MOLECULAR ECOLOGY

Molecular Ecology (2015) **24**, 4042–4051

NEWS AND VIEWS

OPINION

Formalizing the definition of meta-analysis in *Molecular Ecology*

ALTHEA A. ARCHMILLER,* ERIC F. BAUER,† REBECCA E. KOCH,† BHAGYA K. WIJAYAWARDENA,‡ AMMU ANIL,§ JACK J. KOTTWITZ,¶ AMELIA S. MUNSTERMAN** and ALAN E. WILSON§[1]

*School of Forestry and Wildlife Sciences, Auburn University, 3301 Duncan Dr., Auburn AL 36849, USA, †Department of Biological Sciences, Auburn University, Auburn AL 36849, USA, ‡Department of Biological Sciences, Purdue University, West Lafayette IN 47906, USA, §School of Fisheries, Aquaculture, and Aquatic Sciences, Auburn University, Auburn AL 36849, USA, ¶Department of Anatomy, Physiology, and Pharmacology, College of Veterinary Medicine, Auburn University, Auburn AL 36849, USA, **Department of Clinical Sciences, College of Veterinary Medicine, Auburn University, Auburn AL 36849, USA*

Meta-analysis, the statistical synthesis of pertinent literature to develop evidence-based conclusions, is relatively new to the field of molecular ecology, with the first meta-analysis published in the journal *Molecular Ecology* in 2003 (Slate & Phua 2003). The goal of this article is to formalize the definition of meta-analysis for the authors, editors, reviewers and readers of *Molecular Ecology* by completing a review of the meta-analyses previously published in this journal. We also provide a brief overview of the many components required for meta-analysis with a more specific discussion of the issues related to the field of molecular ecology, including the use and statistical considerations of Wright's F_{ST} and its related analogues as effect sizes in meta-analysis. We performed a literature review to identify articles published as 'meta-analyses' in *Molecular Ecology*, which were then evaluated by at least two reviewers. We specifically targeted *Molecular Ecology* publications because as a flagship journal in this field, meta-analyses published in *Molecular Ecology* have the potential to set the standard for meta-analyses in other journals. We found that while many of these reviewed articles were strong meta-analyses, others failed to follow standard meta-analytical techniques. One of these unsatisfactory meta-analyses was in fact a secondary analysis. Other studies attempted meta-analyses but lacked the fundamental statistics that are considered necessary for an effective and powerful meta-analysis. By drawing attention to the inconsistency of studies labelled as meta-analyses, we emphasize the importance of understanding the components of traditional meta-analyses to fully embrace the strengths of quantitative data synthesis in the field of molecular ecology.

Keywords: data synthesis, effect size, F_{ST}, meta-analyses, quantitative review, secondary analysis

Received 4 August 2014; revision received 15 May 2015; accepted 3 June 2015

Introduction

Meta-analysis was most succinctly defined by Glass (1976) as 'the analysis of analyses'. In other words, the authors of a meta-analysis compile and quantitatively synthesize the results of available and pertinent studies using a meaningful common statistic, called an effect size, to address a specific research question (Koricheva *et al.* 2013). Such meta-analytical techniques are recommended to synthesize the literature because they often have higher statistical power than an individual primary study, due to the increased precision of the summary effect size estimate (Borenstein *et al.* 2009; Madden & Paul 2011). However, the power of meta-analysis relies on very specific methodological and statistical treatment of the individual studies reviewed (Rosenberg *et al.* 2004). In this regard, meta-analyses are not secondary analyses, in which original data are collected from published literature and re-analysed; nor are they narrative reviews, in which authors may subjectively choose papers to discuss. A true meta-analysis also does not rely on surveying individual study *P*-values as a vote-counting framework for reaching conclusions (Osenberg *et al.* 1999; Madden & Paul 2011). Meta-analyses do, however, quantitatively synthesize results across all identified, relevant studies to determine not only whether an overall effect is present, but also what underlying covariates may influence the trend across studies.

The purpose of this opinion article is to provide authors and editors with the tools to increase the use and effectiveness of meta-analysis in *Molecular Ecology* and related journals. We first briefly formalize the definition of meta-analysis in this field, and then use the results from our own review of self-described meta-analyses published in *Molecular Ecology* to illustrate specific examples of how these papers failed to meet, met or exceeded the expectations of an effective meta-analysis. An exhaustive or technical review of general meta-analytical techniques is beyond the scope of this paper, but we provide references for

Correspondence: Althea A. ArchMiller, Fax: 334 844 1084; E-mail: althea.archmiller@gmail.com

[1]Present address: Division of Environmental Biology, National Science Foundation, Suite 635, Arlington, VA, 22230, USA

many of the outstanding publications and textbooks that discuss the specific methodology and issues regarding meta-analysis. In this article, we build on this rich meta-analysis literature to discuss how the field of molecular ecology raises unique issues related to the choice of effect sizes, specifically Wright's F_{ST} and its analogues, and propose general recommendations to strengthen future meta-analyses in this field.

Brief introduction to meta-analysis

Meta-analysis has been eloquently detailed in other references, ranging from topics on general meta-analysis (e.g. Borenstein 2009; Borenstein *et al.* 2009), on ecological meta-analysis (e.g. Arnqvist & Wooster 1995; Gurevitch & Hedges 1999; Hedges *et al.* 1999; Osenberg *et al.* 1999; Rosenberg *et al.* 2004; Madden & Paul 2011; Philibert *et al.* 2012; Koricheva *et al.* 2013) and on phylogenetic meta-analysis (Adams 2008; Lajeunesse 2009; Chamberlain *et al.* 2012). Based upon this rich literature, we briefly formalize the definition of effective meta-analyses by describing each component of a meta-analysis, including the initial literature search, effect size choice and data extraction, statistical data synthesis, other statistical considerations, and the documentation of the data set and references.

