A systematic map of studies testing the relationship between temperature and animal reproduction


Abstract

1. Exposure to extreme temperatures can negatively affect animal reproduction, by disrupting the ability of individuals to produce any offspring (fertility), or the number of offspring produced by fertile individuals (fecundity). This has important ecological consequences, because reproduction is the ultimate measure of population fitness: a reduction in reproductive output lowers the population growth rate and increases the extinction risk. Despite this importance, there have been no large-scale summaries of the evidence for effect of temperature on reproduction.

2. We provide a systematic map of studies testing the relationship between temperature and animal reproduction. We systematically searched for published studies that statistically test for a direct link between temperature and animal reproduction, in terms of fertility, fecundity or indirect measures of reproductive potential (gamete and gonad traits).

3. Overall, we collated a large and rich evidence base, with 1654 papers that met our inclusion criteria, encompassing 1191 species.

4. The map revealed several important research gaps. Insects made up almost half of the dataset, but reptiles and amphibians were uncommon, as were non-arthropod...
heat stress is taxonomically widespread, being seen for example in corals (Paxton et al., 2016), insects (David et al., 2005), fish (Breckels & Neff, 2013), livestock (De Rensis et al., 2017; Peña et al., 2019) and humans (Hajdu & Hajdu, 2022; Hoang-Thi et al., 2022). Further, the mechanisms leading to reproductive impairment are highly variable, because fertility and fecundity are emergent products of many physiological, developmental and behavioural processes (Walsh et al., 2019). For example, reproductive impairment can be due to disrupted gonad development (Delorme & Sewell, 2016; McBride et al., 1997), reduced sperm function (Breckels & Neff, 2013; Peña et al., 2019; Pérez-Crespo et al., 2008; Vasudeva et al., 2014), reduced fertilisation and pregnancy rates (De Rensis et al., 2017; Hajdu & Hajdu, 2022), or a reduction in the resources that can be invested into gametes or offspring (Dahike et al., 2020).

While both male and female reproduction can be affected by heat stress, sperm production appears to be particularly sensitive (David et al., 2005; Hansen, 2009; Sales et al., 2018; Schou et al., 2021; Walsh et al., 2019). Conversely, cold temperatures can also lead to impaired reproduction (David et al., 2005; Rinehart et al., 2000; Schou et al., 2021). This can arise due to an overall reduction in activity levels and growth associated with cold in ectotherms (e.g. Pörtner et al., 2001) or due to cold acting directly on gamete function or embryogenesis (e.g. Berger et al., 2008; David et al., 2005; Shine, 2005; Watson, 2000).

Fertility loss caused by extreme temperatures has important ecological consequences, because reproduction is the ultimate measure of population fitness: in evolutionary terms, a sterile individual is the same as a dead one. But even small, persistent reductions in reproductive output of fertile individuals can significantly reduce population growth rate, thus reducing population health and increasing extinction risk (Savage et al., 2004). High-temperature invertebrates. Fecundity was the most common reproductive trait examined, and relatively few studies measured fertility. It was uncommon for experimental studies to test exposure of different life stages, exposure to short-term heat or cold shock, exposure to temperature fluctuations, or to independently assess male and female effects. Studies were most often published in journals focusing on entomology and pest control, ecology and evolution, aquaculture and fisheries science, and marine biology. Finally, while individuals were sampled from every continent, there was a strong sampling bias towards mid-latitudes in the Northern Hemisphere, such that the tropics and polar regions are less well sampled.

5. This map reveals a rich literature of studies testing the relationship between temperature and animal reproduction, but also uncovers substantial missing treatment of taxa, traits, and thermal regimes. This database will provide a valuable resource for future quantitative meta-analyses, and direct future studies aiming to fill identified gaps.

**KEYWORDS**
climate change, egg, evidence map, sperm, sterility, systematic review, thermal fertility limit, thermal tolerance

1 | INTRODUCTION

The world is warming faster than at any time in the last 2000 years (IPCC, 2023), and heatwaves are increasing in severity and frequency (Buckley & Huey, 2016; Murali et al., 2023). High temperatures can have a range of negative effects on organisms. Most well-studied is the effect on basic metabolic functions: at high temperatures metabolism breaks down, and organisms are unable to move or feed, quickly leading to death (Bennett et al., 2021; Clarke, 2017). Organisms living in environments they can no longer tolerate must respond by either evolving increased tolerance, or moving somewhere else (Parmesan, 2006). However, the ability to evolve increased tolerance is limited for species inhabiting regions that are already very warm (Hoffmann et al., 2003; Kellermann et al., 2012; Van Heerwaarden & Sgrò, 2021), because biochemistry sets a hard upper limit on metabolic functions at high temperatures (Bennett et al., 2021; Clarke, 2017). Consequently, warming has led to large-scale range shifts as organisms move to higher latitudes or altitudes (Bebber et al., 2013; Chen et al., 2011; Parmesan & Yohe, 2003). For organisms that cannot adapt or move, extinction is the ultimate outcome of climate change (Duffy et al., 2022; Román-Palacios & Wiens, 2020; Van Heerwaarden & Sgrò, 2021).

