Content Analysis (ACA) (Nunez-Mir, Iannone, Pijanowski, Kong, & Fei, 2016). ACA is a text-mining technology that allows researchers to identify topics underlying a literature corpse. Specifically, articles in the literature database were sorted by topic proportion and irrelevant articles were quickly eliminated by scanning the titles for those with clearly irrelevant topics, which their relevance to my meta-analysis were quantified in the Chapter 2. Next, the articles were assessed manually following by reading abstracts and finally following by full-text assessment. Articles were assessed as relevant when they meet the inclusion criteria described above (types of study, types of subject, types of phenotypic outcome and types of comparators).

4.3.4 Data extraction

Data for calculation of effect sizes and for the sources of heterogeneity was extracted from relevant articles identified after the full-text assessment stage. I retrieved all data that were required to calculate inbreeding effect sizes and their measurement error variance (*mev*) from these articles. I approached the authors to request raw or summarised data if the required data was not presented in an article in the required form. Other information that could potentially cause variance in inbreeding effect sizes was also extracted (sources of heterogeneity). Specific details are described below:

- Mean phenotypic values (μ) and their standard deviation (s): For continuously distributed traits, the mean phenotype for each trait measured was extracted for each group (inbred and non-inbred groups in treatment and control environments). Associated standard deviation values were retrieved directly or calculated from the standard error and sample size where necessary. Sample sizes for phenotypic mean values were also recorded for 4 different groups (inbred and outbred groups in treatment and control environments).

Binomially distributed phenotypic values (e.g. hatched offspring, survival and death) were extracted as a proportion or percentage. The accompanying standard deviation and its corresponding sample size were also retrieved if they were presented in the papers. The following equation was used to calculate standard deviation for binomial data if it was not presented in the article.

$$S(p) = \sqrt{Var(p)} = p(1-p)/n$$

Where p is the probability of success in n observations (Koricheva, Gurevitch, & Mengersen, 2013). WebPlotDigitizer version 4.4 (Rohatgi, 2021) was used to estimate trait means and standard deviations if the phenotypic values were only presented in figures. Data for different traits or where inbred progeny had a differing inbreeding coefficient were retrieved separately.

- Environmental factors: The environmental manipulations applied to treatment groups were recorded as the table below. I categorized any interactions between species on different trophic levels as *between tropic levels interaction*, such as predator, disease and between species competition. Otherwise, when species interactions occurred *within the same trophic level*, which include parental care and competition, I classified it as a *within trophic level* interaction. When more than one environmental treatment was applied at the same time, or the environmental treatment was carried out by placing organisms at different locations (e.g. field, common garden, greenhouse or lab), between which multiple environmental factors vary, I categorized these as multiple factors.

Table 2. Categories of environmental factors used in this meta-analysis.

Category 1	Category 2	Example
Light	Abiotic factors	Sun light availability
Nutrient		Food availability
Water stress		Drought
Temperature		Increased temperature
Chemical stress		Methanol treatment
Salinity		Salt concentration
Ventilation		Ventilation or not
Between trophic levels interaction	Biotic factors	Disease
Within trophic levels		Maternal care

interaction		
Multiple factors	Multiple factors	Lab and field

Category 1 classified every single environmental factor population could experience whereas category 2 tried to distinguish the impacts of abiotic environmental shifts from the impacts of changes in community structure.

4.3.5 Calculation of effect sizes and measurement error variance

The inbreeding effect sizes were calculated for each study and environmental treatment as a log response ratio:

$$ES = -\ln(\mu_i/\mu_o) \quad (1)$$

where μ_i and μ_o are trait mean values for the inbred group and non-inbred group (outbred) respectively (Borenstein, Hedges, Higgins, & Rothstein, 2009). The study measurement error variance (*mev*) was calculated as

$$mev = S^2\left(\frac{1}{n_I\mu_I^2}\right) + S^2\left(\frac{1}{n_O\mu_O^2}\right) \quad (2)$$

where n_I and n_O are the sample sizes of inbred offspring and non-inbred offspring and S_I and S_O are standard deviations of inbred group and non-inbred group respectively (Borenstein *et al*., 2009). These effect sizes quantify relative difference in performance between inbred and non-inbred groups at different environmental treatments. Positive values indicate inbreeding depression whereas negative values mean outbreeding depression. In order to investigate the effects of environmental shift within a population, I also estimated environmental response inbreeding effect sizes (environmental response effect sizes) from the equation:

$$ES_{Environmental\ response} = \frac{ES_{Treatment}}{Inbreeding\ coefficient} - \frac{ES_{Control}}{Inbreeding\ coefficient} \quad (3)$$

, where $ES_{Treatment}$ was the effect sizes calculated from the equation (1) in treatment environments and $ES_{Control}$was the effect sizes calculated from the control environments. The effect sizes were divided by inbreeding coefficient in order to standardise by the degree of inbreeding. The terms $\frac{ES}{Inbreeding\ coefficient}$ can be interpreted as genetic load in treatment and control environments respectively (Charlesworth and Charlesworth, 1987). This environmental effect size represents the change of

inbreeding responses due to changing environmental treatment. The relationship between inbreeding effect sizes and environmental response effect size is displayed in Figure 3.

Stress levels were estimated from the trait means of outbred groups and were calculated from the equation described below:

$$\text{Stress level} = \frac{Outbred_{Benign} - Outbred_{Stress}}{Outbred_{Benign}}$$

, where $Outbred_{Benign}$ is the phenotypic mean value of the outbred group in the benign environment and $Outbred_{Stress}$ is the phenotypic mean value of the outbred group in a stressful environment (Fox & Reed, 2011). These stress levels represent the relative change in phenotypic (or fitness) due to environmental changes. The definition of benign and stressful environments is based on the organism's perception of the environmental conditions. Specifically, the benign environment is the environment where the organisms have highest phenotypic values. The stress levels for organism perception are always positive and all the values lie between 0 and 1, by definition.

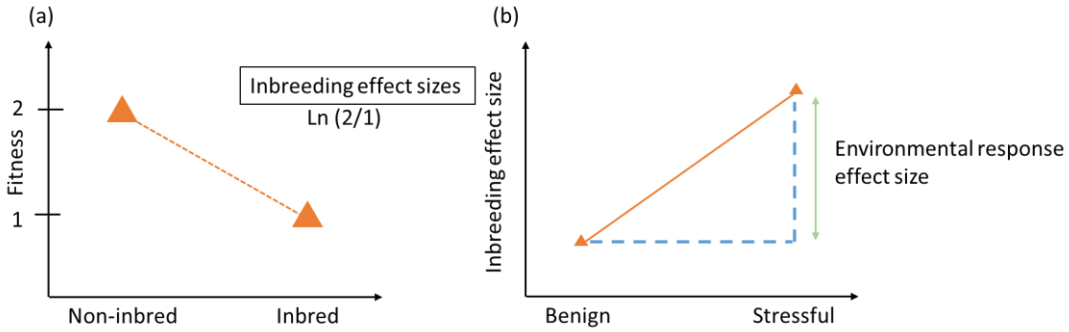


Fig. 3 (a) Inbreeding effect sizes quantify the magnitude of inbreeding depression by calculating the log response ratio of fitness in non-inbred offspring to fitness in inbred offspring. In this example the inbreeding effect size is ln (2). (b) The diagram shows the relationship between inbreeding effect sizes and environmental response effect size. Environmental response effect sizes quantify the shift in inbreeding caused by environmental stress.

4.3.6 Meta-Analysis

Effect sizes were meta-analysed using the MCMCglmm package (version 2.29) in R (Hadfield, 2010; R Core Team, 2019). This package provides functions for modelling

effect size data within Bayesian generalised mixed modelling frameworks, via Markov chain Monte Carlo sampling, which allows for tests of the effects of relevant predictor variables (Hadfield, 2010). MCMCglmm permits models with fixed- and random effects while taking account of the study *mev* (an essential part of meta-analysis). Here, I fitted explanatory variables of primary interest as fixed effects, such as inbreeding coefficient, environmental factors and stress levels. Otherwise, study identity (study ID) was fitted as random effect in order to capture and account for data structure, and to model variation in effect sizes in different contexts.

The MCMC chains for all models were run with a total of 1.1×10^5 iterations with a burn-in of 1×10^4 iterations and then a following thin interval of 100 iterations where parameters were extracted. Total sample size for the posterior distribution resulting from this sampling scheme was 1000 for each fitted model. Priors for random effects were uniform improper distributions on their standard deviation (Gelman, 2006). Posterior distributions were summarised to derive point estimates of the model parameters (posterior means, including pooled effect sizes) and their 95% credible intervals. The meta-effect sizes were considered statistically significant when their 95% credible interval did not overlap with zero. Bayesian p-values, representing the probability of the parameter's location, were estimated by examining the proportion of the posterior distribution overlapping with zero (or with the point location for another parameter). Results were plotted as forest plots, including effect sizes and their corresponding credible intervals. A Bayesian measure of fit was assessed for each model via a deviance information criteria (DIC) (Spiegelhalter, Best, Carlin, & van der Linde, 2002). DIC estimates model adequacy by measuring model fit in relation to complexity (the effective number of parameters) (Spiegelhalter et al., 2002). Due to the variation in DIC arising from separate model runs, each model was run three times and the average DIC values were taken and compared.

To model inbreeding responses with environmental factors, inbreeding effect size was the response variable and a fixed-effect, inbreeding coefficient was fitted into a minimal model, which is

Inbreeding Effect size ~ inbreeding coefficient

Article ID and *mev* were fitted as random effects for all models progressed from the minimal model. Article ID is a unique number assigned to each article within the database. I considered inbreeding coefficient is the predictor variable of the inbreeding effect size since inbreeding coefficient quantifies the magnitude of inbreeding by the probability of identical by descent (Frankham, Ballou, & Briscoe, 2010). This model analysed the variation in inbreeding depression when stress intensity increase. It reflects absolute amount of inbreeding depression we can expected to observed across stress gradient, regardless of species or population. Since this analysis investigated the changes of inbreeding depression across studies, I referred it to "*among study relationship*" in the result section.

Alternatively, environmental response effect sizes were modelled with explanatory variables directly since inbreeding coefficient has been standardised in the effect sizes.

Environmental response effect size ~ 1

Article ID and *mev* were included as random effects for all models proceeded from the minimal model. The reason to exclude inbreeding coefficient as a predictor variable is that it cannot predict the change in inbreeding depression with a population in response to environmental shifts. This model analysed the expected amount of change in inbreeding depression as environments become more stressful, and related the amount of change in the inbreeding depression to the amount of change in the environment (stress level). Since this analysis investigated the change in inbreeding depression within study, I referred it to "*within study relationship*" in the result section. The intercept from these models was expected to be not significant different from zero because there is no difference in inbreeding depression between stressful and benign environments when the intensity of stress is zero.

4.4 Result

4.4.1 Review descriptive statistics

My first stage literature search in combination with Neaves' literature database retrieved 29836 articles from 1941 to 2018. There were 129 research articles that met my criteria for meta-analysis, which included 7 categories of environmental treatments. The types of the environmental treatments are well distributed across the studies.

Between species interaction, which comprised of infection, multiple species competition and present of predators, are the most prevalent category (21%) whereas light and salinity only account for 2 % of total studies. For all of these environmental manipulations, 2 (34%) and 7 (52%) categories are grouped into biotic and abiotic treatments respectively. Among these 129 articles, there are totally 1362 effects sizes, with 682 effect sizes for fitness component traits and 680 effect sizes for non-fitness component traits.

Table 3. Model fitting summarises for meta-analysis containing fix-effect explanatory variables and inbreeding effect sizes.

Model	Fixed effects	DIC	Among-article variance	Within-article variance	Within-study variance	Trait types
A1	~Intercept	532.3	0.0776	0.0945	0.008	FC
A2	~*F*	504.53	0.0794	0.0904	0.008	FC
A3	~*F* + Stress level	491.8	0.0814	0.0882	0.008	FC
A4	~*F* + poly (Stress level, 2)	481.81	0.0799	0.0865	0.008	FC
A5	~*F* + poly (Stress level, 3)	480.57	0.0795	0.0862	0.008	FC
B1	~Intercept	809.38	0.0092	0.1632	0.0104	NFC
B2	~*F*	811.65	0.0099	0.1626	0.0104	NFC
B3	~*F* + Stress level	803.84	0.0074	0.1616	0.0104	NFC
B4	~*F* + poly (Stress level, 2)	804.2	0.0081	0.1609	0.0104	NFC
E1	~Intercept	1418.38	0.0454	0.1325	0.0093	All
E2	~*F*	1405.29	0.0477	0.131	0.0093	All
E3	~*F* + Env_factors (1)	1420.11	0.0383	0.132	0.0093	All
E4	~*F* + Env_factors (1) + Trait	1422.02	0.0351	0.133	0.0093	All
E5	~*F* + Env_factors (1) * Trait	1430.9	0.0332	0.1335	0.0093	All
H1	~Intercept	1418.38	0.0454	0.1325	0.0093	All
H2	~*F*	1405.29	0.0477	0.131	0.0093	All
H3	~*F* + Env_factors (2)	1408.83	0.0421	0.1314	0.0093	All
H4	~*F* + Env_factors (2) + Trait	1410.38	0.0386	0.1318	0.0093	All
H5	~*F* + Env_factors (2) * Trait	1411.21	0.0377	0.1316	0.0093	All

F: inbreeding coefficient; FC: fitness component traits; NFC: non-fitness component traits; All: all types of traits. Env_factors (1): category 1 of environmental factors listed in Table 2; Env_factors (2): category 2 of environmental factors listed in Table 2; Trait: categorial description of the traits, whether they are fitness component trait or not.

4.4.2 Among study relationship between inbreeding depression and stress levels

The relationship between inbreeding depression and environmental stress was investigated across studies for fitness component and non-fitness component traits in stressful environments respectively. I detected a significant second order polynomial relationship between stress levels and inbreeding depression (pMCMC < 0.001; Fig. 3). Adding second order polynomial term increased the goodness of fit of the model (change in DIC = -9.99, Table. 3). The intensity of inbreeding depression rises steadily until stress levels reach around 0.49. Beyond this point, the magnitude of inbreeding depression declines while stress levels increase further (Fig. 4). Inbreeding depression is therefore most acute at intermediate levels of environmental stress (in relation to a benign comparator environment). The inbreeding effect size reaches a value of 0.163 when stress level reaches 1. The degree of inbreeding depression has a similar level when stress intensity is 0 and 1. The credible intervals of this model do not overlap with zero at any point along the environmental stress gradient, indicating inbreeding depression is occurring regardless of the amount of environmental stress the populations face.

For non-fitness component traits, the stress intensity displayed significantly linear relationship with inbreeding depression (Slope = 0.195, pMCMC = 0.008***; Fig. 3). As expected, the intercept of the model was significantly different from zero, implying that inbreeding depression occurs even when there is no environmental stress (Intercept = 0.085, pMCMC = 0.006***).