Literature search

Research synthesis, which includes both meta-analyses and systematic reviews, begins with a literature search (Gurevitch & Nakagawa 2015). Once a testable hypothesis has been determined, an appropriate search protocol is formulated to find all the studies that have addressed this question. Two important aspects of the literature search are that the search is both thorough and repeatable (Gates 2002; Philibert *et al.* 2012). Thorough literature searches of all of the relevant studies must be broad in scope and context to reduce the possible impacts of publication bias and to maintain objectivity (Koricheva & Gurevitch 2014). When available, 'grey literature' (e.g. documents such as theses, dissertations and conference proceedings) and unpublished studies may also be included in the meta-analysis data set. To ensure repeatability, meta-analysis literature search methods should include the specific Boolean operators, search databases and dates, and criteria used for inclusion or exclusion. Furthermore, the use of the PRISMA (http://www.prisma-statement.org/statement.htm) flowchart and checklist for systematic reviews to document the search protocol and results is recommended (Koricheva & Gurevitch 2014).

Effect size choice and data extraction

Picking the appropriate statistical techniques is required in any scientific study and is equally important for meta-analysis. Subsequent to the literature search, the authors of a meta-analysis must pick an effect size that can be statistically tested against an appropriate null hypothesis (Osenberg *et al.* 1999; Madden & Paul 2011). In general, the effect size should be picked or developed with awareness of its underlying statistical properties and assumptions, and also chosen to match the biological objectives and hypotheses of the meta-analysis (Osenberg *et al.* 1997, 1999). Borenstein (2009) suggests four criteria for selecting an appropriate effect size: (i) it must address comparable questions across studies and be unaffected by study design; (ii) it should not require a secondary analysis of data from the original studies but instead be calculated from commonly reported metrics; (iii) it should have known statistical properties, such as sampling distribution, so that variances and confidence intervals can be mathematically derived; and (iv) it should be meaningful to those working in the field in which the meta-analysis is being conducted. Effect sizes commonly used in meta-analyses include standardized mean differences, Hedges' d or g, correlation coefficients, odds ratios or log response ratios (Gurevitch & Hedges 1999; Hedges *et al.* 1999). In some cases, effect sizes may have to be converted from other metrics to be comparable across studies; for instance, a reported chi-squared value may be statistically converted into a correlation coefficient (Chapman *et al.* 2009; Miller & Coltman 2014). After choosing an effect size, authors should also extract related meta-data from each publication, including variance, covariates, sample sizes and notes about the study (e.g. location, perceived study quality).

Most of the traditional meta-analysis effect sizes are equally suitable for meta-analyses in molecular ecology; for example, one can test for heterozygosity–fitness relationships with the correlation coefficient (Chapman *et al.* 2009) or compare mean allelic richness or heterozygosity between populations with Hedges' d (Aguilar *et al.* 2008) or the log response ratio (Johannesson & André 2006). However, some statistics used in molecular ecology are unique to evolutionary sciences, raising important questions about their value as potential effect sizes for meta-analysis. Mengersen & Gurevitch (2013) provide valuable guidelines for using unconventional effect sizes, highlighting the need to use truly comparable and standardized effect sizes with an appropriate sampling variance for weighting. One set of unconventional statistics common in the field of molecular ecology is Wright's F_{ST} and its analogues, including D, F_{IT}, F'_{ST}, G_{ST}, G'_{ST} and R_{ST}, which are indices of the variation in allelic frequencies among populations that provide measures of population differentiation based on the genetic polymorphism or fixation of alleles (Wright 1949; Balloux & Goudet 2002; Avise 2004; Meirmans & Hedrick 2011). In meta-analyses, F_{ST} values can be used as effect sizes if they are computed by consistent means across studies (i.e. addressing similar questions and using similar markers or DNA loci) to meet the criteria of comparability outlined by Borenstein (2009) and Mengersen & Gurevitch (2013). F_{ST} values also meet Borenstein's (2009) additional criteria for effect sizes in that they are commonly reported (and therefore calculable without secondary analyses) and have known statistical properties and distributions (Weir & Cockerham 1984; Long 1986; Leviyang 2011; Meirmans & Hedrick 2011).

In practice, F_{ST} and analogues are constrained by expected within-population heterozygosity (Meirmans & Hedrick 2011), which may hinder cross-study comparisons. Thus, we recommend the use of a standardized analogue of F_{ST}, such as F'_{ST}, G'_{ST}, R_{ST} or Φ'_{ST}, as a meta-analysis effect size to alleviate the confounding effects of different maximum possible values across studies (Jost 2008, 2009; Heller & Siegismund 2009). Authors must also consult the literature for the advantages and disadvantages, as well as the proper application of these statistics, to choose the appropriate metric for the research question (Jost 2008, 2009; Heller & Siegismund 2009; Whitlock 2011). The variance of F_{ST} and its analogues can be estimated with Taylor series expansions for single di-allelic loci and estimated for multi-allelic loci with jackknife permutation (Weir & Cockerham 1984; Long 1986); however, the variance is exceedingly complicated to calculate when it is applied to multi-allelic loci and thus may not be cited or calculable from reported information. Rather than rely upon unweighted analyses in such cases, we recommend contacting authors for more information, calculating variances when possible using the computer programs cited in Meirmans & Hedrick (2011), or using imputation to estimate variances. Standardized analogues of F_{ST} are associated with higher variances than their nonstandardized counterparts (Meirmans & Hedrick 2011). As such, the use of standardized F_{ST} analogues as an effect size may provide more conservative meta-analysis results compared to the nonstandardized counterparts. Finally, F_{ST} and its standardized analogues can range from 0 to 1 and are not normally distributed (Long 1986; Leviyang 2011). Transformations may be applied to F_{ST} to approximate normality, but we caution the use of transformations and recommend that F_{ST} -related effect sizes be analysed within a generalized linear mixed model (see Koricheva *et al.* 2013). Meta-analysis of F_{ST} effect sizes can alternatively be completed by comparing F_{ST} -related parameters between populations with traditional effect sizes with known distributions (e.g. Friesen *et al.* 2007; Phillips *et al.* 2012).