Measuring survival at different temperatures is a common way to determine thermal tolerance (e.g. Bennett et al., 2021). However, sub-lethal consequences of temperature stress are also important to consider, with perhaps the most critical sub-lethal effect being the impairment of reproduction. For example, exposure to high temperatures can disrupt both the ability of individuals to produce any viable eggs or sperm (fertility), and the number of eggs or offspring produced by fertile individuals (fecundity) (Schou et al., 2021; Walsh et al., 2019). Reproductive impairment due to
fertility loss may also have important economic consequences, as heat stress is well-known to impair the reproduction of livestock (e.g., De Rensis et al., 2017; Peña et al., 2019) and species harvested for food (e.g., Vilchis et al., 2005; Yoneda & Wright, 2005). Importantly, reproductive impairment often occurs at conditions extreme than death or loss of metabolic functions. For example, in Drosophila fruit flies, the upper temperature limit for male fertility may be up to 4°C lower than the upper temperature limit for survival (Parratt et al., 2021; Van Heerwaarden & Sgrò, 2021). In such species, we might expect population fitness in the face of warming to be more strongly limited by reproductive capacity than by survival. This is supported by recent work in Drosophilidae, showing that the temperature at which males become sterile is a better predictor of current species distributions, and extinction risk in laboratory populations, than the lethal temperature (Parratt et al., 2021; Van Heerwaarden & Sgrò, 2021). If this is a general trend then attempts to estimate species vulnerability to temperature extremes using laboratory measurements of thermal limits for survival will be overly optimistic.

However, despite the importance outlined above, we are currently lacking a large-scale synthesis of the effect of temperature on reproduction, outside of a very few well-studied species groups (Dahlke et al., 2020; Kellermann et al., 2012; Parratt et al., 2021; Van Heerwaarden & Sgrò, 2021). This contrasts with a large and taxonomically diverse database of studies of temperature effects on survival (Bennett et al., 2018). To facilitate such a synthesis, our aim in this study was to build a database of published studies that test the relationship between temperature and reproduction in animals. A systematic map is a structured overview of the evidence base for a research question (James et al., 2016). It collates, describes and catalogues the available evidence in a systematic way and can be used for future quantitative analysis or to identify research gaps (James et al., 2016). To create the map, we first systematically searched for published studies that statistically test for a relationship between temperature and animal reproduction, in terms of three main types of reproductive trait: fertility (the ability to produce any offspring), fecundity (the number of offspring produced) and indirect measures of reproductive potential (gamete and gonad traits). Overall, we discovered a large and rich evidence base, with 1654 papers (on 1191 species) that met our inclusion criteria. We used this dataset to ask six key questions regarding studies that test the relationship between temperature and animal reproduction:

1. What is the taxonomic diversity of the species studied?
2. What are the biological or ecological characteristics of these study species? For example, are they predominantly terrestrial or aquatic? Endotherms or ectotherms?
3. Which reproductive traits are most often measured?
4. What experimental methods do studies use?
5. Where in the world are animals sampled from?
6. In which types of journals are studies published?

2 | MATERIALS AND METHODS

Throughout we follow the ROSES Reporting standards for Systematic Evidence Syntheses guidelines developed by Haddaway et al. (2018).

2.1 | Search methods

We performed online literature searches as part of a broader project investigating how animal reproduction is likely to be affected by five abiotic factors currently experiencing human-induced changes: temperature, radiation, humidity, pH, and salinity. However, any studies obtained from these searches that did not record temperature were not considered in our systematic map (see further details below). All searches were performed using the ISI Web of Science Core Collection. We did not search for unpublished data or grey literature. Initial searches were performed between December 2020 and February 2021. Final searches were performed on the 24 August 2021. Each search had the same set of reproduction-related terms, but varied in abiotic factor-related terms. We considered all available years. The full search strings are presented in Table 1. In total, the five searches resulted in 25,051 unique hits (after duplicates were removed). Given the large number of hits, we did not perform additional forward or backwards searching.

2.2 | Study inclusion criteria

To be considered eligible for inclusion in the systematic map, a study had to fulfil the following criteria:

1. Be a peer-reviewed (and not retracted) scientific article or book chapter presenting new data.
2. Be conducted on any animal species, except for humans.
3. Measure at least one of the following reproductive traits:
   a. Fertility (the ability to produce offspring).
   b. Fecundity (number of eggs or offspring produced).
   c. Other reproductive traits (conception rate, proportion of individuals breeding or reproducing, male paternity or sperm transfer).
   d. Number of matings.
   e. Gamete traits (sperm or ova number, size, performance, fertilisation ability).
   f. Gonad traits (testes or ovary size, morphology, developmental stage, function; gametogenesis).
4. Report one of the above traits for at least two different temperatures.
5. Record variation within a species (i.e. multiple individuals are measured for each species).

We excluded review papers, meta-analyses (unless there was a clear reason to think they present previously-unpublished data),
mathematical models and opinion articles. We considered studies published in any language, however restricting our search terms to English meant that only 46 non-English language articles were found by our searches, and all had an English abstract or title (14 Portuguese, 12 French, 6 German, 5 Japanese, 4 Spanish 2 Russian, 1 Hungarian, 1 Polish, 1 Turkish).