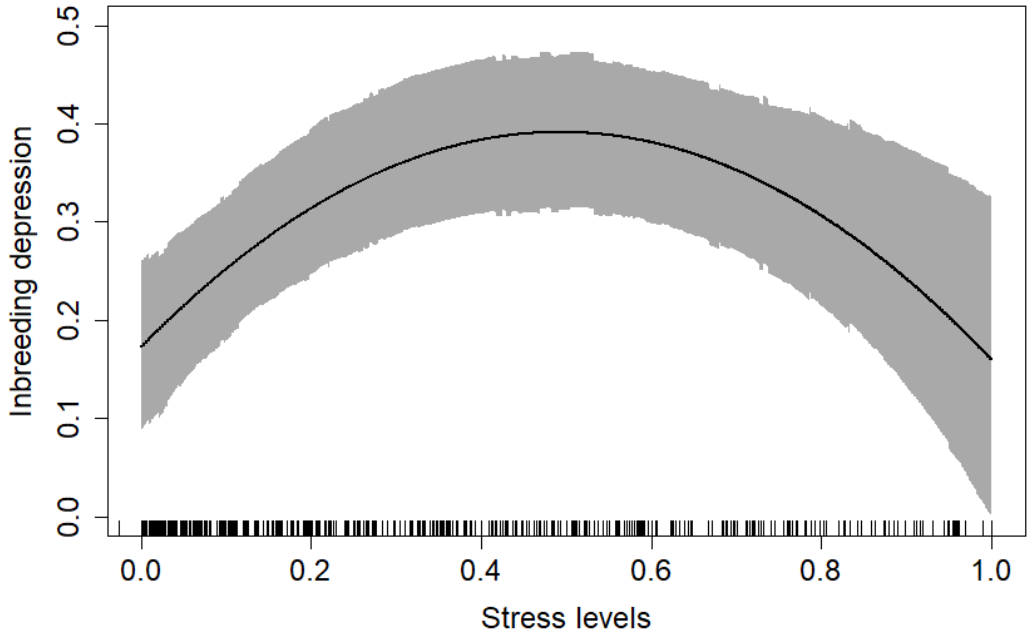


Fig. 4 The second-order polynomial relationship between inbreeding effect sizes of fitness component traits and stress level in stressful environments. Stress level was calculated as the fitness difference between benign and stressful environments divided by the fitness value observed in the benign environment, for outbred individuals. The data represent 682 observations of inbreeding depression in stressful environments. This regression reflects the average association between stress and inbreeding depression across different species and populations. The shaded area represents the 95% credible zone for the regression line, indicating a posterior distribution or a predictive distribution that an unobserved data might show up.

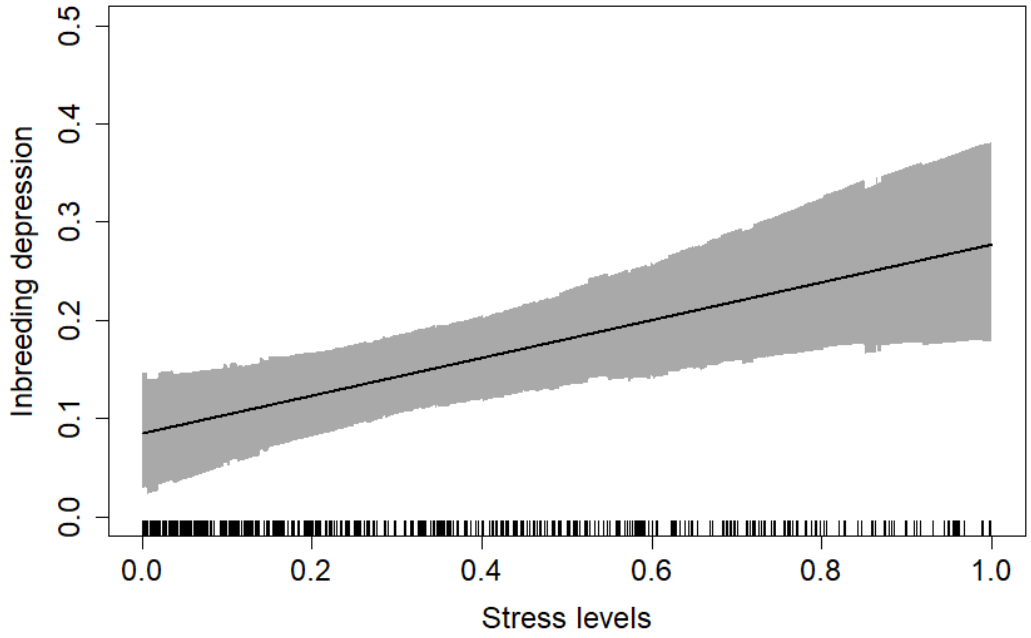


Fig. 5 The positive relationship between inbreeding effect sizes of non-fitness traits and stress levels in stressful environments. Stress level was calculated as the phenotypic values difference between benign and stressful environments divided by the phenotypic value observed in the benign environment, for outbred individuals. The data represent 680 datapoints in stressful environments. This regression predicts an increase of one level of stress corresponds to a rise of 0.195 inbreeding depression. The relationship can be demonstrated as follow: Inbreeding effect size = 0.085 + 0.195 stress. The shaded area represents the 95% credible zone for the regression line, as set out in Figure 4.

Table 4. Model fitting summarises for meta-analysis containing fix-effect explanatory variables and environmental response effect sizes.

Model	Fixed effects	DIC	Among-article variance	Among-study variance	Within-study variance	Trait types
A1	~Intercept	2530.18	1.296	2.108	0.008	FC
A2	~Stress level	2532.14	1.29	2.112	0.008	FC
B1	~Intercept	3424.06	5.374	8.167	0.104	NFC
B2	~Stress level	3423.07	4.802	8.155	0.104	NFC
E1	~Intercept	6010.29	0.879	4.527	0.0093	All
E2	~Env_factors (1)	6027.89	0.669	4.591	0.0093	All
E3	~Env_factors (1) + Fitness_class	6030.40	0.675	4.592	0.0093	All

H1	~Intercept	6009.7	0.884	4.523	0.0093	All
H2	~Env_factors (2)	6013.43	0.887	4.542	0.0093	All
H3	~Env_factors (2) + Fitness_class	6015.78	0.895	4.539	0.0093	All

F: inbreeding coefficient; FC: fitness component traits; NFC: non-fitness component traits; All: all types of traits. Env_factors (1): category 1 of environmental factors listed in Table 2; Env_factors (2): category 2 of environmental factors listed in Table 2. Fitness_class: category description of the trait as to whether it belongs to FC or NFC.

4.4.3 Within study relationship between inbreeding depression and stress levels

The response of inbreeding depression to environmental stress within a study was investigated for fitness component and non-fitness component traits separately. Stress level was not a significant predictor for environmental response of inbreeding depression. Considering fitness component traits, neither the slope nor intercept of environmental response effect sizes differed significantly from zero (Intercept = 0.043, pMCMC = 0.806; Slope = -0.2, pMCMC =0.422; Fig. 6. The credible intervals overlap with zero over all levels of the environmental stress predictor, indicating that regardless of the intensity of the stress, there is no expected net change in inbreeding depression (Fig. 6). However, for non-fitness component traits, I detected a negative relationship between the environmental response effect size and stress level (Fig. 6; slope = -1.52, pMCMC = 0.006). The environmental response effect sizes become negative beyond a stress level of ~ 0.58, meaning that inbreeding depression for non-fitness component traits in benign environments is larger than that in stressful environments (Fig. 7).

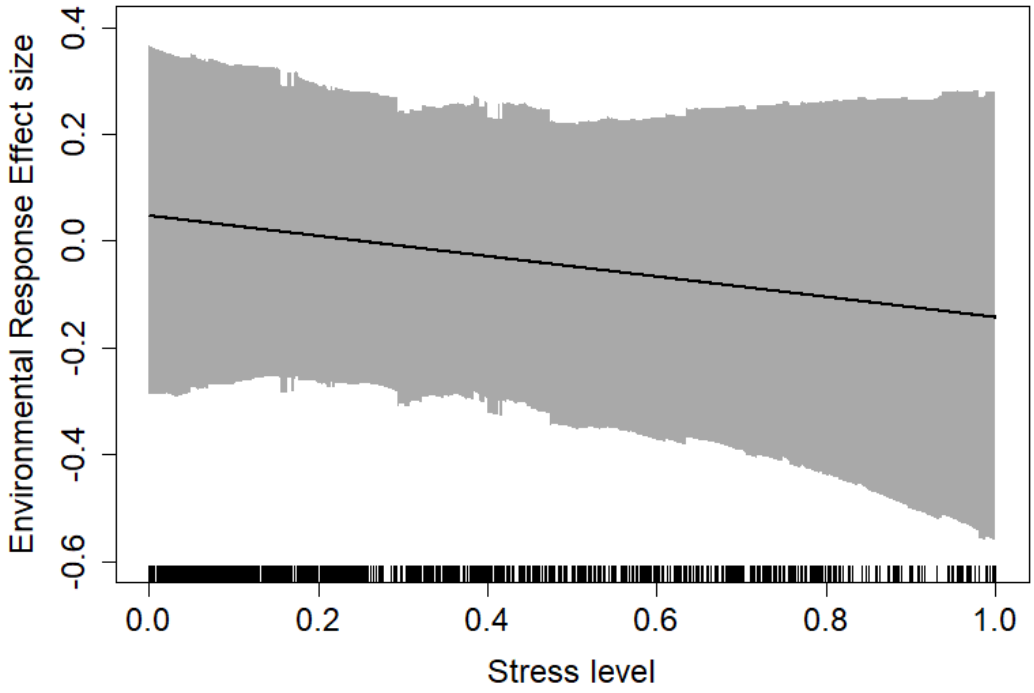


Fig. 6 No relationship between stress-induced change in inbreeding depression and the magnitude of stress induced by environmental change. The figure shows environmental response effect sizes for fitness traits against the environmental stress level caused by environmental treatment. The negative value of y axis implies that inbreeding depression is larger in a benign environment than in a stressful environment and vice versa. Stress level was calculated as the fitness difference between benign and stressful environments divided by the fitness value observed in the benign environment, for outbred individuals. The data represent 682 datapoints comparing the effects of inbreeding depression between stressful and benign environments. This regression reflects expected effect of stress on inbreeding depression within populations. The shaded area represents the 95% credible zone for the regression line, as set out in Figure 4.

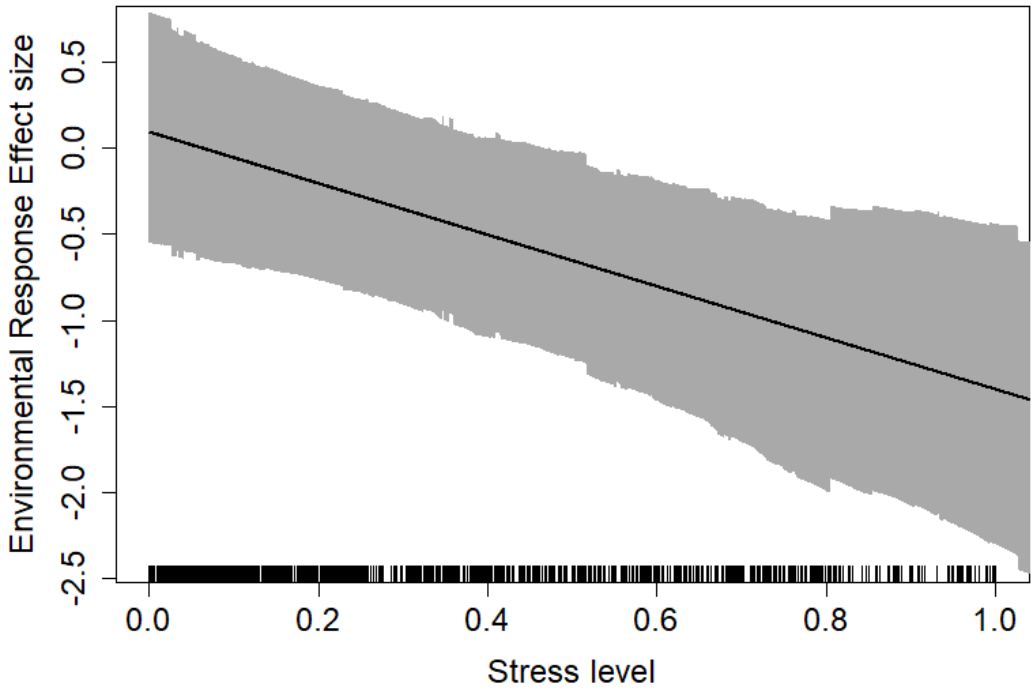


Fig. 7 The negative relationship between inbreeding effect sizes of non-fitness traits after populations experienced environmental treatments. Stress level was calculated as the phenotypic values difference between benign and stressful environments divided by the phenotypic value observed in the benign environment, for outbred individuals. The data represent 680 datapoints comparing the effects of inbreeding depression between stressful and benign environments. This regression predicts a decrease of one level of stress corresponds to a decline of 1.52 inbreeding depression changes between treatment and control environments. Specifically, the change in inbreeding depression when populations are exposed to environmental stress can be shown as: environmental response effect sizes = 0.88 -1.52 Stress. The shaded area represents the 95% credible zone for the regression line, as set out in Figure 4.

4.4.4 The effect of environmental factors on inbreeding depression across studies

The effects of the type of environmental stress factor on inbreeding depression was analysed with 1362 datapoints across species and populations. Typically, most types of environmental treatments were associated with inbreeding depression except salinity (posterior mean effect size = 0.067; credible interval, -0.267–0.399, pMCMC = 0.664; Fig. 8). Adding environmental factor as a predictor variable didn't improve

model goodness of fit (change in DIC = +15.1; Table 4). The variation among different environmental factors appears to be driven by the effects of light stress (Fig. 8), which displays the most severe inbreeding depression (posterior mean effect size = 0.51, credible interval 0.32 – 0.712, pMCMC < 0.001; Fig. 8). Orthogonal contrasts between types of environmental treatments indicate that inbreeding depression under light stress is significantly different from that observed in other types of stressful environmental treatments. Inbreeding depression for water stress and chemical stress are significantly lower than that for light stress (pMCMC < 0.001 and pMCMC = 0.002, respectively; Fig. 8).

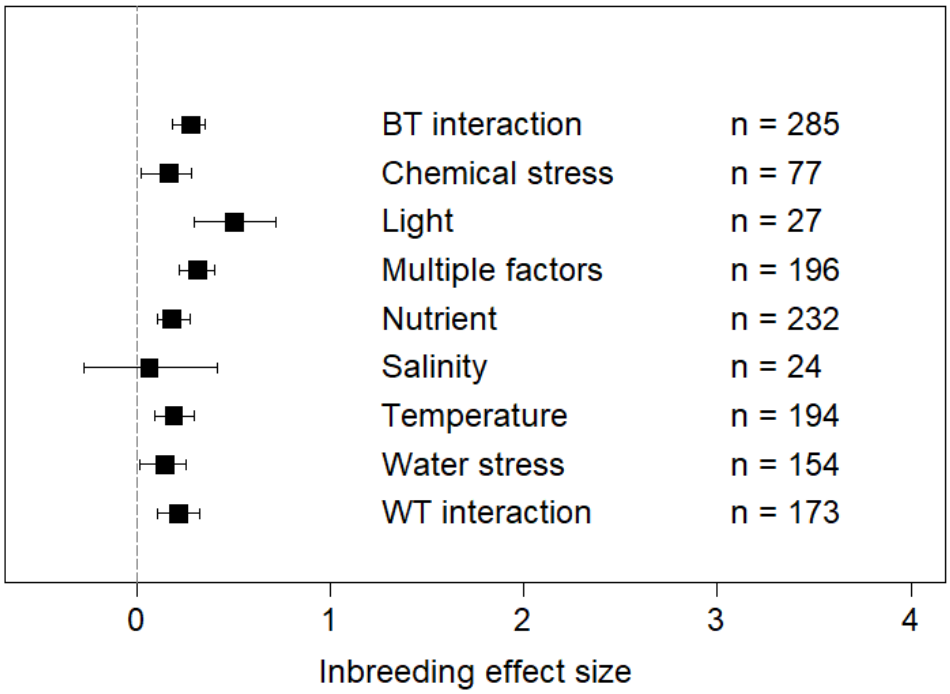


Fig. 8 Pooled effects summarising the variation of inbreeding depression in response to environmental treatments in stressful environments (all traits). BT interaction refers to between trophic levels interaction, including disease, the presence of predator and between species competition. WT interaction refers to within trophic levels interaction. Parental care and within species interaction are within this category. Positive values indicate the expression of inbreeding depression. Error bars show 95 % of credible intervals. The effect size was considered significant if credible intervals do not overlap with zero. n represents the number of effect sizes.

