

ARTÍCULO INVITADO

**HISTORICAL AND CURRENT APPROACHES OF RESEARCH
SYNTHESIS IN PLANT SCIENCES**

RAMIRO AGUILAR ¹

Summary: The progress of science has depended to a great extent on the ability of reaching general conclusions from a body of published research, making reviews essential to scientific development. Research reviews have provided the bases for conceptual syntheses and for development of general theories in many research areas of the natural sciences, as in the case of plant sciences. Historical approaches include narrative and “vote-counting” types of reviews, both of which present serious flaws that limit their ability to obtain robust generalizations. Currently, quantitative reviews such as meta-analyses are able to synthesize results from independent studies in a manner that is both objective and statistically defensible. Meta-analysis has been used to synthesize disparate research findings and identify patterns to arrive at conclusions unavailable to the researchers of the primary studies. Here I outline some of the most important features of meta-analysis, provide a state of the art of meta-analyses in plant sciences, and point out some future perspectives on its use. Hopefully, readers will gain interest and appreciation on meta-analytical techniques to arrive to meaningful generalizations in different areas of plant sciences.

Key words: research synthesis, history of reviews, meta-analysis, narrative reviews, vote-counting, quantitative reviews

Aproximaciones históricas y actuales en la síntesis de investigaciones en ciencias vegetales

Resumen: El progreso de la ciencia ha estado íntimamente ligado a la habilidad de alcanzar conclusiones generales a partir de investigaciones publicadas, haciendo de las revisiones una herramienta esencial para el desarrollo científico. Las revisiones en ciencia han provisto las bases para síntesis conceptuales y para el desarrollo de teorías generales en muchas áreas de investigación de las ciencias naturales como en las ciencias vegetales. Las aproximaciones históricas incluyen las revisiones narrativas y las conocidas como “cuenta votos”, las cuales presentan serias fallas que limitan su capacidad de obtener generalizaciones robustas. Actualmente, las revisiones cuantitativas como el meta-análisis son capaces de sintetizar resultados de estudios independientes en una manera que es tanto objetiva como estadísticamente defendible. Los meta-análisis han sido utilizados para sintetizar resultados dispares de investigaciones e identificar patrones para arribar a conclusiones que de otra forma no estarían disponibles para los investigadores de los estudios primarios. Aquí destaco algunas de las características más importantes del meta-análisis, proveo un estado del arte en las ciencias vegetales y subrayo algunas perspectivas futuras de su uso. Es la esperanza que los lectores adquieran conocimiento e interés sobre las técnicas meta-analíticas para arribar a generalizaciones reveladoras en diferentes áreas de las ciencias vegetales.

Palabras clave: Síntesis en investigación, historia de las revisiones, meta-análisis, revisiones narrativas, “cuenta votos”, revisiones cuantitativas.

¹ Instituto Multidisciplinario de Biología Vegetal, Universidad Nacional de Córdoba – CONICET, CC 495, (5000) Córdoba, Argentina. e-mail: raguilar@imbiv.unc.edu.ar

ON GENERALIZATIONS IN SCIENCE

Science, in the most broad sense, may be regarded as a body of rational, systematic, exact, verifiable, and therefore, fallible knowledge (Bunge, 1988; Klimovsky, 2001). Through scientific research humans have reached a conceptual reconstruction of the natural world that is progressively wider, deeper and more exact. Many seemingly impossible feats have been made possible thanks to scientific research. This distinctive and unique human endeavor is possible through a methodical process that is grounded on previous knowledge (Bunge, 1988).

The empirical sciences in particular, make use of experimental methods to test specific conclusions extracted from general hypotheses (Bunge, 1988; Klimovsky, 2001). An experiment is designed to test a particular hypothesis or set of hypotheses, and the results obtained from the experience will formally validate or refute them. In the natural sciences, however, such experiment will prove to be securely valid for the specific individual organisms studied in one place at one time (Gurevitch & Hedges, 2001; Klimovsky, 2001). That is, in areas such as medicine, genetics, the environmental, agricultural, and plant sciences, among others, it is usually impossible to replicate an experiment in the exact same way in different laboratories or natural systems to confirm the general applicability of the results. Surely enough, testing hypotheses with individual experiments are of no interest if they cannot be generalized (Gurevitch & Hedges, 2001). This inherent problem, however, has not prevented scientists to comply with the "scientific duty" of evaluating the generality of the findings of an experiment within these areas of research (Hunt, 1997). Then, what have scientist in the natural sciences done to arrive to generalizations?