We excluded comparative studies in which only species-average values were analysed. We considered both experimental studies which exposed subjects to controlled environmental manipulations, and observational studies in which subjects were exposed to natural variations in temperature. We excluded studies for which temperature was only indirectly linked to reproduction, for example via seasonal changes (i.e. date was a variable rather than temperature). We did not limit our search to a specific temperature range, and so considered studies exposing subjects to controlled environmental manipulations, irrelevant (e.g. clinical trials on humans, engineering or chemistry) were excluded. Studies that passed the initial title screen (5563) were then imported into Rayyan for abstract screening (Ouzzani et al., 2016). Abstracts were screened using the same eligibility criteria as above, with each study being screened by two people independently. In total, 28 people screened during this stage. For 1407 out of 5563 abstracts (25.3%), the two screeners disagreed over the inclusion decision. These abstracts were then screened for a third time by LRD, AB, CF, TARP, or RRS, and a final decision made using the inclusion criteria described above. Articles that passed the abstract screen (2284) were then downloaded in full for the full-text screening phase. Here, each article was read by a single screener, and was assessed one last time against the inclusion criteria. In total, 28 people screened during this stage. This process resulted in 1654 relevant papers being included in the final dataset. The full list of included papers is available at Dougherty et al. (2023).

### Study screening

See Figure 1 for a summary of the literature search and study screening process. We screened studies found by all searches listed in Table 1. However, only studies that included a measurement of reproductive traits under at least two different temperatures were considered for inclusion (effects of other abiotic factors on reproduction were not recorded).

Studies were screened in three progressively-detailed stages. First, titles were screened by LRD, and studies that were obviously irrelevant (e.g. clinical trials on humans, engineering or chemistry) were excluded. Studies that passed the initial title screen (5563) were then imported into Rayyan for abstract screening (Ouzzani et al., 2016). Abstracts were screened using the same eligibility criteria as above, with each study being screened by two people independently. In total, 28 people screened during this stage. For 1407 out of 5563 abstracts (25.3%), the two screeners disagreed over the inclusion decision. These abstracts were then screened for a third time by LRD, AB, CF, TARP, or RRS, and a final decision made using the inclusion criteria described above. Articles that passed the abstract screen (2284) were then downloaded in full for the full-text screening phase. Here, each article was read by a single screener, and was assessed one last time against the inclusion criteria. In total, 28 people screened during this stage. This process resulted in 1654 relevant papers being included in the final dataset. The full list of included papers is available at Dougherty et al. (2023).
2.4 Data extraction and coding

For all relevant articles, the same 28 screeners also collected detailed qualitative information about the subject species and the study methodology. All screeners used the same data template. For species-level information, we primarily relied on information provided in the papers themselves. In other cases, the screeners searched for information online; however, these sources were not recorded and so we are unable to provide a definitive list here.

For each paper, we recorded the following qualitative information:

1. Bibliometric information (author names, journal title, publication year, doi). We classified journals into one of 20 categories based on title.
2. Species taxonomic information (taxonomic family, class, phylum).
3. Reproductive traits. We sorted reproductive traits into five categories:
   a. Fertility (conception/pregnancy, proportion of reproductive or breeding individuals, sperm transfer or paternity success).
   b. Fecundity (egg or offspring number). NB: unlaid eggs (measured for example by dissecting the ovaries) were included in the gamete traits category (point d below).
   c. Number of matings.
   d. Gamete traits: ova number, ova size, ova quality (viability, longevity or composition), sperm number, sperm morphology (sperm length, frequency of physical abnormalities), sperm performance (swimming speed, proportion of motile sperm, proportion of living sperm, longevity, mitochondrial activity, antioxidant production), fertilisation rate (proportion of ova that are fertilised).
   e. Gonad traits: gonad size (including gonadosomatic index), gonad morphology (size, number, or shape of specific components of the gonads, tissue damage, physical abnormalities), gonad developmental stage, gamete developmental stage, gametogenesis (the production of any gametes).
4. Habitat (species). Whether animals are primarily aquatic or terrestrial. Amphibious species that live both on land and in water, or switch environments during development, were classed as ‘both’. We further characterised aquatic species as living in

FIGURE 1 PRISMA flowchart showing the literature search and study screening process.
freshwater, saltwater or both (for example anadromous salmon spend time in both salt and freshwater).

5. Mode of thermoregulation (species). We classified animals as endothermic (maintains body temperature at a metabolically favourable temperature) or ectothermic (body temperature depends on environmental heat).

6. Reproductive mode (species). Do animals reproduce sexually, asexually (via parthenogenesis) or as hermaphrodites?

7. Fertilisation mode (for sexual species and hermaphrodites only). Does fertilisation occur internally or externally?

8. Degree of mobility (species). Are animals sessile or motile as adults?

9. Is the species of economic importance? We sorted species into one of six categories: farmed (for food or other products), harvested, pests of crops or stored food, biological control agents, pollinators or human disease vectors or parasites.

10. Sex exposed (for sexual species only). Were individuals of both sexes, only males or only females exposed to temperature variation prior to reproductive measurements?

11. Study setting. Was the study performed in a lab setting (here the environment is controlled), a semi-natural setting such as a farm or other type of outdoor enclosure (here the environment is partially controlled, but animals are still exposed to some natural variation in temperature), or a field setting (here the environment is not controlled, and animals are fully exposed to natural variation in temperature)?