The nine types of environmental factors mentioned above were further categorised as either abiotic, biotic or "multiple factors" and their effects on inbreeding depression were analysed in the same way. Three of these environmental treatments are significantly associated with inbreeding depression. Multiple factors show most severe inbreeding consequence in the stressful environments (posterior mean effect size = 0.294, credible interval = 0.198 – 0.372, pMCMC < 0.001) whereas abiotic shows relatively least effect of inbreeding depression (posterior mean effect size = 0.187, credible interval = 0.123 – 0.251, pMCMC < 0.001; Fig. 9). Adding the environmental factor categories to the model did not improve its goodness of fit (change in DIC = +3.46, Table. 3). However, inbreeding depression slightly varied significantly among these environmental factors. Orthogonal contrasts among these environments indicates that abiotic and multiple factors differ significantly in the magnitude of inbreeding depression (pMCMC = 0.03).

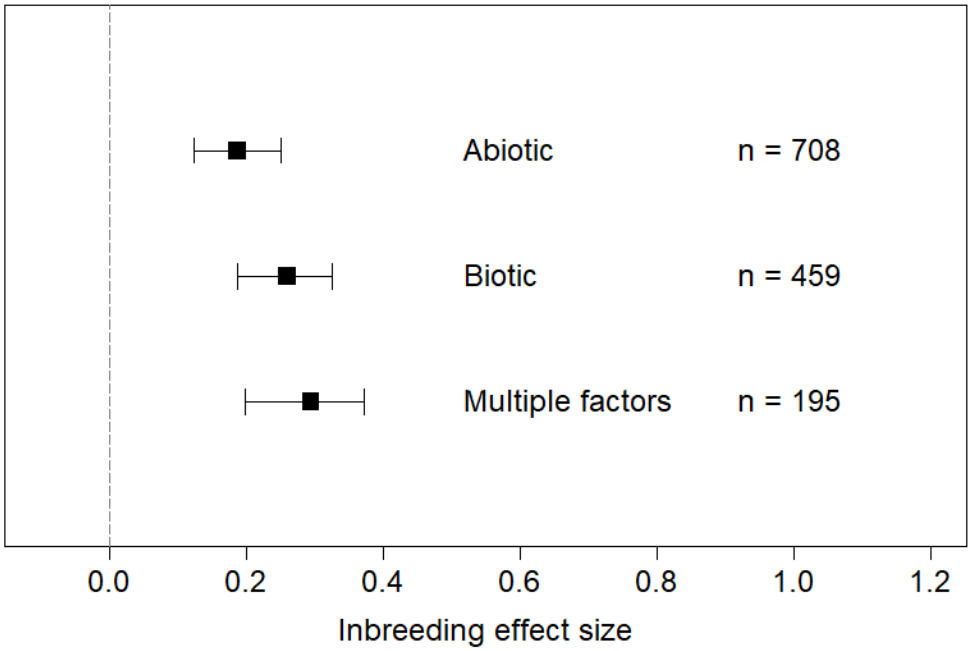


Fig. 9 Pooled effects summarising the variation of inbreeding depression in response to environmental treatments in stressful environments (all traits). Positive values indicate the expression of inbreeding depression. Error bars show 95 % of credible intervals. The effect size was considered significant if credible intervals do not overlap with zero. n represents the number of effect sizes.

4.4.5 The effects of environmental changes on inbreeding depression within populations

The response of inbreeding depression to different types of environmental treatments was analysed with 1362 datapoints within population. For most types of environmental treatments, stressful environments did not change the severity of inbreeding depression (Fig. 10). However, there are two exceptions, light and salinity stress treatments, which show contrasting trends. Populations experiencing light stress suffer more inbreeding depression than the same populations in benign environments (posterior mean effect size = 1.11, credible interval = 0.176 – 2.074, pMCMC = 0.028; Fig. 10). In contrast, the intensity of inbreeding depression is stronger in benign environments than stressful environments for populations subjected to salinity stress (posterior mean effect size = -3.472, credible interval = -5.089 - -2.058, pMCMC < 0.001; Fig. 10). As expected, orthogonal contrasts among these environmental factors reveal most of the variations are stemming from light and salinity.

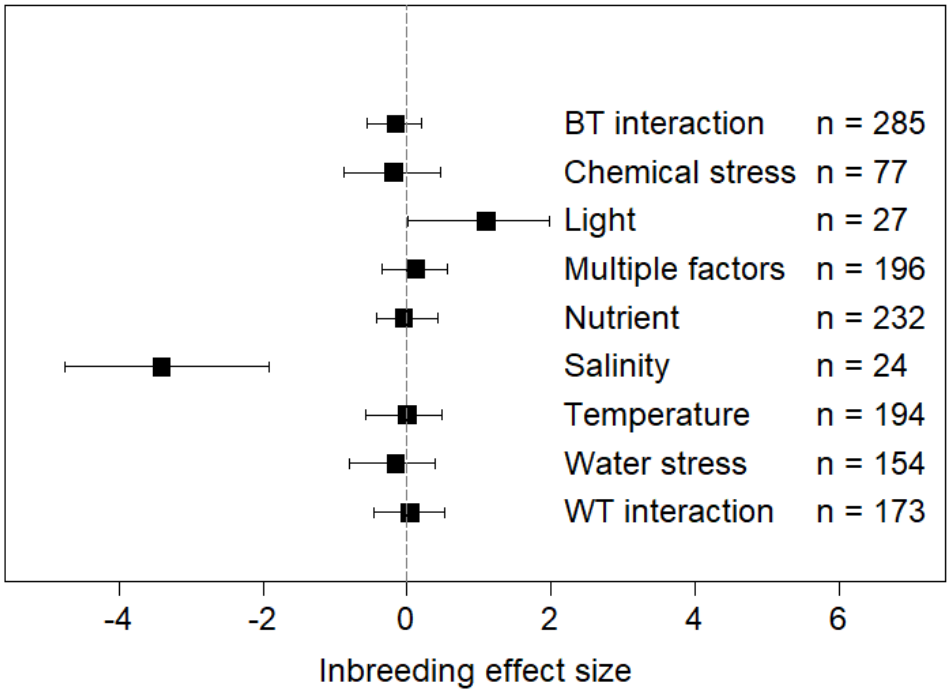


Fig. 10 Pooled effects summarising the variation of the changes in inbreeding depression in response to environmental treatments within population. The interpretations of pooled effect estimates, credible intervals, environmental types and n follow Fig. 7

Likewise, the nine types of environmental factors were further categorised as abiotic, biotic and multiple factors and the response of inbreeding depression to these environmental treatments were analysed over 1362 effect sizes within population. There was no consistent change in the magnitude of inbreeding depression with transition between benign and stressful environments, for any environmental treatment category (Fig. 11).

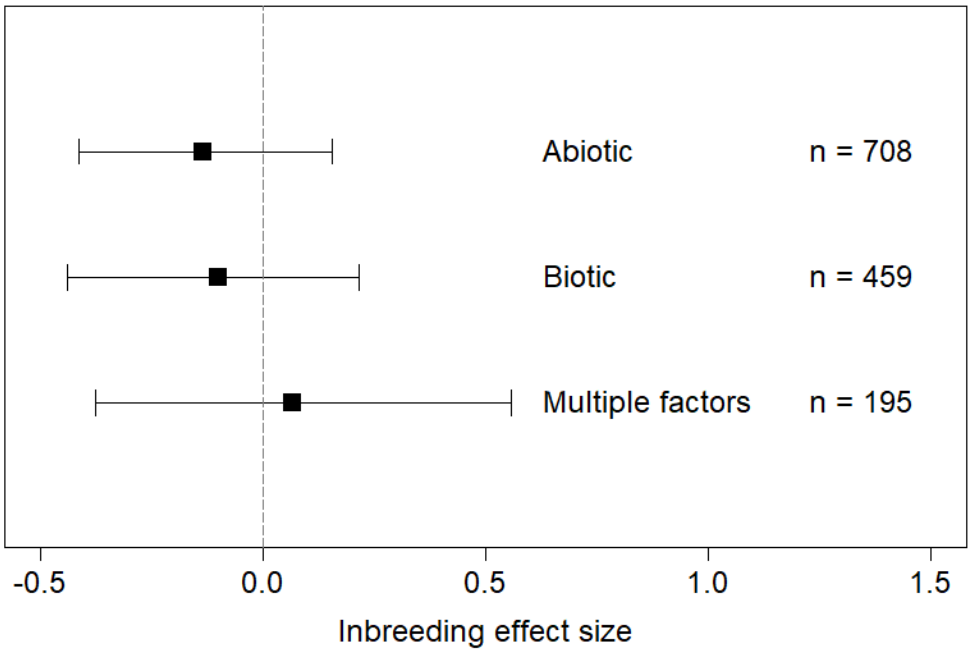


Fig. 11 Pooled effects summarising the variation of the changes in inbreeding depression in response to environmental treatments within population. Positive values imply the expression of inbreeding depression is greater in stressful environments than benign environments. Error bars show 95 % of credible intervals. The effect size was considered significant if credible intervals do not overlap with zero. n represents the number of effect sizes.

4.4.5 Heterogeneity in inbreeding response

Considering minimal models with inbreeding effect sizes, the between-article variance accounted for 43.1% and 5 % of total heterogeneity for fitness component traits and non-fitness component traits, respectively. The among-study variance accounted for 52.5% and 89.3% of total heterogeneity for fitness component traits and non-fitness

component traits. The remaining heterogeneity (4.4% and 5.7% for fitness component traits and non-fitness component traits) was stemmed from *mev*.

Alternatively, in the minimal models with environmental response effect sizes, the between-article variance accounted for 38% and 39.4% of total heterogeneity for fitness component traits and non-fitness component traits, respectively. The among-study variance accounted for 61.8% and 59.9% of total heterogeneity for fitness component traits and non-fitness component traits. The remaining heterogeneity (0.4% and 0.7% for fitness component traits and non-fitness component traits) was attributed to *mev*.

4.5 Discussion

Concern about the possible consequences of ongoing global environmental change for population sustainability has prompted the attempts to synthesise the evidence on the effects of environmental stress on inbreeding depression (Armbruster & Reed, 2005; Fox & Reed, 2011). However, these early reviews were limited in size and scope by the available literature and did not consider the variation of environmental inbreeding depression among types of environments. In the systematic review described in this chapter, I aimed to conduct a comprehensive meta-analysis that gives us an understanding of the variation of environment-dependent inbreeding depression. This study is the first to generalise the variation of environmental factors on inbreeding response. Furthermore, it is the first review to consider the relationship between stress levels and inbreeding depression within populations and between populations separately. Analysing the inbreeding-environmental stress across populations informed us whether inbreeding depression occurs or changes across stress intensity. Alternatively, analysing inbreeding-environmental stress within populations gives us insight into the changes of inbreeding depression after experiencing stress. The key finding of my meta-analyses are that (1) Inbreeding depression occurs across populations and environments. For the fitness component traits, I detected a second order relationship between inbreeding effect sizes and the intensity of stress (Fig. 4), implying that the costs of inbreeding are expected to be greatest at intermediate stress levels. Whereas, for non-fitness components, a positive relationship was identified, indicating increasing inbreeding depression with environmental stress (Fig. 5). (2) After populations experienced a stressful environmental stimulus, I did not find any overall changes in the

magnitude of inbreeding depression, or any relationship of changes in inbreeding repression with the degree of environmental stress, for fitness component traits (Fig. 6). However, for non- fitness component traits, a negative relationship between inbreeding depression changes and stress levels was detected (Fig. 7). (3) When populations experienced two or more environmental changes simultaneously, the impacts of inbreeding depression were more severe (Fig. 9).

4.5.1 Among studies relationship between inbreeding depression and stress levels

The second order polynomial relationship between stress levels and inbreeding depression was first discussed by Springer and his colleagues in 2020 after comparison of their experimental study with Fox and Reed's study. Both of their studies utilised cowpea seed beetles (*Callosobruchus maculatus*) to investigate the effects of environmental stress on inbreeding depression (Fox & Reed, 2011; Springer, Messina, & Gompert, 2020). However, Fox and Reed's study indicated a positive relationship between stress intensity and environment-induced changes in inbreeding depression, for survival traits, whereas Springer failed to detect the interaction between environment and inbreeding depression in the same species. Springer inferred the difference might reflect the magnitude of stress intensity applied in their experiments. In Fox and Reed's study, the intensity of stress they imposed was mild to moderate. At this point, there is still the opportunity for shifts in inbreeding depression and environmental stress to scale positively. However, Springer's study applied more intense stress where stress levels *per se* could force the physiological limiting bounds for the traits observed. Specifically, due to the fact that individuals cannot have fitness values of less than zero, increasing stress intensity towards the limit where non-inbred fitness approaches zero should diminish the expression of inbreeding depression. Similarly, in 2015, Schou and co-workers found the smaller increase in inbreeding depression in *Drosophila* when stress intensity was high (Schou, Loeschcke, & Kristensen, 2015).

Overall, my result in for fitness component traits supported Springer's explanation of physiological limiting bunds and provided a general synthesis on the relationship between stress levels and inbreeding depression across populations and species. When stress levels are mild to moderate, inbreeding depression scales positively with stress

intensity until an intermediate level of stress was reached, when stress level is around 0.49 for fitness components (Fig. 4). After that, inbreeding depression decreased when stress levels increased from there towards the limiting value of one.

However, that limiting bound is absent in non-fitness components. The difference might reflect a difference in genetic architecture between fitness and non-fitness traits (Fig. 5). Morphological traits (or non-fitness components) often undergo stabilising selection for an intermediate optimum, resulting in weakly directional dominance (Charlesworth. & Charlesworth., 1999; DeRose & Roff, 1999; Merila & Sheldon, 1999). Recessive mutations that lead to shifts in trait mean in either direction would be nearly selectively neutral, making them less effective to inbreeding depression (Lynch & Walsh, 1998). In this case, fitness reductions arising from these mutations are smaller. As a result, there is more opportunities for inbreeding depression and environmental stress to interact additively to reduce fitness without reaching physiological limits.

4.5.2 Within study relationship between inbreeding depression and stress levels

Meta-analysis of changes in inbreeding depression associated with environmental stress did not support any consistent relationship between stress intensity and shifts in inbreeding depression associated with fitness component traits.