Statistical data synthesis

As stated earlier, a detailed accounting of the statistics required for meta-analysis is beyond the scope of our review, and instead, we recommend previous publications (e.g. Borenstein *et al.* 2009; Koricheva *et al.* 2013; Mengersen & Gurevitch 2013) for more thorough and detailed descriptions of the statistical techniques of meta-analyses. In general, once the database of effect sizes and meta-data has been compiled, the statistical meta-analysis begins with the choice of using either a fixed-effect or a random-effects model (Hedges *et al.* 1999; Borenstein *et al.* 2009). The fixed-effect model assumes within-study sampling error as the only source of variation between the estimates and considers all the studies to be functionally identical with a common underlying true effect size (Koricheva *et al.* 2013). But, in molecular ecology and related fields, fixed-effect models are rarely appropriate, as studies almost always differ from one another in more ways than the sampling variation, making it rarely logical to estimate a truly common effect size. Therefore, when the true outcomes (effect sizes) differ among studies for a variety of reasons or if inferences about a wider population are desired, the random-effects model is more appropriate and generally suggested. Each effect size is weighted by the inverse of the variance to give studies with high precision and power (i.e. large sample sizes) more impact on the final effect size than those with lower power and to maintain low type II error rates (Gurevitch & Hedges 1999; Mengersen & Gurevitch 2013). In a fixed-effect model, only within-study variance is used to weight an associated effect size, but in a random-effects model, both within-study variation and between-study variation (i.e. heterogeneity) are used to weight effect sizes (Gurevitch & Hedges 1999; Borenstein *et al.* 2009; Madden & Paul 2011; Philibert *et al.* 2012). Thus, the resultant confidence interval of the summary effect size from a random-effects model meta-analysis will be wider if heterogeneity is present or will reduce to the fixed-effect model if heterogeneity is absent. While using a weighted analysis is strongly recommended, it is not always necessary (i.e. conclusions may not differ between weighted and unweighted analyses) or possible (Englund *et al.* 1999; Gurevitch & Hedges 1999; Gurevitch *et al.* 2001; Koricheva & Gurevitch 2014). Some alternative approaches have been suggested for conducting meta-analysis when variance measures are either not available or calculable from individual studies, including resampling mean effect sizes and associated confidence intervals (Adams *et al.* 1997; Gurevitch & Hedges 1999; Gurevitch *et al.* 2001) or using imputation to estimate effect size variances from related publications or previous meta-analyses (Koricheva *et al.* 2013).

Exploring the factors that account for heterogeneity can demonstrate how covariates, such as sampling technique or study time frame, affect the summary effect size and broaden the applicability of the meta-analysis to wider populations. Partitioning and explaining heterogeneity is often accomplished through subgroup analyses or meta-regression (Gurevitch & Hedges 1999; Osenberg *et al.* 1999; Borenstein *et al.* 2009; Stewart 2010; Mengersen & Gurevitch 2013). Meta-regression differs from ordinary least squares (OLS) regression in many ways that are important for an effective synthesis of the data. Most notably, meta-regression weights each effect size by both within- and among-study variance and can be used in molecular ecology meta-analyses as a mixed-effects model, representing heterogeneity as a random effect and the covariates as fixed effects (Gurevitch & Nakagawa 2015). Also, to determine slope significance, meta-regression relies on different probability distributions (*Z*- and *Q*-tests) than OLS regression (*t*- and *F*-tests). In random-effects meta-regression, heterogeneity must be computed at each step of the prediction slope, which results in larger slope confidence and prediction intervals than OLS regression (Borenstein *et al.* 2009). As with OLS regression, the potential impacts of multicollinearity and confounding effects among

covariates should be addressed in a multiple meta-regression analysis (Koricheva & Gurevitch 2014).

Other statistical considerations

The authors of a meta-analysis should address the limitations of their meta-analysis, such as possible nonindependence of effect sizes. Nonindependence can occur when more than one effect size was extracted per study or research group (Philibert *et al.* 2012), when effect sizes were reported from repeated measures (Koricheva *et al.* 2013) or when the studies compiled for meta-analysis used species with common ancestry (Adams 2008; Lajeunesse 2009; Chamberlain *et al.* 2012). Nonindependence can be mitigated by extracting only one effect size per study or by averaging effect sizes within a study (Gurevitch & Hedges 1999; Gates 2002; Philibert *et al.* 2012). However, explicitly incorporating the nonindependence, for example using a hierarchical model to incorporate repeated measures (Koricheva *et al.* 2013) or *phyloMeta* (Lajeunesse 2011) to incorporate phylogenetic nonindependence, is a recommended and less conservative way to deal with nonindependent effect sizes. A second statistical consideration is the potential presence of studies that disproportionately affect the summary effect size. Such influential studies can be identified through sensitivity analysis, which is built into many meta-analytic software packages (Borenstein *et al.* 2009; Gurevitch & Nakagawa 2015).

Another important but often-overlooked component of meta-analysis is the possibility of publication bias, where the literature published and included in a meta-analysis is biased towards favoured results (Gurevitch & Hedges 1999; Borenstein *et al.* 2009; Madden & Paul 2011). In particular, studies showing statistically significant results often have a higher chance of publication, whereas nonsignificant studies may have difficulty being published and may be absent from the meta-analysis data set. Publication bias can also occur through searches for primarily English language publications or due to limitations inherent to electronic search databases (i.e. duplications retrieved, nonindexed journals) (see Lowry *et al.* 2012). Such publication bias may lead to inaccurate conclusions from meta-analysis (Rosenthal 1979). Many methods can be used to assess publication bias, including (i) a fail-safe number method (estimates the number of unpublished studies with nonsignificant results required to nullify the significance of the analysis); (ii) a funnel plot (graphs effect size vs. study variance); (iii) correlation tests (measures the relationship between standardized effect size and sample size); and (iv) the trim and fill method (an iterative technique that adjusts funnel plots to account for possible publication bias) (Møller & Jennions 2001). It should be noted that over the years, these methods have been subjected to many deliberations over their respective weaknesses (Sutton *et al.* 2000). For instance, heterogeneity, data irregularities and chance could also produce asymmetrical funnel plots (Egger *et al.* 1997; Sterne *et al.* 2011). Therefore, we recommend being thorough when identifying studies to include in the meta-analysis, using multiple methods to acquire greater confidence in the effect of publication bias on meta-analysis results and exerting caution when interpreting results. Research bias is another concern that arises due to the deliberate and nonrandom selection of organisms or conditions (Gurevitch & Hedges 1999). Research bias is unfortunately difficult to detect, yet it remains a solid issue in ecological and evolutionary studies and is therefore worth discussing in any published meta-analysis.