12. Stressor duration (for experimental studies only). How long were subjects exposed to variable temperatures? We sorted studies into three categories: less than 24h, 1–5 days, or more than 5 days.

13. Life stage (for experimental studies only). During which life stage did subjects experience variation in temperature? We sorted studies into four categories: gamete (spERM or egg), juvenile (pre sexual maturity), adult (post-sexual maturity), or mixed (if temperature treatments spanned two or more life stages).

14. Temperature variation (for experimental studies only). Were subjects exposed to constant temperatures, fluctuating temperatures or both?

15. Cold (for experimental studies only). Were any subjects in the study exposed to temperatures below an arbitrary cut-off of 10°C?

16. Survival or lifespan. Did the study also record the lifespan or survival of individuals at more than one temperature? Individuals could be the same or different from those whose reproduction was measured.

17. For wild or wild-caught individuals, we recorded the country animals were sampled from. For marine species, we took the country with the closest coastline (a few oceanic species were excluded from this analysis).

Notably, we did not extract quantitative information on the range of temperatures tested in each study, for three reasons. First, because this information is not informative without some knowledge of the temperature range that each species typically experiences or tolerates. Second, for experimental studies this would require detailed extraction of treatment data, which we did not do for practical reasons. Third, a minority of studies were purely observational, recording reproduction of wild animals. Here, the temperature environment experienced by each subject is not known—typically only the average weekly or monthly temperature is reported in the study.

2.5 | Data synthesis and presentation

All analyses and visualisations were performed in the R environment (R version 4.1.2; R Development Core Team, 2021) by LRD and FF. We used the Open Tree of Life database (Hinchliff et al., 2015; https://tree.opentreeoflife.org/), plus the R packages rotl v3.0.12 (Michonneau et al., 2016) and ape v5.6-2 (Paradis et al., 2004) to construct a phylogenetic tree for the species in the data table. For data visualisation we used Mapproj v1.2.11, ggtree v3.2.1 (Yu et al., 2017), and ggplot2 v3.3.6 (Wickham, 2011). We obtained data on the average latitude of each country from (https://github.com/albertyw/avenews). All data are available at Dougherty et al. (2023).

3 | RESULTS

3.1 | Species

The final dataset contains data from 1191 species from 12 animal phyla (Figure 2; Figure S1). 65% of species were arthropods (777 species), and 20% were chordates (244 species). The remaining 15% of species were from Annelida, Cnidaria, Ctenophora, Echinodermata, Mollusca, Nematomorpha, Platyhelminthes, Porifera, Rotifera and Tardigrada (Figure 2). Insects made up 47.5% of the dataset, with the next most common clades being Crustacea (10%), Actinopterygii (8.3%), Arachnida (6.8%), Mollusca (6.2%) and Aves (5.4%) (Figure 2). Twenty-four species were represented in the dataset by seven or more studies (Figure 3). This group includes common model organisms such as the fruit fly Drosophila melanogaster (45 studies), the house mouse Mus musculus (24 studies), the flour beetle Tribolium castaneum (9 studies) and the brown rat Rattus norvegicus (7 studies); domestic chickens Gallus domesticus (21 studies), pigs Sus scrofa domesticus (19 studies), sheep Ovis aries (18 studies) and cattle Bos taurus (13 studies); and economically important species such as the red spider mite Tetanychus urticae (10 studies), the diamondback moth Plutella xylostella (8 studies) and the yellow fever mosquito Aedes aegypti (7 studies).

Species categorisations are summarised in Figure 4. Almost two-thirds (63.9%) of the 1191 species in the dataset live predominantly on land. Of the 429 aquatic species included in the dataset, 42% inhabit freshwater, 54% inhabit saltwater and 2.1% (9 species of fish) inhabit both fresh and saltwater. Endothermic
species (birds and mammals) made up only 7.8% of the total species in the dataset. Similarly, most species in the dataset exhibit internal fertilisation (75%), reproduce sexually (88%), and are motile as adults (96%). Only 7.6% of species reproduce asexually, and 4.3% are hermaphrodites.

We classified species into one of seven categories related to their economic importance. Of 1191 species, 48% had some link to human health or livelihoods. 16% of species are classed as biological control agents, 2.8% are reared for food, 2.3% are harvested from the wild for food, and 0.4% are important pollinators of crops (Figure 5). 24% of species are pests of crops or stored food products, and 1% are human disease vectors (Figure 5).

3.2 | Reproductive traits

The trait data are summarised in Figure 6. Fecundity (egg or offspring number) was the most common trait, measured in 72% of 1654 studies. Also, 36% of studies reported temperature effects on gamete traits and 17% on gonad traits. We further divided gamete traits into one of seven categories, of which the most common were sperm performance (165 studies), ova size (148 studies), sperm number (112 studies) and fertilisation rate (80 studies; Figure 6b). We further divided gonad traits into one of five categories, of which the most common was gonad size (112 studies; Figure 6c).
### Study characteristics

The number of studies published each year shows a sharp increase after 1990 (Figure 7). However, we found no evidence that studies examining the link between temperature and animal reproduction are becoming relatively more likely to be published—the temporal trend for this sample matches the trend for the total number of articles indexed in Scopus over the same duration (Bornmann et al., 2021). In 2020, the last year for which we have full data, 100 relevant studies were published. The earliest study we included was from 1924 (Mavor & Svenson, 1924). We found only 34 relevant studies published before 1970.