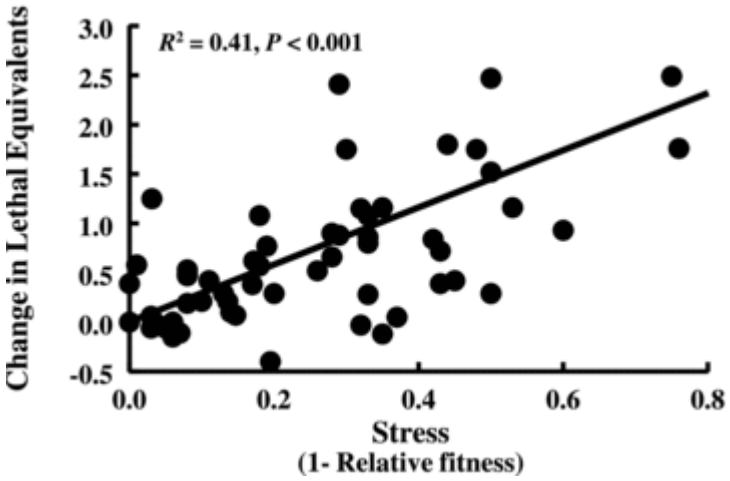


Fig. 12 The association between change in lethal equivalents expressed in the stressful versus benign environments and the intensity of stress in Fox and Reed's study (2011)

This finding overturns our previous understanding that changes in inbreeding depression should scale positively with stress intensity, i.e., greater levels of stress induce greater costs of inbreeding (Armbruster & Reed, 2005; Fox & Reed, 2011). The extent to which there is expected to be a relationship between stress levels and the magnitude of inbreeding depression is highly relevant to conservation practitioners, as it determines whether the costs of inbreeding should be an increasing concern under progressively more stressful environments, and therefore whether and when they should be managed. My result implies that while inbreeding costs undoubtedly exist, and are significant, the expected shift in these costs for fitness traits is zero, regardless of the amount of environmental stress. However, there were significant among-article and within-article variation in this model. The among-article and within-article variation might stem from species-specific or population-specific effects. Therefore, we should still be cautious to draw a conclusion without taking account these case-specific effects. The striking difference between my study and that of Fox & Reed (2011) might be because of large differences in statistical power and study scope. The Fox & Reed meta-analysis (2011) synthesised the impacts of environmental stress on inbreeding depression by comparing the difference in the number of lethal equivalents expressed in the stressful versus benign environment across studies (Fig. 12). However, their review was restricted to a single fitness trait (survival). Other fitness components (e.g., viability and fecundity), which are reliable indicators for population sustainability were not considered in their review. Additionally, their meta-analysis only comprised of 33 articles with 60 effects sizes, which is very small compared with my meta-analysis, which has 129 studies and 1362 effect sizes. There is unlikely to be a difference in the quality of the studies included, as I have used rigorous methods for assessing study relevance against inclusion criteria, as usual in systematic review. Consequently, I argue that my study provides a more rigorous and comprehensive understanding of the interaction between environmental stress and inbreeding depression.

Non-fitness component traits displayed a negative relationship between stress intensity and inbreeding depression, leading to significantly negative shifts in inbreeding depression: as the amount of environmental stress increased, the costs of inbreeding became significantly lower in stressful compared with benign environments (Fig. 7). This decrease in environmental response effect sizes, could indicate the mechanism of

scenario (Fig. 1b), where both inbred and outbred offspring suffer inbreeding depression as stress intensity increases. Therefore, inbreeding depression was only observed in benign environments (introduction; Fig. 1). Specifically, the environmental response effect sizes fall to negative values when stress intensity increases, implying that outbred populations suffer a greater fitness reduction through exposure to environmental stress than inbred populations.

4.5.3 The impacts of environmental factors on inbreeding depression

Another key finding from this meta-analysis is that populations suffer more inbreeding depression when they are exposed to multiple environmental stresses simultaneously compared with abiotic stresses alone regardless of traits measured (Fig. 9). Liao and Reed (2009) reviewed 15 studies (see their discussion) that investigated inbreeding depression in natural populations and found the average number of lethal equivalents to be 2.53 under stressful environments, which is two times higher than Armbruster and Reed's study (1.02), which only retrieved lethal equivalents under manipulated laboratory environments (Liao & Reed, 2009). My results further reinforce the idea that stressors from multiple dimensions found in natural habitats could potentially be more detrimental for synergistically increasing inbreeding costs than simple stresses generated under laboratory conditions. The effects of multiple stressors in natural environments could trigger several inbreeding-related alleles response, resulting in more severe inbreeding depression.

However, there was no significant difference in inbreeding effects sizes between biotic treatments and abiotic or multiple factors treatments, suggesting that the intrinsic characteristics of biotic treatments could be a mixture of multiple biotic factors. For example, a common biotic stressor, within species competition could trigger nutrient and water stress at the same time, depending on the levels of competition researchers applied. Therefore, biotic stressor in some cases might indicate the combination of different abiotic stressors.

4.5.4 Conservation implication

The interaction between inbreeding depression and environmental stress has been recognised as a ubiquitous phenomenon, prompting interest in the consequences for the persistence of populations of conservation concern (Armbruster & Reed, 2005; Fox & Reed, 2011; Kristensen *et al*., 2008; Kristensen, Dahlgaard, & Loeschcke, 2003; Reed, Briscoe, & Frankham, 2002). Conservation biologists are keen to understand not whether an inbreeding-environment interaction exists but whether, and how reliably, this interaction can be translated into effects on population sustainability, by increasing or decreasing population growth or viability. However, my results do not directly indicate the effects of inbreeding-environment interactions on population extinction probability. My meta-analysis fills the knowledge gap by synthesising the effects of inbreeding-environment interaction on fitness reduction, which I expect to contribute to demographic changes in natural populations.

The result of the effects of environmental stress on inbreeding depression across populations (Fig. 4; Fig. 5) suggests that inbreeding-environment interaction are prevalent in the wild, and that these effects, very likely threaten population sustainability via decreasing reproductive output or survival (Fig. 4). In the past, the effects of inbreeding depression on population sustainability have been studied under laboratory conditions or in captivity, in the absence of significant environmental stress (Bijlsma *et al*., 2000; Frankham, 1995; Hedrick & Kalinowski, 2000; Reed *et al*., 2002; Saccheri *et al*., 1998). Since the expected amount of (and expected change in) inbreeding depression varies with the severity of environmental stress, extrapolation of population survival from simple laboratory observations of inbreeding depression is not possible, and might underestimate the phenotypic consequences of inbreeding depression in natural populations of conservation concern. Typically, species of conservation concern in the wild experience multiple stressful stimuli simultaneously (ICUN, 2012). Here, I observed that intersecting sources of environmental stress led to greater fitness reductions as a consequence of inbreeding. (Fig. 9). Future studies focussing on the consequences of inbreeding depression in the wild should not exclude of environmental heterogeneity or global environmental change factors.

In regard to conservation management, reintroduction of captive populations to the wild might still be beneficial, since the levels of inbreeding depression remain similar within populations when populations are exposed to environmental stress. Similarly, application of genetic rescue to mitigate inbreeding depression should be viable because translocation of populations will not deteriorate the negative effects of inbreeding. According to my results, moving populations from benign (captive, native) environments to stressful (natural, foreign) environments does not enhance the severity of inbreeding depression (Fig. 6). Rather, the effects of environmental stress are likely, on average, to outweigh the effects of changes in inbreeding depression when populations are translocated between habitat in conservation management. Therefore, conservation practitioners do not need to worry about the interaction between inbreeding depression and environments when they move populations of conservation concern.

Effective conservation measures that take account of genetic factors are still in their infancy (Frankham, 2010; Hedrick, 2001). Conventionally, population sizes are the major concern when conservation organisation classifies the current state of endangered species. The influence of genetic variation and inbreeding-environment interactions on population sustainability are rarely taken into account (Caughley, 1994; Willi *et al*., 2022). My results show that fitness reductions via inbreeding vary under, and with the level of, environmental stress across populations. However, the magnitude of inbreeding depression did not change in response to environmental shifts. In the face with future climate changes, inbreeding depression or environmental stress *per se* might still be responsible for demographic changes with implications for population persistence. Wherever possible, conservation practitioners could consider genetic issues and environmental stress independently when they design conservation management interventions in the future.

4.6 References

Armbruster, P., & Reed, D. H. (2005). Inbreeding depression in benign and stressful environments. *Heredity, 95*(3), 235-242. doi:10.1038/sj.hdy.6800721

Bijlsma, R., Bundgaard, J., & Boerema, A. C. (2000). Does inbreeding affect the extinction risk of small populations? predictions from Drosophila. *Journal of Evolutionary Biology, 13*(3), 502-514. Retrieved from <Go to ISI>://WOS:000087030000015

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*: John Wiley & Sons, Ltd.

Caughley, G. (1994). Directions in conservation biology. *Journal of Animal Ecology, 63*(2), 215-244. doi:10.2307/5542

Charlesworth. (2003). Effects of inbreeding on the genetic diversity of populations. *Philosophical Transactions of the Royal Society B-Biological Sciences, 358*(1434), 1051-1070. doi:10.1098/rstb.2003.1296

Charlesworth., & Charlesworth. (1987). Inbreeding depression and its evolutionary consequences. *Annual Review of Ecology and Systematics, 18*, 237-268. doi:10.1146/annurev.ecolsys.18.1.237

Charlesworth., & Charlesworth. (1999). The genetic basis of inbreeding depression. *Genetical Research, 74*(3), 329-340. doi:10.1017/s0016672399004152

Charlesworth., & Willis, J. H. (2009). Fundamenal concepts in genetics the genetics of inbreeding depression. *Nature Reviews Genetics, 10*(11), 783-796. doi:10.1038/nrg2664

Cheptou, P. O., & Donohue, K. (2011). Environment-dependent inbreeding depression: its ecological and evolutionary significance. *New Phytologist, 189*(2), 395-407. doi:10.1111/j.1469-8137.2010.03541.x

Crnokrak, P., & Roff, D. A. (1999). Inbreeding depression in the wild. *Heredity, 83*, 260-270. doi:10.1038/sj.hdy.6885530

DeRose, M. A., & Roff, D. A. (1999). A comparison of inbreeding depression in life-history and morphological traits in animals. *Evolution, 53*(4), 1288-1292. doi:10.2307/2640831

Fox, C. W., & Reed, D. H. (2011). Inbreeding depression increases with environmental stress: an experimental study and meta-analysis. *Evolution, 65*(1), 246-258. doi:10.1111/j.1558-5646.2010.01108.x

Frankham, R. (1995). Conservation genetics. *Annual Review of Genetics, 29*, 305-327. doi:10.1146/annurev.ge.29.120195.001513

Frankham, R. (2010). Challenges and opportunities of genetic approaches to biological conservation. *Biological Conservation, 143*(9), 1919-1927. doi:10.1016/j.biocon.2010.05.011

Frankham, R., Ballou, J. D., & Briscoe, D. A. (2010). *Introduction to conservation genetics* (2 nd ed.): Cambridge University press.

Frankham, R., Lees, K., Montgomery, M. E., England, P. R., Lowe, E. H., & Briscoe, D. A. (1999). Do population size bottlenecks reduce evolutionary potential? *Animal Conservation, 2*(4), 255-260. doi:10.1111/j.1469-1795.1999.tb00071.x

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models(Comment on an Article by Browne and Draper). *Bayesian Analysis, 1*(3), 515-533. doi:10.1214/06-ba117a

Hadfield, J. D. (2010). MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. *Journal of Statistical Software, 33*(2), 1-22. doi:10.18637/jss.v033.i02

Hedrick, P. W. (2001). Conservation genetics: where are we now? *Trends in Ecology & Evolution, 16*(11), 629-636. doi:10.1016/s0169-5347(01)02282-0

Hedrick, P. W., & Garcia-Dorado, A. (2016). Understanding Inbreeding Depression, Purging, and Genetic Rescue. *Trends in Ecology & Evolution, 31*(12), 940-952.

doi:10.1016/j.tree.2016.09.005
Hedrick, P. W., & Kalinowski, S. T. (2000). Inbreeding depression in conservation biology. *Annual Review of Ecology and Systematics, 31*, 139-162. doi:10.1146/annurev.ecolsys.31.1.139
Henry, P. Y., Pradel, R., & Jarne, P. (2003). Environment-dependent inbreeding depression in a hermaphroditic freshwater snail. *Journal of Evolutionary Biology, 16*(6), 1211-1222. doi:10.1046/j.1420-9101.2003.00629.x
ICUN. (2012). *IUCN Red list categories and criteria*. Gland, Switzerland and Cambridge, UK: UK: IUCN
Keller, L. F., & Waller, D. M. (2002). Inbreeding effects in wild populations. *Trends in Ecology & Evolution, 17*(5), 230-241. doi:10.1016/s0169-5347(02)02489-8
Koricheva, K., Gurevitch, J., & Mengersen, K. (2013). *Handbook of Meta-analysis in Ecology and Evolution* Princeton University Press.
Kristensen, T. N., Barker, J. S. F., Pedersen, K. S., & Loeschcke, V. (2008). Extreme temperatures increase the deleterious consequences of inbreeding under laboratory and semi-natural conditions. *Proceedings of the Royal Society B-Biological Sciences, 275*(1646), 2055-2061. doi:10.1098/rspb.2008.0426
Kristensen, T. N., Dahlgaard, J., & Loeschcke, V. (2003). Effects of inbreeding and environmental stress on fitness - using Drosophila buzzatii as a model organism. *Conservation Genetics, 4*(4), 453-465. doi:10.1023/a:1024763013798
Kristensen, T. N., Pedersen, K. S., Vermeulen, C. J., & Loeschcke, V. (2010). Research on inbreeding in the 'omic' era. *Trends in Ecology & Evolution, 25*(1), 44-52. doi:10.1016/j.tree.2009.06.014
Kristensen, T. N., Sorensen, P., Pedersen, K. S., Kruhoffer, M., & Loeschcke, V. (2006). Inbreeding by environmental interactions affect gene expression in Drosophila melanogaster. *Genetics, 173*(3), 1329-1336. doi:10.1534/genetics.105.054486
Liao, W., & Reed, D. H. (2009). Inbreeding-environment interactions increase extinction risk. *Animal Conservation, 12*(1), 54-61. doi:10.1111/j.1469-1795.2008.00220.x
Lynch, M., & Walsh, B. (1998). *Genetics and Analysis of Quantitative Traits*: Sinauer.
Merila, J., & Sheldon, B. C. (1999). Genetic architecture of fitness and nonfitness traits: empirical patterns and development of ideas. *Heredity, 83*, 103-109. doi:10.1046/j.1365-2540.1999.00585.x
Neaves, L. E., Eales, J., Whitlock, R., Hollingsworth, P. M., Burke, T., & Pullin, A. S. (2015). The fitness consequences of inbreeding in natural populations and their implications for species conservation - a systematic map. *Environmental Evidence, 4*(1). doi:10.1186/s13750-015-0031-x
Nunez-Mir, G. C., Iannone, B. V., Pijanowski, B. C., Kong, N. N., & Fei, S. L. (2016). Automated content analysis: addressing the big literature challenge in ecology and evolution. *Methods in Ecology and Evolution, 7*(11), 1262-1272. doi:10.1111/2041-210x.12602
Reed, D. H., Briscoe, D. A., & Frankham, R. (2002). Inbreeding and extinction: The effect of environmental stress and lineage. *Conservation Genetics, 3*(3), 301-307. doi:10.1023/a:1019948130263
Rohatgi, A. (2021). WebPlotDigitizer. Retrieved from

https://automeris.io/WebPlotDigitizer
Saccheri, I., Kuussaari, M., Kankare, M., Vikman, P., Fortelius, W., & Hanski, I. (1998). Inbreeding and extinction in a butterfly metapopulation. *Nature, 392*(6675), 491-494. doi:10.1038/33136
Sandner, T. M., & Matthies, D. (2016). The effects of stress intensity and stress type on inbreeding depression in Silene vulgaris. *Evolution, 70*(6), 1225-1238. doi:10.1111/evo.12929
Schou, M. F., Loeschcke, V., & Kristensen, T. N. (2015). Inbreeding depression across a nutritional stress continuum. *Heredity, 115*(1), 56-62. doi:10.1038/hdy.2015.16
Spiegelhalter, D. J., Best, N. G., Carlin, B. R., & van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society Series B-Statistical Methodology, 64*, 583-616. doi:10.1111/1467-9868.00353
Springer, A. L., Messina, F. J., & Gompert, Z. (2020). Measuring the effect of environmental stress on inbreeding depression alone obscures the relative importance of inbreeding-stress interactions on overall fitness inCallosobruchus maculatus. *Evolutionary Applications, 13*(10), 2597-2609. doi:10.1111/eva.13060
R Core Team, R. C. (2019). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/.
Vitousek, P. M., Mooney, H. A., Lubchenco, J., & Melillo, J. M. (1997). Human domination of Earth's ecosystems. *Science, 277*(5325), 494-499. doi:10.1126/science.277.5325.494
Waller, D. M., Dole, J., & Bersch, A. J. (2008). Effects of stress and phenotypic variation on inbreeding depression in Brassica rapa. *Evolution, 62*(4), 917-931. doi:10.1111/j.1558-5646.2008.00325.x
Willi, Y., Kristensen, T. N., Sgro, C. M., Weeks, A. R., Orsted, M., & Hoffmann, A. A. (2022). Conservation genetics as a management tool: The five best-supported paradigms to assist the management of threatened species. *Proceedings of the National Academy of Sciences of the United States of America, 119*(1). doi:10.1073/pnas.2105076119
Winn, A. A., Elle, E., Kalisz, S., Cheptou, P. O., Eckert, C. G., Goodwillie, C., . . . Vallejo-Marin, M. (2011). Analysis of inbreeding depression in mixed-mating plants provides evidence for selective interference and stable mixed mating. *Evolution, 65*(12), 3339-3359. doi:10.1111/j.1558-5646.2011.01462.x