Documentation

In a meta-analysis, it is also important that the authors provide a full list of the studies used in their meta-analysis; similarly, provision of the full data set or a summary of effect sizes and variances associated with each study are recommended for both transparency for the audience and repeatability as described above. Authors should also cite the program used to analyse the data, as well as the specific procedures, when applicable.

Review of *Molecular Ecology* meta-analyses: methods

For our review, we used the aforementioned components of a meta-analysis to develop a rubric to assess self-described meta-analyses published in *Molecular Ecology* up to the time of our literature search (Table 1). The primary purpose of the rubric evaluation was to highlight when essential meta-analytical techniques, as described in the literature (e.g. Gurevitch & Hedges 1999; Borenstein *et al.* 2009; Madden & Paul 2011), were unclear or missing for a particular study. We searched for papers in May 2014, using the search terms 'meta', 'meta-' or 'meta-analys*' on Google Scholar and through the *Molecular Ecology* website search function. We included papers in our review only if they referred to themselves as meta-analyses (or a derivation thereof, such as 'metastudy') and excluded papers that were not self-described meta-analyses (e.g. primary literature that cited other meta-analyses). Two to three reviewers familiar with meta-analysis evaluated the meta-analytical components of each study following the rubric detailed in Table 1. If the techniques used for a particular aspect of the meta-analysis under review were not clear, a zero was given for the corresponding rubric item because of the need for clarity and repeatability in meta-analyses. Because some rubric components are recommended but not necessary for an effective meta-analysis (i.e. subgroup analysis or meta-regression), our expectation for these meta-analyses was the completion and documentation of 15 or more of the 17 traditional meta-analysis components discussed in the Brief Introduction to Meta-Analysis section.

Results

The results of our review of *Molecular Ecology* meta-analysis papers are summarized in Fig. 1 and detailed in Table S1 (Supporting information). We identified 25 papers initially and included 18 studies that used the term

Table 1 Rubric used to review meta-analyses for *Molecular Ecology*

Literature search		
Boolean operators	(1)	Included specific keywords
	(0)	Specific search keywords were not given
Search databases and dates	(1)	Stated where and when search was completed
	(0)	Stated only where or only when or search details were not given
Exclusion/inclusion criteria	(1)	Gave specific information about why papers were retained or rejected
	(0)	No exclusion or inclusion criteria stated
Effect size choice and data extraction		
Effect size choice		What effect size was used? Were effect size conversions used?
Effect size extraction/conversion	(1)	Specified the effect size extracted from each study and the conversions used, if applicable
	(0)	Effect size choice was not stated
Meta-data documentation	(1)	Stated specifically what other information was extracted from each study (e.g. variance, sample size, locations, covariates)
	(0)	Meta-data not described
Statistical data synthesis		
Model Choice	(1)	Stated if fixed-effect or random-effects model was used or discussed between- and within-study variation
	(0)	Model choice or two sources of variation not stated
Heterogeneity estimation	(1)	Calculated heterogeneity statistics (e.g. Q, τ, I)
	(0)	No mention or calculation of heterogeneity
Individual study weight calculation	(1)	Addressed weighted meta-analysis, how each study was weighted, or why unweighted analysis was used
	(0)	Unweighted analysis used without justification or unclear if weighted
Effect size summarization	(1)	Stated or quantified an overall effect size for all studies
	(0)	No summarization of effect size across studies
Subgroup analysis	(1)	Summarized effect sizes for subgroups or categorical covariates
	(0)	Subgroup analysis was not used
Meta-regression	(1)	Used specific meta-regression techniques
	(0)	Used no meta-regression or only ordinary least squares regression
Other statistical considerations		
Nonindependence	(1)	Quantified or stated impacts of nonindependence
	(0)	Did not address nonindependence
Sensitivity analysis	(1)	Quantified impact of individual effect sizes with sensitivity analysis or discussed potential impacts of individual study effect sizes
	(0)	No discussion of the impact of individual studies in these ways
Publication bias	(1)	Discussed or quantified publication bias
	(0)	No specific mention of publication bias
Documentation		
List of references	(1)	Included list of the studies used in analysis in either article or supplementary information (SI)
	(0)	References of the studies used was not given
Data set used in the meta-analysis	(1)	Provided data set used for meta-analysis in article, SI, or online
	(0)	Data set not provided
Statistical program	(1)	Listed the statistical program used for meta-analysis
	(0)	Statistical program not given

'meta-analysis' or a derivative thereof to describe their publication. The seven publications excluded from our review were primary literature papers that referred directly to other meta-analyses in their text (e.g. Rossetto *et al.* 2009), thus being found through our Boolean keyword search on Google Scholar yet inappropriate for comparison with other meta-analyses. Our search results are presented within the PRISMA (http://www.prisma-statement.org/statement.htm) framework in Fig. S1 (Supporting information). We recognize that meta-analytical techniques may be used in other papers in *Molecular Ecology* outside those gathered here, which may limit the scope of our reports. In fact, meta-analysis is not consistently used as a keyword in *Molecular Ecology*, so we recommend that it be included in keywords or article titles for greater clarification of the nature of each study. Based on our 2014 search, the earliest *Molecular Ecology* paper claiming to be a meta-analysis was published in 2003 (Fig. 2; Slate & Phua 2003) and the most recent were in 2014 (Gao *et al.* 2014; Miller & Coltman 2014; Ni *et al.* 2014; Pinsky & Palumbi 2014; Raeymaekers *et al.* 2014; Tedersoo *et al.* 2014). There is an obvious increase in the number of published meta-analyses over