Most (77% of 1654) studies measured reproduction in a lab setting, with temperature usually being controlled. In contrast, 9.6% of studies sampled wild individuals exposed fully to natural variation in temperature, and 11% examined captive individuals exposed to some natural environmental fluctuation, including farm settings (8.8%) and semi-natural enclosures (2.2%). Notably, 50% of studies also recorded survival or lifespan at more than one temperature. Of the 1454 studies performed on sexual species, over two-thirds (67.4%) exposed both males and females to different temperatures at the same time, 18% exposed only females and 15% exposed only males (Figure 8a). The methodological characteristics of experimental studies are summarised in Figure 8b.

**FIGURE 3** The number of studies providing information for each of the top 24 most commonly studied species in the map dataset.

**FIGURE 4** Stacked bar plots summarising species’ ecology and biology. The number of species (out of a total of 1191) for each category is shown for five descriptors: habitat, mode of thermoregulation, mode of reproduction, mode of fertilisation, and degree of motility. Note the bar is lower for the habitat category because nine aquatic species inhabit both fresh and saltwater during their lives.

<table>
<thead>
<tr>
<th>Species</th>
<th>Number of Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vinegar fly Drosophila melanogaster</td>
<td>33</td>
</tr>
<tr>
<td>House mouse Mus musculus</td>
<td>26</td>
</tr>
<tr>
<td>Chicken Gallus gallus</td>
<td>19</td>
</tr>
<tr>
<td>Pig Sus scrofa</td>
<td>10</td>
</tr>
<tr>
<td>Sheep Ovis aries</td>
<td>10</td>
</tr>
<tr>
<td>Spotted-wing Drosophila Drosophila suzukii</td>
<td>9</td>
</tr>
<tr>
<td>Cattle Bos taurus</td>
<td>7</td>
</tr>
<tr>
<td>European rabbit Oryctolagus cuniculus</td>
<td>7</td>
</tr>
<tr>
<td>Red spider mite Tetanychus urticae</td>
<td>6</td>
</tr>
<tr>
<td>Red flour beetle Tribolium castaneum</td>
<td>5</td>
</tr>
<tr>
<td>Atlantic salmon Salmo salar</td>
<td>5</td>
</tr>
<tr>
<td>Diamondback moth Plutella xylostella</td>
<td>5</td>
</tr>
<tr>
<td>Mediterranean mussel Mytilus galloprovincialis</td>
<td>5</td>
</tr>
<tr>
<td>Cowpea seed beetle Callosobruchus maculatus</td>
<td>5</td>
</tr>
<tr>
<td>Squinting bush brown Bicyclus anynana</td>
<td>4</td>
</tr>
<tr>
<td>Yellow dung fly Scathophaga stercoraria</td>
<td>4</td>
</tr>
<tr>
<td>Bird cherry-oat aphid Rhopalosiphum padi</td>
<td>4</td>
</tr>
<tr>
<td>Brown rat Rattus norvegicus</td>
<td>4</td>
</tr>
<tr>
<td>Great tit Parus major</td>
<td>4</td>
</tr>
<tr>
<td>Water flea Daphnia magna</td>
<td>4</td>
</tr>
<tr>
<td>Codling moth Cydia pomonella</td>
<td>4</td>
</tr>
<tr>
<td>Goat Capra hircus</td>
<td>4</td>
</tr>
<tr>
<td>Honey bee Apis mellifera</td>
<td>4</td>
</tr>
<tr>
<td>Yellow fever mosquito Aedes aegypti</td>
<td>4</td>
</tr>
</tbody>
</table>
FIGURE 5 The number of studies in the map dataset containing at least one species which is economically important to humans. We considered six categories: farmed (is reared for food or another animal product), food (is harvested from the wild for food), pollinator (is involved in the pollination of crops), vector (transmits at least one human disease), pest (feeds on crops or stored food) and control agent (predates on pests of crops or stored food).

FIGURE 6 The number of studies in the dataset that measure at least one trait from: (a) each of five reproductive trait categories, (b) one of seven gamete trait categories, and (c) one of five gonad trait categories.
Of the 1425 studies that experimentally-manipulated temperature, 75.3% exposed subjects to temperature variation for more than five days, 5.6% exposed subjects for between one and five days, and 12% exposed subjects for less than 24 h. Animals were mostly exposed either as adults or across several life stages. In addition, 7.2% of experimental studies exposed only sperm or eggs, and 5.9% exposed only juveniles. Only 19% of 1425 experimental studies exposed subjects to temperatures below 10°C. Finally, 82.7% of experimental studies kept temperatures constant during the experiment.