The systematic revision of bodies of research on a subject has historically allowed scientists to arrive to generalizations. While there still may be a debate on the heuristic value of reviews and the level at which to make generalizations (Gurevitch & Hedges, 2001), summarizing evidence is a fundamental task in all research (Hunt, 1997). Research reviews have provided the bases for conceptual syntheses and for development of general theories in many areas of the natural sciences, as in the case of plant sciences. In this regard, the progress of science has depended greatly

on the ability of scientists to reach general conclusions from a body of published research, making reviews essential to scientific development (Arnqvist & Wooster, 1995; Hunt, 1997; Gurevitch & Hedges, 2001).

RESEARCH SYNTHESIS EVOLUTION

Historically, reviews have been conducted in a qualitative fashion. The first and most common was the narrative review, which was typically conducted by well-recognized scientists on the particular subject areas of their expertise. Narrative reviews were the primary and prevailing criteria in the scientific community to find trusted generalizations on a certain matter (Hunt, 1997). The conclusions raised from the review were obtained in a simple way: research results were colloquially described and a consensus emerged from the description. One of the problems of narrative reviews was, however, their subjectivity to cope with contradictory results; i.e., for any given particular treatment, different research studies may report positive effects, negative effects or no effects at all. Thus, in the end, the conclusions raised depended to a great extent on the individual interpretation of the reviewer to ponder a study or a group of studies with dissimilar outcomes. Such dependency on the expert's personal opinion makes narrative reviews seriously flawed (Hedges & Olkin, 1985; Hunter & Schmidt, 1990). It is important to remark, however, that for certain environmental sciences in particular, as in some areas of zoology, botany or geology, narrative reviews have made very valuable contributions in synthesizing information on the natural history of organisms, ecosystems, and landscapes, providing systematic and meticulous descriptions at all these scales.

Another common approach for synthesizing research findings has been the gathering and formal counting of the number of statistically significant and non-significant results. This approach, also known as the "vote-counting" technique, is often coupled with statistical tests to compare the frequency or proportion of occurrence between studies with significant and non-significant outcomes to decide whether there is support for or against a particular hypothesis or whether a treatment effect exists or not (Hedges & Olkin,

1985; Cooper & Hedges, 1994; Gurevitch & Hedges, 2001). The rather simple and straightforward characteristics of this technique have made it a quite successful and intuitive tool that has been widely used in research syntheses in plant sciences and in many other areas. Nevertheless, this technique has low statistic power making vote counts results seriously biased toward finding no effect (e.g., Gurevitch & Hedges, 1999; 2001, Borenstein *et al.*, 2009). Regrettably, the problems associated with vote-count procedures do not improve by including a greater number of studies in the research synthesis; on the contrary, the power of the test actually decreases as more studies are included, tending to zero as the number of studies become large (Hedges & Olkin, 1985; Cooper & Hedges, 1994; Borenstein *et al.*, 2009).

On the whole, both narrative and vote-counting reviews suffer from serious biases as a result of exclusively basing conclusions about the existence and frequency of a particular phenomenon on the statistical significance of the outcomes (based on probability levels), without considering sample size and statistical power of each of the studies included in their synthesis (Hedges & Olkin, 1985; Cooper & Hedges, 1994, Borenstein *et al.*, 2009). The significance level of a study is a function not only of the magnitude of the effect, but also of its sample size. Given two different studies with a hypothetical identical treatment effect magnitude, the study with small sample sizes is less likely to show statistical significant results (i.e., it has lower power) compared to a study with large sample sizes (Cooper & Hedges, 1994; Gurevitch & Hedges, 2001). Therefore, by simply counting up or describing the number of significant results is not a reliable indication of whether an effect is real, how important and frequent it is, or under what circumstances the effect exists (Hedges & Olkin, 1985; Gurevitch & Hedges, 2001, Borenstein *et al.*, 2009).

Finally, quantitative reviews, in contrast to the qualitative methods described, offer a different perspective on the results of independent studies. Instead of providing a definite demonstration on a particular phenomenon, quantitative reviews, widely known as “meta-analyses”, treat individual published results as if they were subjected to sampling uncertainty, just like primary studies do with their own raw data (Cooper & Hedges, 1994; Gurevitch & Hedges, 1999; 2001, Borenstein *et al.*,

2009). Quantitative syntheses estimate not only the magnitude and direction of the effects but also how variable the effects are among individual studies. This kind of synthesis has proven to be a powerful tool to establish generalizations in many research areas as diverse as the medical, physical, behavioral, and environmental sciences, allowing to answer a wider variety of questions and providing a considerable advance in scientific rigour over traditional narrative or vote-counting reviews (Hedges & Olkin, 1985; Arnqvist & Wooster, 1995; Gurevitch & Hedges, 1999; 2001, Borenstein *et al.*, 2009).