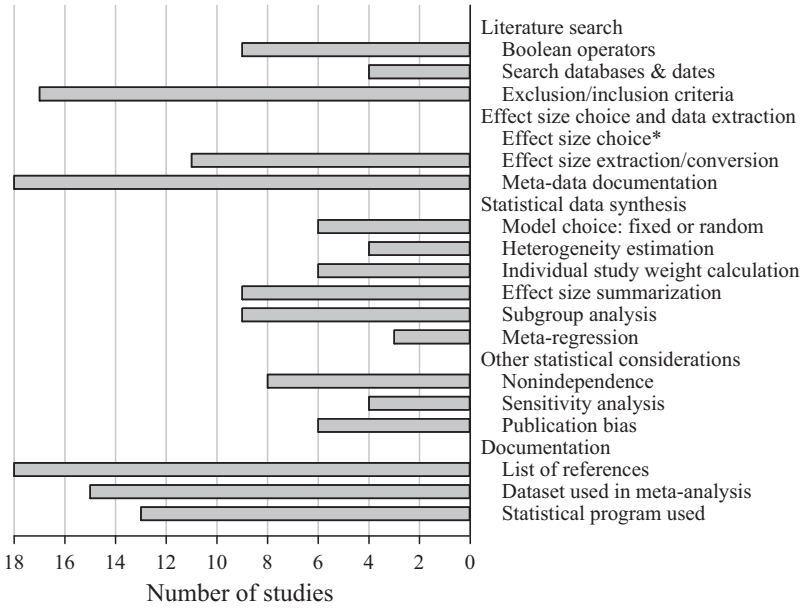

Fig. 1 The components of a traditional meta-analysis and the number of studies that successfully accomplished each component based on our review. Full review results can be found in Table S1 (Supporting information). *Note: Effect size choice was a categorical review category and thus was not tallied.

time in *Molecular Ecology* as well as in the broader scientific literature (Fig. 2). We reviewed a total of 17 meta-analysis components in our rubric, and the range of components completed was 3 (Johannesson & André 2006) to 16 (Aguilar *et al.* 2008; Chapman *et al.* 2009; Table S1, Supporting information).

Literature search

Four of the eighteen studies we reviewed successfully described their literature search as to be fully repeatable (3/3 rubric components); five studies gave specific Boolean operators and exclusion/inclusion criteria but did not state when the search was completed (2/3 components), and nine studies provided little to no information on their searches (≤1/3 components; Table S1, Supporting information).

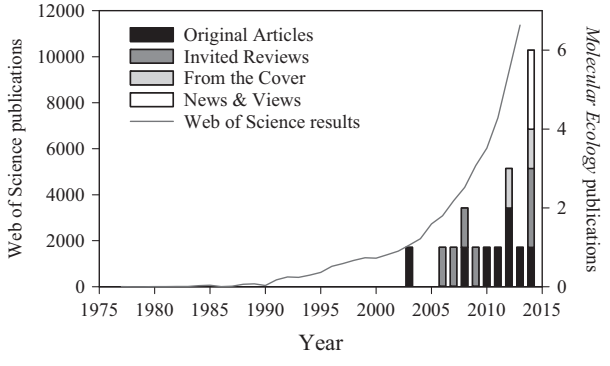

Fig. 2 Published meta-analyses by year. Publications shown are those found when searching all Web of Science journals for 'meta-analys*' (July 2014; left *Y*-axis; line) and the publications in *Molecular Ecology* reviewed in this work by article category (May 2014; right *Y*-axis; bars).

Within the studies we reviewed from *Molecular Ecology*, Chapman *et al.* (2009) and Sutton *et al.* (2011) provide outstanding examples of how the literature search can be handled in a meta-analysis, including systematic and detailed search descriptions as well as specific inclusion and exclusion criteria. However, other articles, such as Slate & Phua (2003) and Raeymaekers *et al.* (2014), failed to adequately describe their literature search protocol in a thorough and repeatable manner.

Effect size choice and data extraction

Every publication that was reviewed successfully documented the meta-data that they extracted from each study; however, seven studies did not mention the specific effect size or dependent variable that the authors extracted from each study (Fig. 1; Table S1, Supporting information). Among the exemplary studies in this regard were Sutton *et al.* (2011) and Miller & Coltman (2014), which included exhaustive and authoritative descriptions of how their respective effect sizes were calculated, what the underlying distribution and assumptions were about these effect sizes, and how they summarized these metrics across studies.

Of the five studies using F_{ST} or related derivatives for effect sizes, two compared the mean F_{ST} between two populations (Table S1, Supporting information; Friesen *et al.* 2007; Phillips *et al.* 2012). For example, Friesen *et al.* (2007) analysed the effects of breeding range, colony dispersion, nonbreeding distribution, foraging range, population size and climate zone on population genetic structure of seabirds using mean difference of F_{ST} or Φ_{ST} between categories. They notably took the necessary step of ensuring only studies with comparable genetic markers, specifically Domain I markers were used. Ni *et al.* (2014) and Raeymaekers *et al.* (2014) were the only studies to use F_{ST} or analogues as direct effect sizes. Finally, as part of their

meta-analysis, Johannesson & André (2006) performed a direct OLS regression analysis of F_{ST} vs. geographical distance from the entrance of the Baltic Sea to determine how geographical isolation affects genetic diversity.

Statistical data synthesis

Our review of meta-analyses in *Molecular Ecology* revealed that studies were divided between those that completed a traditional meta-analysis vs. those that fell well short of the established meta-analytical guidelines; six studies completed at least four of six components required for effective meta-analysis, but the rest only completed two or fewer of the components (Fig. 1; Table S1, Supporting information). Johannesson & André (2006) compared species inside and outside the Baltic Sea with a log response ratio, but did not refer to it as such, and they did not use meta-analysis techniques to synthesize this effect size (Table S1, Supporting information). Likewise, Hendry *et al.* (2008) described their two effect sizes (phenotypic change quantified as Haldane and Darwin numerators) and the related meta-data that they extracted from each study, statistically summarized the effect sizes with subgroup analysis, and even discussed publication bias; yet, they did not provide clear information as to how they dealt with random effects, weighting of individual studies or between-study heterogeneity. Finally, a number of the reviewed studies performed regression analyses, but used OLS regression analyses rather than meta-regression (Johannesson & André 2006; Gao *et al.* 2013, 2014; Pinsky & Palumbi 2014; Tedersoo *et al.* 2014). For example, Tedersoo *et al.* (2014) discussed how sampling effects of individual studies may have a profound effect on the results, but then did not formally incorporate the variation of the studies into the regression model, as would occur with meta-regression.