We found that 34.9% of studies used animals not born in captivity. Individuals in these studies were sampled from 78 countries, across all seven continents (Figure 9a). The USA was the most well-sampled country by far, with 114 studies, followed by Australia, China and Japan, with 49, 37 and 36 studies, respectively. Europe, North America, and Asia were the most sampled continents, with 222, 166 and 135 studies, respectively. Only 56 studies were carried out in South America, 39 in Africa, 63 in Oceania and 5 in Antarctica. This results in a sampling bias towards mid-latitudes in the Northern Hemisphere (Figure 9b).
Studies were published in 416 journals. We classified journals into one of 20 categories (Table 2). Entomological journals were the most well represented in the dataset, both in terms of the number of journals and the number of articles (Table 2). Other common subject areas included aquaculture and fisheries, ecology and evolution, marine biology and pest control (Table 2). Only 37 journals were represented by more than 10 papers. The top 10 most well-represented journals included the specialist entomology journals *Environmental Entomology*, the *Journal of Economic Entomology*, *Entomologia Experimentalis et Applicata*, *Applied Entomology and Zoology* and *The Annals of the Entomological Society of America* (Figure 10).

### TABLE 2 The number of journals and papers in the map dataset assigned to one of 20 journal categories.

<table>
<thead>
<tr>
<th>Journal category</th>
<th>No. journals</th>
<th>No. papers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entomology</td>
<td>47</td>
<td>442</td>
</tr>
<tr>
<td>Ecology and evolution</td>
<td>36</td>
<td>155</td>
</tr>
<tr>
<td>Pest control</td>
<td>36</td>
<td>144</td>
</tr>
<tr>
<td>Aquaculture and fisheries</td>
<td>34</td>
<td>123</td>
</tr>
<tr>
<td>Marine biology</td>
<td>22</td>
<td>122</td>
</tr>
<tr>
<td>Taxon-specific</td>
<td>25</td>
<td>95</td>
</tr>
<tr>
<td>General biology</td>
<td>34</td>
<td>83</td>
</tr>
<tr>
<td>All topics</td>
<td>11</td>
<td>69</td>
</tr>
<tr>
<td>Zoology</td>
<td>29</td>
<td>68</td>
</tr>
<tr>
<td>Physiology</td>
<td>14</td>
<td>56</td>
</tr>
<tr>
<td>Reproduction</td>
<td>17</td>
<td>51</td>
</tr>
<tr>
<td>Other</td>
<td>28</td>
<td>42</td>
</tr>
<tr>
<td>Thermal biology</td>
<td>1</td>
<td>36</td>
</tr>
<tr>
<td>Veterinary</td>
<td>15</td>
<td>36</td>
</tr>
<tr>
<td>Parasites and disease</td>
<td>20</td>
<td>32</td>
</tr>
<tr>
<td>Ornithology</td>
<td>12</td>
<td>28</td>
</tr>
<tr>
<td>Animal breeding</td>
<td>10</td>
<td>26</td>
</tr>
<tr>
<td>Behaviour</td>
<td>10</td>
<td>26</td>
</tr>
<tr>
<td>Genetics</td>
<td>15</td>
<td>19</td>
</tr>
</tbody>
</table>

Note: Categories are listed in descending order of the number of papers.

### DISCUSSION

Our systematic literature searches revealed that the number of studies testing the relationship between temperature and reproduction in animals is very large—we have identified over 1600 relevant articles. There has been a sharp increase in the number of articles published each year since 1990, with no sign of this trend slowing down. It thus seems likely that at least 100 relevant studies will be published per year for the foreseeable future. While exciting, this scale brings its own practical challenges. Notably, we found studies in often specialised journals that covered a broad range of research fields, and synthesis of this question may be hindered if the communication between research fields is poor.

Arthropods are the most well-studied taxonomic group in the dataset, probably because of the ease with which they can be reared in temperature-controlled lab environments, and their significance as pests of crops and stored foods. We obtained data for only 241 vertebrate species, with fish and birds the most well-represented. Amphibians and reptiles are surprisingly uncommon, given their prevalence in studies of temperature effects on survival and other
FIGURE 10 The number of included studies for each of the top 20 most common journals in the map dataset.

phenotypic traits (e.g. Bennett et al., 2018; Pottier, Lin, et al., 2022; While et al., 2018). Mammals are also less well-studied in terms of species numbers, except for livestock species. In terms of applied relevance, many studies focus on species of economic importance. For example, many species are either important food pests (almost 300 species), or their biological control agents (almost 200 species). This dataset is therefore an excellent resource for quantifying the effects of climate change on the reproduction of agricultural pests or our ability to control them. Most major livestock species are also present in the dataset. However, our searches found few studies examining reproduction in species that are vectors of human diseases.

In broad terms, the dataset consists predominantly of studies of ectothermic species, which live on land, reproduce sexually, have internal fertilisation, and are motile as adults. This is not to suggest there is any research bias here- to the best of our knowledge, these traits are representative of most animal species (Angilletta et al., 2010; De Meeûs et al., 2007; Kahrl et al., 2021; May, 1994). We hope this dataset can be used to test for biological or ecological differences between species that moderate the relationship between temperature and reproduction. There may also be important biological or ecological features, which we have not recorded. For example, seed beetle reproduction is more temperature-sensitive in populations that have evolved under strong sexual selection, suggesting that the mating system plays an important role (Baur et al., 2022). In reptiles, the evolution of viviparity (live-bearing of offspring) has been suggested as an adaptation to increase the hatching success of offspring in cold climates (Shine, 2005; Zimin et al., 2022). It would be interesting to compare the relationship between temperature and reproduction in viviparous vs oviparous species, or between species with different degrees of egg incubation (Webb, 1987) or types of egg laying behaviour.