Chapter 5: General discussion and synthesis

5.1 Abstract

Anthropogenic climate change has altered environmental conditions and habitat structures across the globe, leading to population fragmentation and demographic isolation in animals and plants. Inbreeding represents a potential problem for these small and isolated populations, by unmasking recessive deleterious alleles, causing fitness reductions. This phenomenon is called inbreeding depression and the associated fitness reductions can consequently limit population viability. If impacts of environmental change and inbreeding depression have synergistic costs, it will be critical for understanding population sustainability in a conservation context. In this thesis, I have used meta-analysis to synthesise and understand the phenotypic and fitness consequences of inbreeding. In chapter 2, I evaluated the viability of automated content analysis (ACA) to accelerate literature assessment during meta-analysis. I demonstrated the ability of ACA to reduce workload by ~45% with only a limited risk (6%) of losing relevant articles. I also showed the ability of ACA to speed up literature assessment by prioritising the article assessment process. In chapter 3, I investigated the sources of heterogeneity in inbreeding depression. I detected variation in inbreeding responses related to trait types and life-history stage. The association between life-history stage and inbreeding depression also depends on the mating system in plants. In chapter 4, environmental factors and stress were analysed with inbreeding effect sizes, in order to understand the possible consequences of ongoing global environmental changes for organismal phenotypes and fitness during inbreeding. My results overturned our previous understanding that environmental stress increases the magnitude of inbreeding depression. The expected costs of inbreeding are significant, but my analysis provides no evidence that environmental changes will intensify inbreeding depression. Additionally, when examining the association between inbreeding depression and stress intensity across studies, I detect a unimodal (humped) relationship between stress intensity and the intensity of inbreeding depression for fitness traits but for non-fitness traits, stress levels intensify inbreeding depression. In this chapter, I synthesise the results from previous chapters, evaluating practical conservation measures for endangered species. Finally, my results suggest creating gene-flow between populations will be viable options to mitigate the effects of inbreeding depression for conservation

measure since the fitness costs of inbreeding depression when a population experiences environmental changes is negligible.

5.2 Assessment of automated content analysis for literature evaluation during meta-analysis

The need for rigorous evidence to support conservation measures has prompt the adaption of systematic review to ecology and conservation biology from medical sciences (Roberts, Stewart, & Pullin, 2006). However, the exponential growth of published studies in digitised era has made systematic review labour-intensive and time-consuming (Westgate & Lindenmayer, 2017). Text-miming approach has offered a potential solution by reducing the number of articles need to be assessed or accelerating screening process (O'Mara-Eves, Thomas, McNaught, Miwa, & Ananiadou, 2015). In my chapter 2, I constructed an end-to-end pipeline to retrieve relevant articles from a literature corpus, for systematic review. I have showcased how automate content analysis (ACA) could facilitate the process of screening articles by (1) reducing workload with automatic classification (2) accelerating the process with screening prioritisation (Roberts, Stewart, & Tingley, 2019).

The application of text-mining method in the field of ecology and conservation biology is still in its infancy. Majority of studies utilised text-mining approach to discover underlying topics or a trend in a period of time (Altaweel, Bone, & Abrams, 2019; McCallen *et al*., 2019; Nunez-Mir, Iannone, Curtis, & Fei, 2015). A recent study carried out by Cornford and his colleagues (2022) attempted to extract information of taxonomic groups, geographic ranges and estimated population trends from the Living Plant Database. Their result demonstrated the accuracy of retrieving taxonomic and geographic data while poor performance in obtaining in population trends at article-level with ACA (Cornford *et al*., 2022). The limitation in extracting population trend might stem from the linguistic complexity or lack of information in abstracts compared to main texts (Cornford *et al*., 2022). For example, a decreased population might be interpreted as an increased population if a paragraph opens with negative words, such as “slowly increase” or “negative growth”. Overall, this study highlighted the barrier

between text-mining approach and extracting more complex information in synthesising current understanding in the field of ecology in literature database.

Despite the difficulty in summarising our current understanding, my study illustrated the ability of text-mining approach to facilitate the process of performing systematic reviews, allowing researchers to quickly identify the articles requiring for literature synthesis. Unlike Cornford's research that required text-mining approach to detect key information at article-level, my study discovered the most prevalent concepts in a large literature corpus. Therefore, the concepts identified with my ACA process are less vulnerable to the combinations of words in individual articles (or linguistic complexity) since these key concepts were identified on the base of the exclusivity and prevalence of words in the whole database. Additionally, ACA are able to detect an assemblage of words that occurs simultaneously at high frequency in a large literature corpus (Roberts *et al*., 2019). For example, in my inbreeding database, studies focus on inbreeding depression in plants and humans were identified by two different clusters of words (topics), plant, seed, flower for plant-related articles or family, marriage, consanguinity for human-related articles. This ability might be limited when applying textmiming approach at article-level. As a result, the application of ACA at article-level to synthesise knowledge might require the development of advanced algorithms to handle linguistic complexity. However, when it comes to the database-level, my study illustrated the ability of ACA to facilitate literature synthesis by identifying and classifying the topics to individual studies.

The application of text-mining approach on systematic review might be novel and innovative in the field of ecology. However, there were already several studies in the field of medical sciences investigating the utility of text-mining approach on systematic reviews by reducing workload with automatic classification (Cohen, Hersh, Peterson, & Yen, 2006; Jonnalagadda & Petitti, 2013; Wallace, Trikalinos, Lau, Brodley, & Schmid, 2010), speed up the process with screening prioritisation (Cohen, 2008; Cohen, Ambert, & McDonagh, 2012) and functioning as a second screener (Bekhuis, Tseytlin, Mitchell, & Demner-Fushman, 2014; Frunza, Inkpen, & Matwin, 2010). The core mechanisms underlying these text-mining approach above involved two stages: (1) extracting features from a literature database. (2) utilising these features

to build up classifiers with machine-learning algorithms in a training database. (3) applying this trained machine-learning algorithm to classify the rest article in the dataset. The "feature" mentioned above indicates a combination of vectors, usually a group of words that represent a specific biomedical concept.

My ACA-LDA process developed in the Chapter 2 shares the similarity with these text-mining approach investigated in biomedical sciences above. In reality, my method can be considered as an adaption of text-mining techniques from medical sciences in ecology. The features used to build the classifiers were a group of topics (a group of words shares a similar concept) discovered by my stm model (Roberts *et al.*, 2019). However, in my study, the choice of these topics (or features) to include in the classifier requires manual assessment from stm model. In biomedical sciences, the choice of features is frequently based on the terms recorded in MeSH (Medical Subject Headings) and UMLS (Unified Medical Language System)(Bekhuis *et al*., 2014; Cohen, 2008; Cohen *et al*., 2012; Cohen *et al*., 2006; Wallace *et al*., 2010). MeSH and UMLS are two platforms integrate thesaurus and terminology in medical sciences in organised way, in order to promote cataloguing and searching of biomedical information. A study conducted by Aphinyanaphongs (2006) showed a good performance of using MeSH terms as features with a specific machine-learning algorithm. Since we still don't have analogous platform to integrate thesaurus and terminology, like MeSH and UMLS in the field of ecology, assessing topics from the result of stm model with relevant knowledge background might still be necessary for researchers to select appropriate features to build up classifier.

Instead of employing machine-learning algorithms as classifiers like the previous biomedical studies, my study utilised linear discriminant analysis (LDA) and set up cut-off values from LDA results to classify relevant and irrelevant articles. My result demonstrated the viability of using LDA model as a classifier, with 45% of workload reduction accompanied by the risk of losing 6 % of relevant articles. There are some benefits of using LDA as a classifier. First, it allows researchers to adjust its cut-off values. Thus, researchers are able to determine the trade-off between workload reduction and losing of relevant articles, depending on the needs for their systematic reviews. Second, it provides an all-in-one function for automatic classification and screening

prioritisation. Finally, it would be relatively straight forward and intuitive for ecologists without computing background and knowledges, comparing to these machine-learning algorithms. However, it is still questionable if these machine-learning algorithms have better performance than my LDA results. We cannot still rule out the potential benefits of these machine-learning algorithms to facilitate systematic reviews. As a result, for future research I would recommend applying different text-mining approaches on systematic review in the field of ecology to discover other approaches' features and abilities. Lastly, for the long-term goal, we might still need to construct similar platforms, like MeSH, which allows researchers to index and classify ecological and conservation-related information for text-mining approaches.

5.3 The sources of heterogeneity in inbreeding depression

The adverse effects of inbreeding depression have been recognised as a ubiquitous phenomenon, fuelling concern in the consequences for population sustainability in conservation context (Frankham, 1995; Keller & Waller, 2002). To bridge the knowledge gap between inbreeding depression and its implications for population demography in natural population, there is a need to disentangle the heterogeneity underlying the phenotypic consequences of inbreeding. Several empirical studies have tried to investigate the associations between traits (Brzeski, Rabon, Chamberlain, Waits, & Taylor, 2014; Nielsen *et al*., 2012; Pico, Mix, Ouborg, & Van Groenendael, 2007; Roff, 1998; Thiele, Hansen, Siegismund, & Hauser, 2010; Vange, 2002), mating systems (Agren & Schemske, 1993), population size (Michaels, Shi, & Mitchell, 2008; Ouborg & Vantreuren, 1994) or life history stages (Ishida, 2008; Pico, Ouborg, & Van Groenendael, 2004; Thiele *et al*., 2010) and the outcome of inbreeding.

Angeloni and co-workers (2011) conducted a meta-analysis between inbreeding depression and population size, traits, species characteristics (tree, shrub and herb) and mating systems in plants. Additionally, DeRose and Roff (1999) compared the slopes of inbreeding coefficient between life-history traits and morphological traits in animals and detected life-history traits suffered significantly more inbreeding depression than morphological traits. Furthermore, Husband and Schemske (1996) investigated the interaction between mating systems and life history stages in plants and found selfing species expressed the majority of inbreeding depression in late history stage whereas

outcrossing species experienced across different life history stages. Among three of these reviews, only Angeloni's studies (2011) used meta-analysis and weighed the effect sizes appropriately to avoid bias from the contributions of the studies with small sample sizes. DeRose and Roff study (1999) applied Kruskal-Wallis tests while Husband and Schemske study (1996) performed Mann-Whitney U-tests. Both of these studies failed to control within-study variance (*mev*). Moreover, all of these studies limited their scopes in either plants or animals.

My study is the first one to systematically synthesise the impacts of various factors on inbreeding depression in both animals and plants. The goal is to provide a fundamental framework that allows conservation biologists to comprehensively understand the heterogeneity of inbreeding depression in conservation concerns. The results also deliver a preliminary understanding when I investigated the effects of global environmental changes on inbreeding depression in chapter 4. These factors included trait type, fitness class, taxonomic group, life history stage, life expectancy, and mating systems.

Overall, there is no heterogeneity in inbreeding depression among taxonomic group or *mating system 1* (see chapter 3, source of heterogeneity). Additionally, I did not detect any association between life expectancy and inbreeding effect sizes. Inbreeding coefficient is the most important factor to predict the severity of inbreeding depression. Inbreeding depression also occurred in most types of traits. The only exception was behaviour, which did not show inbreeding depression when I incorporated trait type as a variable into my inbreeding model. The possible reasons were already discussed in the chapter 3. Briefly, behaviours are frequently controlled by multiple behavioural components, a deleterious mutation exposed by inbreeding would in one behavioural component have relatively less impact if other components are still able to function properly (Arbuthnott, 2009; Mackay *et al*., 2005). Additionally, behaviour has relatively large degree of plasticity, allowing to mitigate inferior phenotypes arisen from inbreeding depression (Reale, Dingemanse, Kazem, & Wright, 2010; Reale & Roff, 2002). Lastly, behaviour is believed to be mainly maintained by balancing selection (Boon, Reale, & Boutin, 2007; Dingemanse, Kazem, Reale, & Wright, 2010; Wolf, van Doorn, Leimar, & Weissing, 2007). Recessive mutations that shift in either direction of trait mean will be selectively neutral (Lynch & Walsh, 1998). This result

suggests conservation biologists should be cautious if they draw a conclusion on the phenotypic consequence of inbreeding with behavioural traits this might underestimate the adverse effects of inbreeding that populations experience in the wild.

Another critical finding is that I did not detect the variation of inbreeding depression between fitness component traits and non-fitness component traits. Traditionally, it is well acknowledged that fitness traits suffer more inbreeding depression than non-fitness traits (Falconer, 1989). The general explanation underlying this statement is the genetic architecture of fitness traits have high ratio of dominance versus additive variance and low heritability (Merila & Sheldon, 1999). However, my result only partially agreed with this statement. In particular, only fitness and fecundity experienced significantly more inbreeding depression in comparison with other traits except defence and growth rate. Other fitness-related traits, such as survival and viability were not significantly different from the majority of other non-fitness traits. Therefore, my result could suggest that fitness traits have high ratio of dominance versus additive variance and low heritability might not be a general pattern in the nature. A narrative review conducted by Merila and Sheldon (1999) indicated that some fitness traits harbour high levels of additive genetic variation despite these traits typically have low heritability. Additionally, high intensity in inbreeding depression in fecundity might be corresponding to the result that the traits in late life history stage suffered more inbreeding depression than the traits in early and middle life history stages. My models indicated more severity of inbreeding depression in late life history stages, possibly due to the removal of deleterious alleles in early or middle life history stages by natural selection (Hamilton, 1966; Williams, 1957). Other fitness component traits such as survival could occur in middle or late life history stages whereas fecundity only occurs after sexual maturity, which is defined at late life history stage. Lastly, the categorisation of fitness and non-fitness component traits could be arbitrary, given that we have limited information about the form of selection on these traits in nature. Therefore, in order to translate the risk of phenotypic outcomes of inbreeding into population sustainability, future research should avoid drawing a conclusion of inbreeding depression from a single trait.

5.4 The inbreeding response of populations to environmental changes

The interaction between environmental stress and inbreeding caused by isolated and fragmentary population size has raised a conservation concern whether small populations are able to persist in the face with global environmental changes. This concern has prompted the investigation on the influence of environmental variation on phenotypic consequences of inbreeding.