Other statistical considerations

In our review, eight of 18 studies addressed the potential for nonindependence within their data sets; only four studies successfully utilized sensitivity analysis to determine the impact of anomalous effect sizes, and six studies quantified or described the potential for publication bias in their meta-analysis (Fig. 1; Table S1, Supporting information). Hendry *et al.* (2008) provided the best discussion of bias and study limitation of all the studies we reviewed, but they did not provide a quantitative analysis of publication bias, which would have further strengthened their paper. Finally, the fact that some studies neither provide a repeatable search protocol nor discuss potential publication bias (e.g. Friesen *et al.* 2007) is cause for concern, because there is no evidence in such a case that the studies selected by the authors were chosen thoroughly and objectively.

Documentation

All 18 meta-analyses that we reviewed successfully listed the references used to create the meta-analytical database; however, only 15 provided the data set directly and 13 documented the statistical program (Fig. 1; Table S1, Supporting information).

Discussion

In summary, of the 18 studies we reviewed, only four (22%) met our expectations of an effective meta-analysis, which required completing and documenting at least 15 of the 17 fundamental meta-analysis components listed in Fig. 1 (Aguilar *et al.* 2008; Chapman *et al.* 2009; Sutton *et al.* 2011; Miller & Coltman 2014). Pinsky & Palumbi (2014) and Slate & Phua (2003) met many of the requirements of effective meta-analyses, accomplishing 12 and 10 of the components, respectively, but they both lacked critical components. Specifically, Pinsky & Palumbi (2014) used OLS regression rather than meta-regression, and Slate & Phua (2003) lacked a thorough and repeatable search procedure. Ten studies (56%) accomplished nine or fewer of the 17 components (Johannesson & André 2006; Friesen *et al.* 2007; Hendry *et al.* 2008; Phillips *et al.* 2012; Sullam *et al.* 2012; Tedersoo *et al.* 2012, 2014; Gao *et al.* 2013, 2014; Raeymaekers *et al.* 2014) and thus were well below our expectations for effective meta-analyses. In each of these cases, a fairly thorough search procedure was followed, an adequate effect size was chosen, and pertinent meta-data was extracted from each study. However, in most of these studies, standard parametric statistics were then utilized rather than meta-analysis techniques (e.g. heterogeneity estimation, weighted analysis, subgroup or meta-regression analyses). For instance, Tedersoo *et al.* (2012) completed a spatial analysis of extracted species richness values from published literature, and we believe they may have been able to incorporate heterogeneity and weighted analysis within the spatial analysis framework to strengthen their statistical conclusions. Thus, these 10 studies began as strong meta-analyses but failed to harness the statistical power of meta-analytical techniques, which were specifically designed to outperform standard parametric statistics. Ignoring all of the literature that details the proper statistical methods of meta-analysis is simply not justified.

We reviewed one study that did not extract means and error terms from published studies, but rather used a secondary analysis of published studies to analyse rDNA sequences in a phylogenetic framework (Bonito *et al.* 2010). We acknowledge that in this case, these authors used the appropriate, *but not meta-analytical*, statistical techniques for their data structure. Thus, the traditional meta-analysis definition is an inappropriate fit for these types of studies, and we suggest that such studies be more appropriately labelled as a 'secondary analyses' rather than as 'meta-analyses'. Depending on the research question, we agree that these types of analyses may be as relevant and statistically sound as meta-analyses. We are not denouncing these authors and their work; instead, we argue for the proper delineation of meta-analysis vs. secondary analysis to encourage the proper understanding and use of meta-analytical techniques.

Finally, we conclude that F_{ST} and its analogues meet the criteria laid out by Borenstein (2009) and Mengersen & Gurevitch (2013) for proper effect sizes. They are commonly reported in the primary literature, are a critical component of molecular ecology analyses, can be standardized among comparable studies and can be calculated without secondary analyses. However, we recommend using a standardized version of F_{ST} (e.g. F'_{ST}, G'_{ST}) as an effect size for meta-analysis to avoid the difficulty of calculating error distributions for F_{ST} values.

Conclusions

In summary, our review of the meta-analyses published in *Molecular Ecology* reveals a systemic misunderstanding of the components required for comprehensive meta-analyses. In this opinion article, we briefly described the steps of a meta-analysis and completed a review of meta-analyses published in *Molecular Ecology* from 2003 to 2014. Based on our review, similar recommendations should be made for molecular ecology meta-analyses such as those previously made for plant ecology (Koricheva & Gurevitch 2014), general ecology (Gates 2002) and agronomy (Philibert *et al.* 2012), specifically:

1 Where, when and how publications were searched should be detailed, such as the specific dates, search databases, Boolean operators and inclusion/exclusion criteria.
2 The effect size extracted from each study and/or conversion methods used should be explicitly stated and be relevant to the primary research questions.
3 Effect sizes should be synthesized according to meta-analysis statistical procedures, including using either a fixed-effect or random-effects (preferred) model, weighted analysis (when applicable), and tests for heterogeneity. Subgroup and meta-regression analyses are also recommended to expand the scope of the meta-analysis, when applicable.
4 Other statistical considerations, including nonindependence of effect sizes, influential studies and publication bias, should be addressed.
5 We also recommend that authors use our review rubric given in Table 1 as a guide during meta-analysis and as a publication checklist. This checklist is also recommended for peer reviewers, when reviewing meta-analyses that have been submitted for publication.