The most common reproductive trait examined in the dataset is fecundity (egg or offspring production), followed by gamete traits. Notably, relatively few studies reported fertility data, in terms of the ability to produce any offspring. The reason for the bias towards fecundity over fertility is not clear, and cannot be attributed to our search methods: both fecundity and fertility were included in our search strings (Table 1). This could be explained if fecundity loss occurs at less extreme temperatures than fertility loss, making the former more easily detectable, especially in observational studies in which extreme warm temperatures do not occur. Nevertheless, the fact that sterility carries a very high fitness cost, and may be induced at ecologically-relevant temperatures, suggests that the thermal tolerance of fertility could play a key role in defining current species distributions and extinction risk following warming (e.g. Parratt et al., 2021; Van Heerwaarden & Sgrò, 2021). We therefore reiterate recent calls to quantify the thermal fertility limits in a broad range of animal taxa (Walsh et al., 2019).

There is no standard method for measuring temperature effects on reproduction. Accordingly, studies vary in their methodological choices. Few studies examined reproduction at temperatures below 10°C, probably because many researchers seek to understand high-temperature effects in light of current climate warming. Most studies examined effects of long-term exposure (i.e. temperatures manipulated for more than five days) and often throughout the whole developmental period of the focal species. We found relatively few studies that tested for temperature effects on reproduction following short-term temperature spikes. We
suggest that more studies should examine effects of short-term temperature spikes, given that extreme short-term temperature fluctuations are becoming more common (Buckley & Huey, 2016; Murali et al., 2023). Most studies examined responses to constant experimental temperatures. However, temperatures in the wild are not constant, even over a single day, and so studies incorporating some degree of temperature fluctuation are much more ecologically realistic (e.g. Rodrigues et al., 2022; Van Heerwaarden & Sgrò, 2021, and see Raynal et al., 2022). We also suggest it would be useful for studies to separate juvenile and adult temperature exposure, in order to determine the ecological significance of exposure at different life stages (Pottier, Burke, et al., 2022). It will be especially useful to test for the presence of sensitive periods during development, such as when gonads are first developing (Canal Domenech & Fricke, 2023; Sales et al., 2021). Another important issue is the relevant temperature sensitivity of male versus female reproduction (lossa, 2019). Male reproduction appears to be more sensitive to heat stress than female reproduction in some species groups (e.g. David et al., 2005). However, there has been no large-scale quantification of this effect (but see Baur et al., 2022). Most studies in this dataset exposed both males and females to temperature treatments simultaneously, which, while ecologically realistic, precludes estimates of the relative contribution of each sex to fertility loss. We suggest that future studies should test for thermal effects on each sex separately.

Around one-third of studies in the dataset examined temperature effects in wild individuals. Individuals were sampled from every continent including Antarctica. However, there was a strong sampling bias towards mid-latitudes in the Northern Hemisphere, and specifically in North America, Europe and Asia. This is a common trend seen in global collections of ecological data (e.g. de los Rios et al., 2018; White et al., 2021). While this might reflect differences in publication rates, it could also be an artefact of our search methods, which did not include non-English keywords or search the grey literature (see below). Whatever the reason, this biased coverage potentially limits our ability to predict global change impacts (White et al., 2021). It also means that the tropics and polar regions are under-represented in our sample. This is especially problematic given that these regions are currently experiencing the largest temperature anomalies due to climate change (Buckley & Huey, 2016). It has also been suggested that tropical ectotherms are particularly vulnerable to future warming, because they are close to their thermal maximum already (e.g. Deutsch et al., 2008; Duarte et al., 2012; Huey et al., 2009; but see Bennett et al., 2021; Duffy et al., 2022; Johansson et al., 2020; Kingsolver et al., 2013; Lancaster, 2016; Pinsky et al., 2019).

While our literature searches identified a large number of relevant articles, our searches were limited in several key ways. First, we only used a single search engine (Web of Science). This is potentially problematic given that the overlap between large academic databases can be as low as 50% for some subjects (Mongeon & Paul-Hus, 2016). Second, we did not search the grey literature, either in terms of unpublished research (e.g. dissertations, or preprints), or articles published outside of traditional academic publishing (e.g. governmental reports, policy statements, technical reports). We expect the grey literature to contain some relevant data (e.g. commercial agricultural pest research), but the amount of accessible data is not yet known. Third, our searches did not include non-English keywords, and thus we were mostly limited to articles written in English. In some fields this can be a problem, as a large amount of data is only available from non-English sources (e.g. biodiversity data: Amano et al., 2023; Chowdhury et al., 2022; Zenni et al., 2023). We are not able to assess the impacts of these decisions on our search process.