An early synthesis conducted by Armbruster and Reed (2005) compared inbreeding depression between benign and stressful environments across 34 studies. They detected that in 48% of studies, inbreeding depression increased significantly in stressful environments in compared to benign environments. Several years later, Fox and Reed (2011) performed a meta-analysis, investigating the relationship between environmental stress and inbreeding depression with 33 articles. They identified that the changes in magnitudes of inbreeding depression scaled linearly with the intensity of environmental stress. Despite both of these studies suffered limited scopes and statistical problems (which has been discussed in the chapter 4), they provided a preliminary framework that environmental stress could deteriorate inbreeding depression. Besides, even if both of studies suggested more intense of inbreeding depression under stressful environments, the interpretations of their results should not be considered the same. Armbruster and Reed compared the levels of inbreeding depression across populations and studies while Fox and Reed's meta-analysis quantified the change in inbreeding depression in response to stress intensity when population experienced environmental treatments.

My meta-analysis investigated the magnitude of inbreeding depression across stress levels and populations and how it changes in response to environmental stress within population. A remarkable finding is a second-order polynomial relationship between stress intensity and inbreeding depression across populations in fitness component traits. There were several attempts tried to explain the interaction between stress intensity and inbreeding depression. One explanation is that deleterious alleles have smaller selection coefficient in benign environments and therefore inbred offspring have similar fitness values to outbred offspring, resulting in minor inbreeding depression (Cheptou & Donohue, 2011; Fox & Reed, 2011). When magnitude of stress

intensifies, selection coefficient of deleterious alleles in inbred offspring increases, subsequently deteriorating inbreeding depression. However, this explanation also exists a threshold effect where both inbred and outbred offspring fail to adapt to high intensity of environmental stress, leading to the decline in inbreeding depression. In fitness component traits, the second order polynomial relationship between inbreeding depression and stress level indicated the existence of this threshold effect. However, this pattern was absent in non-fitness component traits. In non-fitness component traits, I detected a positive association between stress intensity and inbreeding depression across populations. In chapter 3, my statistical model did not detect any significant difference between fitness component traits and non-fitness component traits without taking account into stress intensity. However, when stress intensity increases, inbreeding depression in these two types of traits displayed contrasting trends, suggesting the difference of selection pattern across stress levels between fitness component and non-fitness component traits. Therefore, for the future studies, researchers should still be cautious about interpreting the potential outcomes of inbreeding between fitness and non-fitness component traits since zero intensity of environmental stress will be extremely rare in the reality.

When it comes to the changes in inbreeding depression in response to environmental treatments within population, my results overturned our previous understanding suggested by Fox and Reed (2011). In their meta-analysis, environmental stress scaled positively with the changes in inbreeding depression. Populations experienced stressful treatments suffered more severe inbreeding depression. In contrast, my finding showed that increasing environmental stress did not change the magnitude of inbreeding depression. In particular, the relative fitness difference between inbred and outbred offspring did not change in response to stress intensity. The striking difference between my meta-analysis and Fox and Reed's might stem from the study scopes and difference in statistical power. Fox and Reed's study included 33 published papers and only used survival to measure inbreeding depression whereas my meta-analysis comprised of 129 studies and covered a variety of fitness component and non-fitness component traits. This might provide a broader understanding of demographic outcomes of inbreeding under environmental shifts since other fitness-component traits such as fecundity and viability also influence demography changes in natural populations. In

addition, the effect sizes of Fox and Reed's meta-analysis were not weighed by sample sizes, making their results vulnerable to the bias stemming from small studies. Therefore, I argue my result provide more rigorous evidence of the changes in inbreeding depression after populations experiences environmental stress. However, there is a caveat to this conclusion. Due to the high proportions of heterogeneity from between study and between article variations, we cannot consider this as a universal rule. It is still highly possible to detect increases or decreases in inbreeding depression after populations suffer environmental stress for individual cases.

However, for non-fitness component traits, it showed a different pattern. Increasing stress intensity decreased the change in inbreeding depression. When stress level kept rising, the severity of inbreeding depression in benign environments overtook the severity of inbreeding depression in stressful environments, leading to the negative values of the effect sizes. In summary, the changes in inbreeding depression for fitness and non-fitness component traits can be illustrated by the following 2 scenario (1) the fitness values of inbred and outbred offspring decrease at the same pace when stress intensity increases, leading to the similar levels of inbreeding depression in benign and stressful environments (2) both inbred and outbred fail to adapt to the stressful environments, resulting lower inbreeding depression compared to the benign environments. Incorporating the finding from the among studies analysis, inbreeding depression always exists regardless of environmental stress across. Thus, even if inbreeding depression is decreased in the stressful environments, we can still detect the gap between the fitness values of inbred and outbred offspring.

In addition to the environmental stress, my results also offered an insight into the influences of different types of environmental factors on inbreeding depression. When analysing the causes of environmental factors on inbreeding depression (category 1, Table 2. Chapter 4), I did not detect any variation of inbreeding effect sizes in response to environmental factors across populations. Instead, significant difference was found when we categorised environmental factors into abiotic, biotic and multiple factors. When populations experience different environmental factors (multiple factors) simultaneously (category 2, Table 2. Chapter 4), they suffered more inbreeding depression in comparison to the populations which only experienced single abiotic stress. This

indicates that studying the interaction between inbreeding depression and environmental stress under laboratory environments with single environmental treatment might underestimate the real magnitude of inbreeding depression that natural populations suffered in the wild. Alternatively, when analysing the causes of environmental factors on changes in inbreeding depression after environmental treatments within populations, I did not detect any variation of environmental responses effect sizes regarding two categories of environmental factors (Table 2. Chapter 4). This finding implied that the level of inbreeding depression will not changes within populations in response to the types of environmental shifts they experienced.

In summary, even though the above findings cannot be translated into extinction probability of small, isolated populations, my results offered a connection between inbreeding depression and population sustainability under the influences of ongoing global climate changes. The models suggested environmental stress or inbreeding depression independently threaten population sustainability. Inbreeding depression threatens population sustainability via reducing survival rates and reproductive output. The impacts of inbreeding depression in reproductive output are more severe than survival or viability. Additionally, the types of environmental changes, for example, global warming or increasing frequency of drought, will not affect the severity of inbreeding depression in fitness component traits but these types of changes could potentially reduce individual fitness in both inbred and outbred offspring. Therefore, conservation practitioners should not ignore the adverse effects of inbreeding depression in natural populations.

5.5 Conservation implications

The influences of anthropogenic climate changes have posed an issue that whether small and isolated population can persist with genetic issues under a changing word. The ongoing global climate changes and increasing human activities have decreased habitats available for animals and plants, leading to isolated and fragmented populations (Vitousek, Mooney, Lubchenco, & Melillo, 1997). When populations are fragmented and isolated, gene flow is limited, inbreeding and genetic drift are more likely to happen (Frankham, 1995; Hedrick, 2001; Keller & Waller, 2002). Consequently, loss of genetic diversity and reduction in fitness might hamper the populations to

survive with ongoing environmental changes. To prevent populations from extinction, conservation practitioners typically applied the following genetic management to mitigate inbreeding depression and enhance genetic diversity in isolated and fragmented populations (Frankham, Ballou, & Briscoe, 2010).

(1) Increasing the rate of gene flow by translocation such as genetic rescue or creating habitat corridors.
(2) Reintroduction of populations to natural habitats where they have been locally extinct.

There are several cases that illustrated the benefits of translocating between populations to reverse the adverse effects of inbreeding depression such as Florida panther (*Felis concolor coryi*) and adders (*Vipera berus*) in southern Sweden (Johnson *et al*., 2010; Madsen, Stille, & Shine, 1996). However, there are some concerns that could potentially impede the success of genetic rescue due to the occurrences of outbreeding depression and the loss of local adaptation.

5.5.1 Implication for genetic rescue and gene flow

My result indicates that the inbreeding coefficient is the decisive predictor of the severity of inbreeding depression. This suggests reinforcing gene flower or maintaining habitat connections to mitigate inbreeding depression is essential to recover individual fitness in natural populations. Genetic rescue is one of the conservation managements to reinforce gene flow. However, conservation practitioners might be concerned about the swamping of local adapting genotypes by alien alleles, outbreeding depression and whether translocated population can adapt to foreign habitats.

When implementing genetic rescue, conservation practitioners are eager to find a balance between the loss of heterozygosity within populations and the divergence in allele frequencies among populations. Insufficient gene flow might fail to recover individual fitness in fragmented and isolated populations. Conversely, too many migrants could potentially swamp locally adapted genotypes, subsequently leading to the loss of local alleles. A common rule of thumb to reinforce gene flow is the one-migration-per-generation rule proposed by Wright's work on the infinite-island model (1931). Several empirical studies have suggested one-migration-per-generation could improve the

fitness of subpopulations. An experiment conducted by Spielman and Frankham (1992) found that inbred *Drosophila melanogaster* significantly increased reproduction after receiving one migration per generation. Similarly, Newman's experiment (1996) indicated a significantly positive effect for 5 fitness components in the subpopulations of *Brassica campestris* after they received one migration per generation. These examples illustrated the importance of genetic connectivity to maintain fitness in populations.

Outbreeding depression leads to fitness reductions due to the crossing between populations. A meta-analysis conducted by Whitlock and his colleagues (2013) demonstrated that there are phenotypic benefits of outbreeding in the first generation of outbred offspring. However, fitness costs began to appear in the following generations. Despite the outbreeding effects in later generations, the between-study heterogeneity in their meta-analysis models indicated outbreeding is not a universal phenomenon. In 2011, Frankham and co-workers developed a decision tree based on Breeder's equation suggested that outbreeding depression was more liable to occurs when two populations did not exchange genes in the last 500 years or inhabited different environments. As a result, conservation practitioners should not ignore the benefits of genetic rescue due the possibility of outbreeding depression. The studies above provided information to prevent the potential adverse effects of outbreeding depression under specific circumstances.

Another concern of the opposite effects of genetic rescue is translocated populations could not adapt to the new environments. Inbreeding and genetic diversity are primary genetic issues in adaptation to the foreign environments in translocated populations (Frankham *et al*., 2010). Foreign environments typically represent more stressful environments compared to their native environments. Maladaptation to the new environment will result in the failure of genetic rescue. Additionally, whether environmental stress will deteriorate inbreeding depression in translocated populations is still a concern in executing genetic rescue. My result imply that the magnitude of inbreeding depression remains unchanged after population experienced environmental stress in fitness component traits. However, individual fitness did decline in response to environmental stress and inbreeding depression happened across stress intensity. Therefore,

conservation practitioners should be able to consider the risk of inbreeding depression or environmental stress independently when they conduct genetic rescue.

Furthermore, a meta-analysis carried out by Frankham (2015), synthesising the effects of genetic rescue over 156 published articles. His study indicating 92.9% of these cases benefited from genetic rescue and the effects were more significant in stressful environments compared to benign environments. Taken together with my result, genetic rescue should still be a beneficial and viable measure to alleviate the negative effects of inbreeding depression. Conservation practitioners should not worry about the joint effects of inbreeding depression and environmental changes.

5.5.2 Implication for reintroduction programme

Conventionally, it is well acknowledged that the effects of inbreeding depression are more severe in harsher, wild environments than in captivity. This has posed an issue that reintroduction of captive population to the wild will deteriorate inbreeding depression, subsequently leading to the failure of reintroduction programme. Additionally, captive populations are more likely to suffer from inbreeding depression due to small population sizes (Frankham *et al*., 2010). Whether reintroduce captive populations to the wild will worsen inbreeding depression is still questionable. Similar to the implications from genetic rescue, my finding indicated that increasing environmental stress did not deteriorate inbreeding depression. In specific, the severity of inbreeding depression would not increase when moving populations from captive environments to natural habitats. However, conservation practitioners should still try to minimise inbreeding since fitness reductions due to inbreeding still occur across stress intensity. Once inbreeding is minimised in captive populations, the viability of reintroduction programme should be relied on the adaptation of captive populations to their natural environments.

5.6 Review limitations

Overall, my meta-analysis broadly addresses how inbreeding depression varies under different environmental conditions by synthesising phenotypic consequences of inbreeding in related to different sources of heterogeneity and environmental variations.

However, there are still some limitations in my meta-analysis. To begin with, the fitness reduction caused by inbreeding depression can not be translated into population trajectory directly. Some of the phenotypic outcomes which were measured are of demographic relevance. However, my analyses cannot predict the growth or demographic outcome of natural populations. This is because population persistence is balanced between different demographic processes, including recruitment, mortality and migration. The adverse effects of inbreeding depression on individual fecundity and survival in my study (fitness) could potentially influence recruitment in natural populations, subsequently shaping population demography. In a situation inbreeding depression negatively impacts fitness, the expectation is that population persistence will be undermined. However, additional data on demography are likely to be required to forecast population growth rate in any given case. Furthermore, the large proportions of between-study and between-article heterogeneity in the models of environmental response effect sizes and stress intensity indicated there might be some cases did not follow this pattern due to population-specific and study-specific effects. This study and article variations should be taken account carefully when interpreting how inbreeding depression changes in response to environmental stress within populations. Additionally, it's impossible to include every source of heterogeneity in my meta-analysis due to the limitation to access available data. For example, population size has been reported to influence the variation in inbreeding depression (Angeloni *et al*., 2011). However, only minority of studies reported this information in their studies, consequently hindering the generalisation between population size and inbreeding depression. Lastly, my synthesis was limited to the inbreeding coefficient calculated from pedigree data. As a result, we cannot generalise the phenotypic outcomes of inbreeding depression measured by F_{IS}, the deviation of the observed heterozygosity frequency relative to the expected heterozygosity frequency under random mating.

5.7 Conclusion

Ongoing global climate change is an immediate threat to natural populations, especially to small and isolated populations. Genetic drift and inbreeding are more likely to happen in these populations, subsequently leading to the erosion of genetic diversity and fitness reductions. As a result, population might suffer from demographic declines and restriction of evolutionary potential. Conservation biologists are keen to

understand whether small and isolated populations are able to persist with these genetic threats under anthropogenic environmental changes. My thesis utilised meta-analysis to synthesise phenotypic consequence of inbreeding in related to types of traits, species characteristics and environmental changes. My results showed that inbreeding depression is more severe in fecundity and fitness and every phenotypic trait except behavioural traits suffered inbreeding depression. This indicated that inbreeding depression could still affect population persistence via demographic changes, such as reproductive output and survival. Additionally, traits at late life history stage suffered more inbreeding depression than early and middle life history stage and significant interaction between life history stage and *mating system 2* was observed in plants. This suggests that natural selection could still be a potential factor influencing the heterogeneity in inbreeding depression by the removal of deleterious alleles.

Moreover, my synthesis offers an insight into how climate variation could interact with inbreeding depression in small and isolated populations. I showed that inbreeding depression did happen across populations and species regardless of environmental stress even though fitness and non-fitness component traits displayed different patterns of changes in inbreeding depression. When populations experienced environmental stress, the magnitude of inbreeding depression in fitness component traits did not change within populations. This finding implied that we should not worry about worsening inbreeding depression when we translocate populations for conservation purpose. The harmful effects of environmental changes and inbreeding depression might be considered independently. Lastly, my models also suggested that the types of environmental factors did not predict the level of inbreeding depression. However, when populations experienced multiple dimensions of environmental stress, they might suffer more severity of inbreeding depression. As a result, future researcher should be more conservative when they extrapolate the fitness consequences of inbreeding depression under uniform environmental conditions.

5.7.1 Take-home messages

For conservation practitioners here are the key suggestions for future conservation management: First, inbreeding depression is an intrinsic component within populations. It typically reduces individual fitness regardless of traits and taxonomic groups.

Inbreeding *per se*, which indicated by inbreeding coefficient is the most important factors determining fitness costs of inbreeding depression in natural populations. Therefore, gene flow or habitat connection is crucial to mitigate inbreeding depression and maintain population sustainability. Second, in most scenarios, inbreeding depression does not increase in response to environmental stress in fitness component traits. Population moved to a new habitat or experienced environmental stress or does not deteriorate inbreeding depression. Therefore, genetic rescue is still a viable option to mitigate the negative outcomes of inbreeding. Conservation practitioners should be able to consider the threats from inbreeding and environmental stress separately.