We conclude by urging authors of forthcoming meta-analyses to refer to this comment, the wealth of references on meta-analysis and the exemplary articles highlighted in our review to publish sound and comprehensive meta-analyses. By dissuading authors, editors and readers from the improper treatment of meta-analysis research and encouraging them to embrace the potential power in meta-analyses, we aim to boost the impact and reliability of this powerful analytical tool in *Molecular Ecology*.

Acknowledgements

The authors would like to acknowledge Craig Osenberg, Jessica Gurevitch and four anonymous reviewers for their constructive feedback; the support of Dawn Boothe and Mary Delaney for early encouragement and assistance with our literature review; and Matthew ArchMiller for his critique of our methodology and review of previous drafts.

References

Adams DC (2008) Phylogenetic meta-analysis. *Evolution*, **62–3**, 567–572.
Adams DC, Gurevitch J, Rosenberg MS (1997) Resampling tests for meta-analysis of ecological data. *Ecology*, **78**, 1277–1283.
Aguilar R, Quesada M, Ashworth L, Herrerias-Diego Y, Lobo J (2008) Genetic consequences of habitat fragmentation in plant populations: susceptible signals in plant traits and methodological approaches. *Molecular Ecology*, **17**, 5177–5188.
Arnqvist G, Wooster D (1995) Meta-analysis: synthesizing research findings in ecology and evolution. *Trends in Ecology and Evolution*, **10**, 236–240.
Avise JC (2004) *Molecular Markers, Natural History, and Evolution*, 2nd edn. Sinauer, Sunderland, Massachusetts.
Balloux F, Goudet J (2002) Statistical properties of population differentiation estimators under stepwise mutation in a finite island model. *Molecular Ecology*, **11**, 771–783.
Bonito GM, Gryganskyi AP, Trappe JM, Vilgalys R (2010) A global meta-analysis of *Tuber* ITS rDNA sequences: species diversity, host associations and long-distance dispersal. *Molecular Ecology*, **19**, 4994–5008.
Borenstein M (2009) Effect sizes for continuous data. In: *The Handbook of Research Synthesis and Meta-Analysis* (eds Cooper H, Hedges LV, Valentine JC), pp. 221–226. Russell Sage Foundation, New York.
Borenstein M, Hedges LV, Higgins JPT, Rothstein HR (2009) *Introduction to Meta-Analysis*. John Wily & Sons Ltd, West Sussex, UK.
Chamberlain SA, Hovick SM, Dibble CJ *et al.* (2012) Does phylogeny matter? Assessing the impact of phylogenetic information in ecological meta-analysis. *Ecology Letters*, **15**, 627–636.
Chapman JR, Nakagawa S, Coltman DW, Slate J, Sheldon BC (2009) A quantitative review of heterozygosity—fitness correlations in animal populations. *Molecular Ecology*, **18**, 2746–2765.
Egger M, Davey Smith G, Schneider M, Minder C (1997) Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, **315**, 629–634.
Englund G, Sarnelle O, Cooper SD (1999) The importance of data-selection criteria: meta-analyses of stream predation experiments. *Ecology*, **80**, 1132–1141.
Friesen VL, Burg TM, McCoy KD (2007) Mechanisms of population differentiation in seabirds. *Molecular Ecology*, **16**, 1765–1785.
Gao C, Shi N-N, Liu Y-X *et al.* (2013) Host plant genus-level diversity is the best predictor of ectomycorrhizal fungal diversity in a Chinese subtropical forest. *Molecular Ecology*, **22**, 3403–3414.
Gao C, Shi N-N, Liu Y-X *et al.* (2014) Host plant richness explains diversity of ectomycorrhizal fungi: response to the comment of Tedersoo *et al.* (2014). *Molecular Ecology*, **23**, 996–999.
Gates S (2002) Review of methodology of quantitative reviews using meta-analysis in ecology. *Journal of Animal Ecology*, **71**, 547–557.
Glass GV (1976) Primary, secondary, and meta-analysis of research. *Educational Researcher*, **5**, 3–8.

Gurevitch J, Hedges LV (1999) Statistical issues in meta-analyses. *Ecology*, **80**, 1142–1149.

Gurevitch J, Nakagawa S (2015) Research synthesis methods in ecology. In: *Ecological Statistics: Contemporary Theory and Application* (eds Fox GA, Negrete-Yankelevich S, Sosa VJ), pp. 200–227. Oxford University Press, Oxford.

Gurevitch J, Curtis PS, Jones MH (2001) Meta-analysis in ecology. *Advances in Ecological Research*, **32**, 199–247.

Hedges LV, Gurevitch J, Curtis PS (1999) The meta-analysis of response ratios in experimental ecology. *Ecology*, **80**, 1150–1156.

Heller R, Siegismund HR (2009) Relationship between three measures of genetic differentiation G_{ST}, D_{EST}, and G'_{ST}: how wrong have we been? *Molecular Ecology*, **18**, 2080–2083.

Hendry AP, Farrugia TJ, Kinnison MT (2008) Human influences on rates of phenotypic change in wild animal populations. *Molecular Ecology*, **17**, 20–29.

Johannesson K, André C (2006) Life on the margin: genetic isolation and diversity loss in a peripheral marine ecosystem, the Baltic Sea. *Molecular Ecology*, **15**, 2013–2029.

Jost L (2008) G_{ST} and its relatives do not measure differentiation. *Molecular Ecology*, **17**, 4015–4026.

Jost L (2009) D vs G_{ST}: response to Heller and Siegismund (2009) and Ryman and Leimar (2009). *Molecular Ecology*, **18**, 2088–2091.

Koricheva J, Gurevitch J (2014) Uses and misuses of meta-analysis in plant ecology. *Journal of Ecology*, **102**, 828–844.

Koricheva J, Gurevitch J, Mengersen K (2013) *Handbook of Meta-Analysis in Ecology and Evolution*. Princeton University Press, Princeton, New Jersey.

Lajeunesse MJ (2009) Meta-analysis and the comparative phylogenetic method. *The American Naturalist*, **174**, 369–381.