In summary, our systematic map reveals several key research gaps in the study of temperature effects on animal reproduction. The first is taxonomic: more studies are needed in reptiles and amphibians, and non-arthropod invertebrates. The second is methodological: relatively few studies examine short-term exposure, exposure of different life stages, exposure to temperature fluctuations, exposure to low temperatures, or can separate male versus female contributions to temperature-induced fertility loss. The third is geographic: more sampling is needed of species that live in tropical or polar regions and in the Southern Hemisphere. While we have mapped this literature and drawn attention to some gaps, we further hope to perform formal analyses of the data we do have, in order to answer a range of key questions relating to temperature effects on animal reproduction: For example, to what extent do negative reproductive effects occur at lower-than-lethal temperatures? With respect to the available data, to what extent do the sexes respond differently? Is reproduction more sensitive to temperature at some life stages over others? And, are species from some taxonomic groups or geographic regions more sensitive than others?

**AUTHOR CONTRIBUTIONS**

**Conceptualisation:** Amanda Bretman, Liam R. Dougherty, Claudia Fricke, Maarit I. Maenpaa, Rhonda R Snook, and Tom A. R. Price.

**Data curation:** Fay Frost and Liam R. Dougherty.

**Funding acquisition:** Amanda Bretman, Claudia Fricke, Liam R. Dougherty, Rhonda R Snook, and Tom A. R. Price.


**Methodology:** Amanda Bretman, Claudia Fricke, Fay Frost, Liam R. Dougherty, Maarit I. Maenpaa, Malgorzata Lagisz, Rhonda R Snook, Shinichi Nakagawa, and Tom A. R. Price.

**Project administration:** Amanda Bretman, Claudia Fricke, Liam R. Dougherty, Rhonda R Snook, and
CONFLICT OF INTEREST STATEMENT
The authors declare no conflicts of interest.

PEER REVIEW
The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1002/2688-8319.12303.

DATA AVAILABILITY STATEMENT
All data are available on Figshare: https://doi.org/10.6084/m9.figshare.24182745.v1 (Dougherty et al., 2023).

ORCID
Liam R. Dougherty https://orcid.org/0000-0003-1406-0680
Fay Frost https://orcid.org/0000-0003-2382-2990
Mairit I. Maenpaa https://orcid.org/0000-0002-1906-5811
Melissa Rowe https://orcid.org/0000-0001-9747-041X
Benjamin J. Cole https://orcid.org/0000-0002-2747-2580
Ramakrishnan Vasudeva https://orcid.org/0000-0002-3831-0384
Patrice Pottier https://orcid.org/0000-0003-2106-6597
Eva Schultner https://orcid.org/0000-0002-5069-9732
Erin L. Macarthy https://orcid.org/0000-0003-3866-143X
Jamie L. Smith https://orcid.org/0009-0009-2388-2134
Pau Carazo https://orcid.org/0000-0002-1525-6522
Marco Graziano https://orcid.org/0000-0001-8850-9036
Hester Weaving https://orcid.org/0000-0002-9093-0556
Berta Canal Domenech https://orcid.org/0000-0003-1447-5704
David Berger https://orcid.org/0000-0003-0196-6109
Abhishek Meena https://orcid.org/0000-0001-9731-3924
Tom Rhys Bishop https://orcid.org/0000-0001-7061-556X
Daniel W. A. Noble https://orcid.org/0000-0001-9460-8743
Pedro Simões https://orcid.org/0000-0002-4253-1200
Julian Baur https://orcid.org/0000-0002-4739-2756
Merel C. Breedveld https://orcid.org/0000-0002-3198-7683
Erik I. Svensson https://orcid.org/0000-0001-9006-016X
Lesley T. Lancaster https://orcid.org/0000-0002-3135-4835
Jacintha Ellers https://orcid.org/0000-0003-2665-1971
Alessio N. De Nardo https://orcid.org/0000-0003-3708-1492
Marta A. Santos https://orcid.org/0000-0003-1463-6037
Steven A. Ramm https://orcid.org/0000-0001-7786-7364
Szymon M. Drobiński https://orcid.org/0000-0001-8101-6247
Matteo Redana https://orcid.org/0000-0003-0371-9347
Cristina Tuni https://orcid.org/0000-0002-7190-1143
Natalie Pilakouta https://orcid.org/0000-0001-8503-520X
Z. Valentina Zizzari https://orcid.org/0000-0002-9945-3071
Graziella Iossa https://orcid.org/0000-0001-6813-4361
Stefan Lüpold https://orcid.org/0000-0002-5069-1992
Marelke Koppik https://orcid.org/0000-0002-3433-7248
Regan Early https://orcid.org/0000-0003-4108-5904
Clelia Gasparini https://orcid.org/0000-0001-9712-1142
Shinichi Nakagawa https://orcid.org/0000-0002-7765-5182
Malgorzata Logisz https://orcid.org/0000-0002-3993-6127

ACKNOWLEDGEMENTS
We thank two anonymous reviews for suggestions that improved the manuscript. This work was funded by the European Society for Evolution (which funds a Special Topic Network on Evolutionary Ecology of Thermal Fertility Limits to CF, AB, RRS and TARP), the Natural Environment Research Council (NE/P002692/1 to TARP, AB and RRS, NE/X011550/1 to LRD and TARP), the Biotechnology and Biological Sciences Research Council (BB/W016753/1 to AB, TARP and RRS) and a Heisenberg fellowship from the German Research Foundation (FR 2973/11-1 to CF).
REFERENCES


SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Figure S1: Phylogenetic tree showing the relationship between the 1191 species in the systematic map. Note that branch lengths are standardised because branch length information was not available.