5.7.2 Future directions

In summary, my study has synthesised that inbreeding depression and environmental stress might threaten population sustainability independently via fitness reduction. However, a critical question regarding whether inbred populations are able to persist under climate changes is their ability to adapt and evolutionary potential. Genetic diversity plays a pivotal role underpinned adaptability and evolutionary potential for natural populations. As a result, two connections remain unsolved: (1) the relationship between inbreeding depression and genetic diversity; (2) the adaptability of inbred populations under stressful environments. Therefore, for the future empirical studies, I recommend researchers to record genetic variation and population sizes, in some cases a proxy for genetic diversity in their inbreeding studies to facilitate future reviews. For the future review, I propose two directions: (1) to synthesise the relationship between genetic variation and inbreeding depression; (2) to generalise the relationship between adaptability to stressful environments in related to inbreeding depression. By doing that, conservation biologists will be able to gain comprehensive understandings of population sustainability of small and isolated populations under a changing word.

5.8 Reference

Agren, J., & Schemske, D. W. (1993). Outcrossing rate and inbreeding depression in 2 annual monoecious herbs, Begonia-hirsuta and B-semiovata. *Evolution, 47*(1), 125-135. doi:10.1111/j.1558-5646.1993.tb01204.x

Altaweel, M., Bone, C., & Abrams, J. (2019). Documents as data: A content analysis and topic modeling approach for analyzing responses to ecological disturbances. *Ecological Informatics, 51*, 82-95.

doi:10.1016/j.ecoinf.2019.02.014
Angeloni, F., Ouborg, N. J., & Leimu, R. (2011). Meta-analysis on the association of population size and life history with inbreeding depression in plants. *Biological Conservation, 144*(1), 35-43. doi:10.1016/j.biocon.2010.08.016
Aphinyanaphongs, Y., Statnikov, A., & Aliferis, C. F. (2006). A comparison of citation metrics to machine learning filters for the identification of high quality MEDLINE documents. *Journal of the American Medical Informatics Association, 13*(4), 446-455. doi:10.1197/jamia.M2031
Arbuthnott, D. (2009). The genetic architecture of insect courtship behavior and premating isolation. *Heredity, 103*(1), 15-22. doi:10.1038/hdy.2009.22
Armbruster, P., & Reed, D. H. (2005). Inbreeding depression in benign and stressful environments. *Heredity, 95*(3), 235-242. doi:10.1038/sj.hdy.6800721
Bekhuis, T., Tseytlin, E., Mitchell, K. J., & Demner-Fushman, D. (2014). Feature Engineering and a Proposed Decision-Support System for Systematic Reviewers of Medical Evidence. *Plos One, 9*(1). doi:10.1371/journal.pone.0086277
Boon, A. K., Reale, D., & Boutin, S. (2007). The interaction between personality, offspring fitness and food abundance in North American red squirrels. *Ecology Letters, 10*(11), 1094-1104. doi:10.1111/j.1461-0248.2007.01106.x
Brzeski, K. E., Rabon, D. R., Chamberlain, M. J., Waits, L. P., & Taylor, S. S. (2014). Inbreeding and inbreeding depression in endangered red wolves (Canis rufus). *Molecular Ecology, 23*(17), 4241-4255. doi:10.1111/mec.12871
Cheptou, P. O., & Donohue, K. (2011). Environment-dependent inbreeding depression: its ecological and evolutionary significance. *New Phytologist, 189*(2), 395-407. doi:10.1111/j.1469-8137.2010.03541.x
Cohen, A. M. (2008). Optimizing feature representation for automated systematic review work prioritization. *AMIA Annu Symp Proc, 2008*, 121-125.
Cohen, A. M., Ambert, K., & McDonagh, M. (2012). Studying the potential impact of automated document classification on scheduling a systematic review update. *Bmc Medical Informatics and Decision Making, 12*. doi:10.1186/1472-6947-12-33
Cohen, A. M., Hersh, W. R., Peterson, K., & Yen, P. Y. (2006). Reducing workload in systematic review preparation using automated citation classification. *Journal of the American Medical Informatics Association, 13*(2), 206-219. doi:10.1197/jamia.M1929
Cornford, R., Millard, J., Gonzalez-Suarez, M., Freeman, R., & Johnson, T. F. (2022). Automated synthesis of biodiversity knowledge requires better tools and standardised research output. *Ecography, 2022*(3). doi:10.1111/ecog.06068
DeRose, M. A., & Roff, D. A. (1999). A comparison of inbreeding depression in life-history and morphological traits in animals. *Evolution, 53*(4), 1288-1292. doi:10.2307/2640831
Dingemanse, N. J., Kazem, A. J. N., Reale, D., & Wright, J. (2010). Behavioural reaction norms: animal personality meets individual plasticity. *Trends in Ecology & Evolution, 25*(2), 81-89. doi:10.1016/j.tree.2009.07.013
Falconer, D. S. (1989). *An introduction to quantitative genetics* (3 rd ed.): Wiley.
Fox, C. W., & Reed, D. H. (2011). Inbreeding depression increases with environmental stress: an experimental study and meta-analysis. *Evolution,*

65(1), 246-258. doi:10.1111/j.1558-5646.2010.01108.x
Frankham, R. (1995). Conservation genetics. *Annual Review of Genetics, 29*, 305-327. doi:10.1146/annurev.ge.29.120195.001513
Frankham, R. (2015). Genetic rescue of small inbred populations: meta-analysis reveals large and consistent benefits of gene flow. *Molecular Ecology, 24*(11), 2610-2618. doi:10.1111/mec.13139
Frankham, R., Ballou, J. D., & Briscoe, D. A. (2010). *Introduction to conservation genetics* (2 nd ed.): Cambridge University press.
Frankham, R., Ballou, J. D., Eldridge, M. D. B., Lacy, R. C., Ralls, K., Dudash, M. R., & Fenster, C. B. (2011). Predicting the probability of outbreeding depression. *Conservation Biology, 25*(3), 465-475. doi:10.1111/j.1523-1739.2011.01662.x
Frunza, O., Inkpen, D., & Matwin, S. (2010). *Building systematic reviews using automatic text classification techniques.* Paper presented at the Proceedings of the 23rd International Conference on Computational Linguistics: Posters, Beijing.
Hamilton, W. D. (1966). The moulding of senescence by natural selection. *Journal of Theoretical Biology, 12*(1), 12-45. doi:https://doi.org/10.1016/0022-5193(66)90184-6
Hedrick, P. W. (2001). Conservation genetics: where are we now? *Trends in Ecology & Evolution, 16*(11), 629-636. doi:10.1016/s0169-5347(01)02282-0
Husband, B. C., & Schemske, D. W. (1996). Evolution of the magnitude and timing of inbreeding depression in plants. *Evolution, 50*(1), 54-70. doi:10.1111/j.1558-5646.1996.tb04472.x
Ishida, K. (2008). Effects of inbreeding on the magnitude of inbreeding depression in a highly self-fertilizing tree, Magnolia obovata. *Ecological Research, 23*(6), 995-1003. doi:10.1007/s11284-008-0467-3
Johnson, W. E., Onorato, D. P., Roelke, M. E., Land, E. D., Cunningham, M., Belden, R. C., . . . O'Brien, S. J. (2010). Genetic restoration of the florida panther. *Science, 329*(5999), 1641-1645. doi:10.1126/science.1192891
Jonnalagadda, S., & Petitti, D. (2013). A new iterative method to reduce workload in systematic review process. *Int J Comput Biol Drug Des, 6*(1-2), 5-17. doi:10.1504/ijcbdd.2013.052198
Keller, L. F., & Waller, D. M. (2002). Inbreeding effects in wild populations. *Trends in Ecology & Evolution, 17*(5), 230-241. doi:10.1016/s0169-5347(02)02489-8
Lynch, M., & Walsh, B. (1998). *Genetics and Analysis of Quantitative Traits*: Sinauer.
Mackay, T. F. C., Heinsohn, S. L., Lyman, R. F., Moehring, A. J., Morgan, T. J., & Rollmann, S. M. (2005). Genetics and genomics of Drosophila mating behavior. *Proceedings of the National Academy of Sciences of the United States of America, 102*, 6622-6629. doi:10.1073/pnas.0501986102
Madsen, T., Stille, B., & Shine, R. (1996). Inbreeding depression in an isolated population of adders Vipera berus. *Biological Conservation, 75*(2), 113-118. doi:10.1016/0006-3207(95)00067-4
McCallen, E., Knott, J., Nunez-Mir, G., Taylor, B., Jo, I., & Fei, S. L. (2019). Trends in ecology: shifts in ecological research themes over the past four decades. *Frontiers in Ecology and the Environment, 17*(2), 109-116. doi:10.1002/fee.1993
Merila, J., & Sheldon, B. C. (1999). Genetic architecture of fitness and nonfitness

traits: empirical patterns and development of ideas. *Heredity, 83*, 103-109. doi:10.1046/j.1365-2540.1999.00585.x

Michaels, H. J., Shi, X. J., & Mitchell, R. J. (2008). Effects of population size on performance and inbreeding depression in Lupinus perennis. *Oecologia, 154*(4), 651-661. doi:10.1007/s00442-007-0861-5

Newman, D. K. (1996). Importance of genetics on survival of small populations: genetic drift, inbreeding, and migration. Ph.D. dissertation. University of Montana, Missoula.

Nielsen, J. F., English, S., Goodall-Copestake, W. P., Wang, J. L., Walling, C. A., Bateman, A. W., . . . Pemberton, J. M. (2012). Inbreeding and inbreeding depression of early life traits in a cooperative mammal. *Molecular Ecology, 21*(11), 2788-2804. doi:10.1111/j.1365-294X.2012.05565.x

Nunez-Mir, G. C., Iannone, B. V., Curtis, K., & Fei, S. L. (2015). Evaluating the evolution of forest restoration research in a changing world: a "big literature" review. *New Forests, 46*(5-6), 669-682. doi:10.1007/s11056-015-9503-7

O'Mara-Eves, A., Thomas, J., McNaught, J., Miwa, M., & Ananiadou, S. (2015). Using text mining for study identification in systematic reviews: a systematic review of current approaches. *Systematic Reviews, 4*. doi:10.1186/2046-4053-4-5

Ouborg, N. J., & Vantreuren, R. (1994). The significance of genetic erosion in the process of extinction .4. inbreeding load and heterosis in relation to population-size in the mint Salvia-pratensis. *Evolution, 48*(4), 996-1008. doi:10.2307/2410361

Pico, F. X., Mix, C., Ouborg, N. J., & Van Groenendael, J. M. (2007). Multigenerational inbreeding in Succisa pratensis: effects on fitness components. *Biologia Plantarum, 51*(1), 185-188. doi:10.1007/s10535-007-0037-5

Pico, F. X., Ouborg, N. J., & Van Groenendael, J. M. (2004). Evaluation of the extent of among-family variation in inbreeding depression in the perennial herb Scabiosa columbaria (Dipsacaceae). *American Journal of Botany, 91*(8), 1183-1189. doi:10.3732/ajb.91.8.1183

Reale, D., Dingemanse, N. J., Kazem, A. J. N., & Wright, J. (2010). Evolutionary and ecological approaches to the study of personality. *Philosophical Transactions of the Royal Society B-Biological Sciences, 365*(1560), 3937-3946. doi:10.1098/rstb.2010.0222

Reale, D., & Roff, D. A. (2002). Quantitative genetics of oviposition behaviour and interactions among oviposition traits in the sand cricket. *Animal Behaviour, 64*, 397-406. doi:10.1006/anbe.2002.3084

Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). stm: An R Package for Structural Topic Models. *Journal of Statistical Software, 91*(2), 1-40. doi:10.18637/jss.v091.i02

Roberts, P. D., Stewart, G. B., & Pullin, A. S. (2006). Are review articles a reliable source of evidence to support conservation and environmental management? A comparison with medicine. *Biological Conservation, 132*(4), 409-423. doi:10.1016/j.biocon.2006.04.034

Roff, D. A. (1998). Effects of inbreeding on morphological and life history traits of the sand cricket, Gryllus firmus. *Heredity, 81*, 28-37. doi:10.1038/sj.hdy.6883630

Spielman, D., & R. Frankham. (1992). Modeling problems in conservation genetics using captive *Drosophila* populations: improvement of reproductive fitness

due to immigration of one individual into small partially inbred populations. *Zoo Biology* 11, 343-351. doi: 10.1002/zoo.1430110506

Thiele, J., Hansen, T., Siegismund, H. R., & Hauser, T. P. (2010). Genetic variation of inbreeding depression among floral and fitness traits in Silene nutans. *Heredity, 104*(1), 52-60. doi:10.1038/hdy.2009.103

Vange, V. (2002). Breeding system and inbreeding depression in the clonal plant species Knautia arvensis (Dipsacaceae): implications for survival in abandoned grassland. *Biological Conservation, 108*(1), 59-67. doi:10.1016/s0006-3207(02)00090-3

Vitousek, P. M., Mooney, H. A., Lubchenco, J., & Melillo, J. M. (1997). Human domination of Earth's ecosystems. *Science, 277*(5325), 494-499. doi:10.1126/science.277.5325.494

Wallace, B. C., Trikalinos, T. A., Lau, J., Brodley, C., & Schmid, C. H. (2010). Semi-automated screening of biomedical citations for systematic reviews. *Bmc Bioinformatics, 11*. doi:10.1186/1471-2105-11-55

Westgate, M. J., & Lindenmayer, D. B. (2017). The difficulties of systematic reviews. *Conservation Biology, 31*(5), 1002-1007. doi:10.1111/cobi.12890

Whitlock, R., Stewart, G. B., Goodman, S. J., Piertney, S. B., Butlin, R. K., Pullin, A. S., & Burke, T. (2013). A systematic review of phenotypic responses to between-population outbreeding. *Environmental Evidence, 2*(1), 13. doi:10.1186/2047-2382-2-13

Williams, G. C. (1957). Pleiotropy, natural-selection, and the evolution of senescence. *Evolution, 11*(4), 398-411. doi:10.1111/j.1558-5646.1957.tb02911.x

Wright, S. (1931). Evolution in Mendelian populations. Genetics 16:97-259. doi: 10.1093/genetics/16.2.97

Wolf, M., van Doorn, G. S., Leimar, O., & Weissing, F. J. (2007). Life-history trade-offs favour the evolution of animal personalities. *Nature, 447*(7144), 581-584. doi:10.1038/nature05835

Appendix

Appendix 1

Reference papers used to identify inbreeding-related search strings by Neaves and her colleagues.