Lajeunesse MJ (2011) *phyloMeta*: a program for phylogenetic comparative analysis with meta-analysis. *Bioinformatics*, **27**, 2603–2604.

Leviyang S (2011) The distribution of F_{st} and other genetic statistics for a class of population structure models. *Journal of Mathematical Biology*, **62**, 203–289.

Long JC (1986) The allelic correlation structure of Gainj- and Kalam-speaking people. I. The estimation and interpretation of Wright's *F*-statistics. *Genetics*, **112**, 629–647.

Lowry E, Rollinson EJ, Laybourn AJ *et al.* (2012) Biological invasions: a field synopsis, systematic review, and database of the literature. *Ecology and Evolution*, **3**, 182–196.

Madden LV, Paul PA (2011) Meta-analysis for evidence synthesis in plant pathology: an overview. *Phytopathology*, **101**, 16–30.

Meirmans PG, Hedrick PW (2011) Assessing population structure: F_{st} and related measures. *Molecular Ecology Resources*, **11**, 5–18.

Mengersen K, Gurevitch J (2013) Using other metrics of effect size in meta-analysis. In: *Handbook of Meta-Analysis in Ecology and Evolution* (eds Koricheva J, Gurevich J, Mengersen K), pp. 72–88. Princeton University Press, Princeton, New Jersey.

Miller JM, Coltman DW (2014) Assessment of identity disequilibrium and its relation to empirical heterozygosity fitness correlations: a meta-analysis. *Molecular Ecology*, **23**, 1899–1909.

Møller AP, Jennions MD (2001) Testing and adjusting for publication bias. *Trends in Ecology and Evolution*, **16**, 580–586.

Ni G, Li Q, Kong L, Yu H (2014) Comparative phylogeography in marginal seas of the northwestern Pacific. *Molecular Ecology*, **23**, 534–548.

Osenberg CW, Sarnelle O, Cooper SD (1997) Effect size in ecological experiments: the application of biological models in meta-analysis. *The American Naturalist*, **150**, 798–812.

Osenberg CW, Sarnelle O, Cooper SD, Holt RD (1999) Resolving ecological questions through meta-analysis: goals, metrics, and models. *Ecology*, **80**, 1105–1117.

Philibert A, Loyce C, Makowski D (2012) Assessment of the quality of meta-analysis in agronomy. *Agriculture, Ecosystems, and Environment*, **148**, 72–82.

Phillips RD, Dixon KW, Peakall R (2012) Low population genetic differentiation in the Orchidaceae: implications for the diversification of the family. *Molecular Ecology*, **21**, 5208–5220.

Pinsky ML, Palumbi SR (2014) Meta-analysis reveals lower genetic diversity in overfished populations. *Molecular Ecology*, **23**, 29–39.

Raeymaekers JAM, Konijnendijk N, Larmuseau MHD, Hellemans B, Meester LD, Volckaert FAM (2014) A gene with major phenotypic effects as a target for selection vs. homogenizing gene flow. *Molecular Ecology*, **23**, 162–181.

Rosenberg MS, Garrett KA, Su Z, Bowden RL (2004) Meta-analysis in plant pathology: synthesizing research results. *Phytopathology*, **94**, 1013–1017.

Rosenthal R (1979) The "file drawer problem" and tolerance for null results. *Psychological Bulletin*, **86**, 638–641.

Rossetto M, Crayn D, Ford A, Mellick R, Sommerville K (2009) The influence of environment and life-history traits on the distribution of genes and individuals: a comparative study of 11 rainforest trees. *Molecular Ecology*, **18**, 1422–1438.

Slate J, Phua SH (2003) Patterns of linkage disequilibrium in mitochondrial DNA of 16 ruminant populations. *Molecular Ecology*, **12**, 597–608.

Sterne JAC, Sutton AJ, Ioannidis JPA *et al.* (2011) Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *British Medical Journal*, **342**, 1–8.

Stewart G (2010) Meta-analysis in applied ecology. *Biology Letters*, **6**, 78–81.

Sullam KE, Essinger SD, Lozupone CA *et al.* (2012) Environmental and ecological factors that shape the gut bacterial communities of fish: a meta-analysis. *Molecular Ecology*, **21**, 3363–3378.

Sutton AJ, Song F, Gilbody SM, Abrams KR (2000) Modelling publication bias in meta-analysis: a review. *Statistical Methods in Medical Research*, **9**, 421–445.

Sutton JT, Nakagawa S, Robertson BC, Jamieson IG (2011) Disentangling the roles of natural selection and genetic drift in shaping variation at MHC immunity genes. *Molecular Ecology*, **20**, 4408–4420.

Tedersoo L, Bahram M, Toots M *et al.* (2012) Towards global patterns in the diversity and community structure of ectomycorrhizal fungi. *Molecular Ecology*, **21**, 4160–4170.

Tedersoo L, Bahram M, Dickie IA (2014) Does host plant richness explain diversity of ectomycorrhizal fungi? Re-evaluation of Gao et al. (2013) data sets reveals sampling effects. *Molecular Ecology*, **23**, 992–995.

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Whitlock MC (2011) G'_{ST} and D do not replace F_{ST}. *Molecular Ecology*, **20**, 1083–1091.

Wright S (1949) The genetic structure of populations. *Annals of Eugenics*, **15**, 323–354.

A.A.A. led the writing and revision of the manuscript; E.F.B. contributed F_{ST} effect size review and discussion; B.K.W., A.A. and R.E.K. wrote components of the introduction to meta-analysis and the discussion; A.A.A., E.F.B., R.E.K., B.K.W., A.A., J.J.K., A.S.M. and A.E.W. evaluated meta-analyses for the literature review; and A.E.W. conceived of the manuscript.

doi: 10.1111/mec.13264

Data accessibility

The summary results of our review are provided in the Table S1 (Supporting information).

Supporting information

Additional supporting information may be found in the online version of this article.

Figure S1 PRISMA diagram of our search protocol and results (http://www.prisma-statement.org/statement.htm).

Table S1 Complete results of review of meta-analyses published in *Molecular Ecology*.