Year	Authors	Journal/ Publisher	Title
2002	Keller LF, Waller DM.	Trends in Ecology & Evolution	Inbreeding effects in wild populations.
2004	Spielman D, Brook BW, Frankham R.	PNAS	Most species are not driven to extinction before genetic factors impact them.
2010	Frankham R, Ballou JD, Briscoe DA.	Cambridge: Cambridge University Press	Introduction to Conservation Genetics.
2009	Charlesworth D, Willis JH.	Nature Reviews Genetics	The genetics of inbreeding depression.
1999	Charlesworth B, Charlesworth D.	Genetics Research	The genetic basis of inbreeding depression.
2011	Hedrick PW, Adams JR, Vucetich JA.	Conservation Biology	Reevaluating and broadening the definition of genetic rescue.
2010	Hedrick P, Fredrickson R.	Conservation Genetics	Genetic rescue guidelines with examples from Mexican wolves and Florida panthers.
2010	Johnson WE, Onorato DP, Roelke ME, Land ED, Cunningham M, Belden RC,	Science	Genetic restoration of the Florida panther.
2004	Madsen T, Ujvari B, Olsson M.	Biological Conservation	Novel genes continue to enhance population growth in adders (*Vipera berus*).
1999	Madsen T, Shine R, Olsson M, Wittzell H.	Nature	Conservation Biology: restoration of an inbred adder population.
2007	Willi Y, van Kleunen M, Dietrich S, Fischer M.	Proceeding of Royal Society B Biological Sciences.	Genetic rescue persists beyond first-generation outbreeding in small populations of a rare plant.
2007	Edmands S.	Molecular Ecology	Between a rock and a hard place: evaluating the relative risks of inbreeding and

			outbreeding for conservation and management.
1991	Lynch M.	Evolution	The genetic interpretation of inbreeding depression and outbreeding depression.
2013	Whitlock R, Stewart GB, Goodman SJ, Piertney SB, Butlin RK, Pullin A.	Environmental Evidence	A systematic review of phenotypic responses to between-population outbreeding.
2011	Frankham R, Ballou JD, Eldridge MDB, Lacy RC, Ralls K, Dudash MR.	Conservation Biology	Predicting the probability of outbreeding depression.
2000	Hedrick PW, Kalinowski ST.	Annual Review of Ecology, Evolution, and Systematics	Inbreeding depression in conservation biology.
1999	Crnokrak P, Roff DA.	Heredity	Inbreeding depression in the wild.

Appendix 2

Description of sources of heterogeneity in the meta-analysis (Chapter 3 and Chapter 4), reasons for the choice of heterogeneity and methods for the relevant assessment.

Taxonomic group

Domain-level taxonomic classification was record for studied population as either animal or plant. Animals were further classified as phylum-level as invertebrate and vertebrate. Taxonomic group was recorded in order to determine whether heterogeneity of inbreeding depression exists among different species. The difference of inbreeding depression occurs due to taxonomic-correlated variation in mating system, life-history and behaviour that affect individuals to breed and how deleterious alleles are maintained or removed by natural selection. Understanding the association between taxonomic group and inbreeding depression is essential for conservation practices since it allows conservation biologists to assess the risks of endangered species under the threats of inbreeding depression.

Trait type

I classified phenotypic traits where inbreeding depression was measured in relevant articles. One of behaviour, defence, development, fecundity, growth rate, physiology, size, survival or viability. Majority of behaviour traits were related to mating behaviour such as mating efforts or copulation duration. Defence traits were measured as organism's ability to defend pathogen, predator, herbivore or extreme manipulated environmental conditions, such as heat shock. Developmental traits included the time at reaching developmental stages, such as age when flowered in plants or developmental disorder, such as fluctuating asymmetry. Fecundity traits included measures of reproductive output; measures of number or probability of reproductive organisms produced or measures of mating success. Fitness were fitness metrics that integrated multiple phenotypic traits into one value. These fitness metrics contained at least one of fitness component traits. Growth rate was the traits that measured the increase in body size within a specific timeframe. Physiological traits mainly had two categories. One was the measures of body composition, such as protein, fat, nitrogen, carotenoid, chlorophyll content. The other was the measures of physiological functions, such as photosynthetic rate or water use efficiency. Size traits were the measures of organisms' total body or parts of body's mass, volume or length. Survival traits were the measures of survival proportions after morphological maturity or establishment or the length of lifespan. Viability traits were the survival traits before morphological maturity or entablement or the measures of germination or hatching rates. Other traits were the other non-fitness component traits that did not belong to the categories described above. Studying the association between trait types and inbreeding depression is essential since the effects of inbreeding depression in different traits influence population demography in various ways via fitness reductions.

Fitness class

The trait types described above were further categorised as fitness component traits or non-fitness component traits. Fitness component traits included fecundity, fitness, survival or viability. The remaining trait types were less directly linked to fitness and therefore categorised as non-fitness component traits. Understanding the relationship between fitness class and inbreeding depression is important due to these two classes

of traits influence population demography in different ways. Additionally, the genetic constitutions of fitness component traits could differ from non-fitness component traits, consequently making them have different degree of vulnerability to inbreeding depression (Merila & Sheldon, 1999).

Life history stage

I classified traits according to which life history stages they were measured. Life history stage was categorised as early, middle and late stage. Early life history stage was defined when traits were measured before morphological maturity or establishment, such as germination, egg hatching rates. Trait type "viability" was also in this category. Late life history stage was defined when traits were measured after sexual maturity. Trait type "fecundity", such as reproduction output, flowering proportion was in this category. Any trait measured between early and late life history stages was classified as middle life history stage. Understanding the relationship between life history stage and inbreeding depression is critical since the efficiency of purging differs among life history stages. Therefore, the magnitude of inbreeding depression might vary among life history stages.

Types of population

I categorised whether populations are natural populations or not based on the extent of human intervention. Populations obtained from the wild or populations stayed in the human-maintained environments less than 5 generations are classified as natural populations. The reason to separate these two populations is that human-maintain populations might experience different selection regimes and harbour different genetic compositions in comparison to natural populations. In addition, populations maintained under artificial environments could suffer population bottleneck, making them already inbred or have less genetic diversity. Therefore, they could be more vulnerable to inbreeding depression.

Mating system

Two definitions of mating system were included in this review. In the first definition (*mating system 1*), animals were classified as either polygamy and monogamy. Alternatively, plants were classified based on outcrossing rates measured from molecular markers at population-level. Outcrossing rate less than 0.2 was defined as selfing whereas outcrossing rate larger than 0.8 was defined as outcrossing (Winn *et al*., 2011). Any population with outcrossing rate between 0.2 and 0.8 was defined as mix-mating population. In the second definition (*mating system 2*), the classification was based on study taxa. Taxa was classified as mix-mating if their offspring were produced either from selfing or outcrossing, such as plants or gastropods. Instead, taxa with separate sexes or with self-incompatibility mechanism were classified as outcrossing. When the data was not provided in the papers, literature search was conducted to retrieve the information of these mating systems. I investigated the association between inbreeding depression and mating systems for the following reasons. To begin with, mating system affects the magnitude of standing genetic variation and heterozygosity, and therefore influences the susceptibility to inbreeding depression. In addition, mating system also influences genetic architecture of populations and therefore shaping the pattern of natural selection. This will in turn affect the efficiency of purging (the removal of deleterious alleles), subsequently modified the severity of inbreeding depression.

Life expectancy

Life expectancy was recorded for animals and plants separately. Life expectancy of animals was recorded in the units of days. Alternatively, life expectancy of plants was classified as annul, biannual or perennial. When data were not provided in the articles, literature search was conducted to retrieve the relevant information. In some reviews, life expectancy was reported to be associated with mating system and therefore influences the magnitude of inbreeding depression. For example, annual plants are more likely to have mix-mating systems while perennial plants are mostly outcorsser (Scofield & Schultz, 2006; Winn *et al*., 2011).

Observation environment

Observation environment was recoded for each trait where it was measured. Observation environment was classified as laboratory, greenhouse, common garden and field. Different observation environment might indicate different magnitude of stress intensity. It well acknowledged that laboratory represents benign environment in comparison to field environments. As a result, the level of inbreeding depression might differ depending on which environment it measured.

Appendix 3

Details of all relevant articles used in the meta-analysis and the corresponding studies for each article. Article ID is the arbitrary number assigned for each article in the database; No.env.studied is the numbers of environmental treatments applied in each article except control treatments. No.effect sizes is the numbers of the effect sizes contributed by each article.

Article ID	Species	Taxon	No. env. studied	No. effect sizes
68	*S. vulgaris*	Plant	6	28
366	*A. lyrate*	Plant	3	24
399	*P. astrigera*	Invertebrate	2	2
533	*S. vulgaris*	Plant	7	11
737	*N. vespilloides*	Invertebrate	2	4
857	*G. holbrooki*	Vertebrate	1	1
1202	*S. canaria*	Vertebrate	2	2
1445	*N. femorata*	Invertebrate	1	2
1511	*S. vulgaris*	Plant	8	39
1514	*D. plexippus*	Invertebrate	1	1
1517	*M. luteus*	Plant	1	4
	M. cupreus	Plant	1	4
1624	*N. vespilloides*	Insect	2	8
1703	*D. melanogaster*	Invertebrate	8	16
2133	*I. purpurea*	Plant	2	58
2262	*P. quinata*	Plant	2	8
2326	*D. magna*	Invertebrate	1	12
2363	*N. vespilloides*	Invertebrate	1	4
2551	*D. rerio*	Vertebrate	4	8
2804	*A. asclepiadis*	Invertebrate	1	11

2835	*D. melanogaster*	Invertebrate	20	20
2916	*B. nigra*	Plant	1	6
2940	*T. castaneum*	Invertebrate	1	1
3243	*S. carolinense*	Plant	2	8
3333	*P. acuta*	Invertebrate	2	24
3357	*D. melanogaster*	Invertebrate	1	2
3440	*V. hirundinaria*	Plant	1	1
	A. asclepiadis	Invertebrate	1	4
3520	*D. teres*	Plant	2	2
3742	*P. acuta*	Invertebrate	2	10
3802	*C. maculatus*	Invertebrate	1	2
3836	*B. globosus*	Invertebrate	2	16
4157	*A. petiolata*	Plant	1	4
4174	*B. anynana*	Invertebrate	2	36
4207	*S. carolinense*	Plant	2	10
4512	*E. wildpretii*	Plant	1	1
4578	*D. melanogaster*	Invertebrate	1	15
4649	*S granulata*	Plant	2	4
4651	*C. tectorum*	Plant	1	2
4925	*S. carolinense*	Plant	1	7
5047	*M. guttatus*	Plant	3	48
5241	*S. carolinense*	Plant	1	2
5965	*C. maculatus*	Invertebrate	3	36
5990	*D. melanogaster*	Invertebrate	3	9
6146	*D. melanogaster*	Invertebrate	1	8
6224	*P. acuta*	Invertebrate	2	12
6334	*M. ringens*	Plant	2	16
6449	*P. acuta*	Invertebrate	1	1
6602	*F. virginiana*	Plant	2	2
7053	*C. maculatus*	Invertebrate	3	12
7399	*M. musculus*	Vertebrate	1	2
7492	*C. parviflora*	Plant	1	40
7642	*D. melanogaster*	Invertebrate	5	15
8172	*C. riparius*	Invertebrate	4	24
8239	*I. capensis*	Plant	3	3
8268	*P. major*	Vertebrate	1	2
8789	*C. convexa*	Invertebrate	3	15
	C. fornicata	Invertebrate	3	9
9202	*D. melanogaster*	Invertebrate	2	2
9204	*H. trionum*	Plant	2	7
9290	*L. stagnalis*	Invertebrate	2	4
9564	*M. guttatus*	Plant	1	9

9826	*I. hederacea*	Plant	1	3
9947	*V. pumila*	Plant	1	2
	V. stagnina	Plant	1	2
10004	*C. pepo*	Plant	1	6
10122	*C. pepo*	Plant	1	12
10168	*D. rerio*	Vertebrate	1	1
10276	*C. pepo*	Plant	1	36
10482	*M. guttatus*	Plant	1	20
10565	*P. coronopus*	Plant	2	10
10988	*R. robini*	Invertebrate	1	2
11043	*D. buzzatii*	Invertebrate	3	9
11061	*M. guttatus*	Plant	1	4
11063	*T. urticae*	Invertebrate	2	12
11170	*D. stramonium*	Plant	1	2
11519	*M. guttatus*	Plant	1	4
11628	*G. holbrooki*	Vertebrate	1	12
12038	*A. geniculatus*	Invertebrate	1	8
12082	*C. sancta*	Plant	2	6
12098	*D. melanogaster*	Invertebrate	1	2
12210	*C. sancta*	Plant	1	36
12349	*C concinna*	Plant	2	56
13012	*P. coronopus*	Plant	1	3
13105	*P. dubia*	Plant	1	12
13167	*B. grandiflor*	Plant	1	1
13401	*P. menziesii*	Plant	1	4
13634	*M. guttatus*	Plant	1	12
13686	*L. fios-cuculi*	Plant	2	69
14071	*S. lydgatei*	Plant	1	34
14114	*D. buzzatii*	Invertebrate	2	9
14199	*A. caerulea*	Plant	1	4
14300	*D. verticillatus*	Plant	5	16
14346	*T. castaneum*	Insect	1	2
14655	*G. pulchella*	Plant	1	1
16687	*C. allenii*	Plant	1	7
17518	*L. cardinalis*	Plant	1	4
	L. siphilitica	Plant	1	2
17526	*S. columbaria*	Plant	1	12
17529	*H. appendiculatum*	Plant	1	5
17540	*S. angularis*	Plant	2	6
17801	*S. oregana*	Plant	1	4
19897	*C. sancta*	Plant	1	8
20150	*D. magna*	Invertebrate	2	2

20774	*L. viscaria*	Plant	1	24
21635	*Z. angusticollis*	Invertebrate	2	6
22925	*D. melanogaster*	Invertebrate	3	12
22927	*D. bipectinate*	Invertebrate	1	2
	D. birchii	Invertebrate	1	2
	D. bunnanda	Invertebrate	1	2
	D. pseudoananassae	Invertebrate	1	2
	D. serrata	Invertebrate	1	2
	D. hydei	Invertebrate	1	2
	D. replete	Invertebrate	1	2
	D. melanogaster	Invertebrate	1	2
	D. simulans	Invertebrate	1	2
	D. sulfurigaster	Invertebrate	1	2
31399	*D. buzzatii*	Invertebrate	1	9
31485	*H. subalpina*	Plant	1	1
31736	*M. guttatus*	Plant	1	6
31743	*A. arbustorurn*	Invertebrate	1	1
32744	*C. chilensis*	Plant	1	3
32929	*D. magna*	Invertebrate	1	2
33100	*D. melanogaster*	Invertebrate	1	2
33266	*L. bicolor*	Plant	1	3
	L. jepsonii	Plant	1	3
33901	*S. malacitanus*	Plant	1	24
33978	*B. glabrata*	Invertebrate	2	3
34113	*D. melanogaster*	Invertebrate	3	3
34393	*D. melanogaster*	Invertebrate	2	12
34566	*B. tabaci*	Invertebrate	1	6
34794	*S. ciliata*	Plant	2	8
34944	*S. carolinense*	Plant	2	4
35097	*D. stramonium*	Plant	1	1
35650	*G. holbrooki*	Vertebrate	1	15
35896	*S. canaria*	Vertebrate	2	2
C1188	*T. castaneum*	Invertebrate	1	3
C1231	*D. melanogaster*	Invertebrate	1	1
C1242	*G. holbrooki*	Vertebrate	1	2
C1253	*N. vespilloides*	Invertebrate	3	7
C287	*T. holothuriae*	Invertebrate	2	12
C971	*D. melanogaster*	Invertebrate	2	8
E139	*T. castaneum*	Invertebrate	1	2

Reference

Merila, J., & Sheldon, B. C. (1999). Genetic architecture of fitness and nonfitness traits: empirical patterns and development of ideas. *Heredity, 83*, 103-109. doi:10.1046/j.1365-2540.1999.00585.x

Scofield, D. G., & Schultz, S. T. (2006). Mitosis, stature and evolution of plant mating systems: low-Phi and high-Phi plants. *Proceedings of the Royal Society B-Biological Sciences, 273*(1584), 275-282. doi:10.1098/rspb.2005.3304

Winn, A. A., Elle, E., Kalisz, S., Cheptou, P. O., Eckert, C. G., Goodwillie, C., & Vallejo-Marin, M. (2011). Analysis of inbreeding depression in mixed-mating plants provides evidence for selective interference and stable mixed mating. *Evolution, 65*(12), 3339-3359. doi:10.1111/j.1558-5646.2011.01462.x