A BRIEF STORY OF META-ANALYSIS

Meta-analysis is a relatively recent tool. It was in the mid 1970's when Gene Glass (1976) first coined the term “meta-analysis”. Soon after, him and Mary Smith developed and conducted the first meta-analytical study to determine whether psychotherapy produced beneficial results (Smith & Glass, 1977). The name meta-analysis was used as a simple way of speaking of the “the analysis of analyses” (Glass, 1976). At the time of its development, the volume of research in many fields was growing at such a rate that traditional narrative or qualitative approaches to summarizing and integrating research were beginning to break down. By then, as Glass himself revealed, he began to conceive a new conceptual framework as a result of his growing skepticism of statistical significance testing; his notion that all research was imperfect in one respect or another (i.e., there are no “perfectly valid” studies), and his beginning to question a taken-for-granted assumption that we progress toward truth by doing what everyone commonly refers to as “studies” (Glass, 2000).

Throughout these years, meta-analysis has grown rapidly from a preoccupation of a few statisticians working on problems of research integration in education and psychotherapy to a vast academic enterprise (Hunt, 1997; Glass, 2000). Its popularity in the social sciences and education is irrelevant compared to its influence in medicine (see Egger & Smith 1997 and references therein). A fast web search using the keyword “meta-analysis” yields c.a. 61,000 different published document types from 1977 to the present; astonishingly, nearly

75% of these records were produced in the medical sciences. Today, systematic quantitative reviews of healthcare research are being coordinated by an international network of individuals and institutions, known as the Cochrane Collaboration, which has become a cornerstone of the evidence on which medical research is based (see <http://www.cochrane.org>). In this regard, meta-analyses have served in several instances to guide social and healthcare policies, particularly in the United States of America (Hunt, 1997).

The use of meta-analysis in plant sciences is still in its infancy. Comparatively to the medical or social sciences, its impact within the many research areas comprised in plant sciences is still insignificant. However, as we will see below, some specific topics of research in plant sciences have greatly benefited from this type of quantitative synthesis, where interesting trends and overall response patterns have been found.

HOW DOES META - ANALYSIS WORK?

Here, I briefly depict some of the most important concepts and procedures to conduct a meta-analysis. I do not provide a detailed account on the equations of meta-analysis, which can be found in detail elsewhere (Hedges & Olkin, 1985; Cooper & Hedges, 1994; Gurevitch & Hedges, 2001; Borenstein *et al.*, 2009).

Just like any other statistical analysis, meta-analysis requires data and statistical models to analyze data. The scope and criteria of gathering the published data is crucial. As in any type of review, the quality and nature of the data collected will affect the extent to which we can extrapolate the conclusions drawn. The quantitative nature of meta-analytical reviews brings certain extra issues into sharper focus when gathering published data, which I will not cover in this article. Specific suggestions on collecting and handling data can be found in Cooper & Hedges (1994), Gurevitch & Hedges (1999), and Borenstein *et al.* (2009). After a complete list of studies is collected, numerical data from each study is extracted. Such numerical data are typically mean values, standard deviations, and sample sizes of control and experimental groups, or regression and correlation coefficients and sample sizes. These measures can be obtained from text or

tables and also from figures by digitalizing them.

The data in a meta-analysis takes the form of standardized metrics of the magnitude and sampling variance of effects. Results of each study are represented by an "effect size", which will be chosen to reflect either (a) the differences between experimental and control groups (i.e., categorical approach) or (b) the degree of relationship between the independent and dependent variables (i.e., correlational approach) depending on the characteristic of the data and the hypotheses being tested (Cooper & Hedges 1994; Gurevitch & Hedges, 2001). There are several different metrics of effect size used in meta-analysis that are calculated from the extracted data of independent studies. Some of the most commonly used include the standardized mean difference, d ; Pearson's correlation coefficient, r (analyzed as the z -transform); the log response ratio, lr (the natural log of the ratio between the mean of the experimental and control group; Hedges *et al.* 1999), and the odds ratio for categorical outcomes (Cooper & Hedges, 1994). The choice of the appropriate metric of effect size to be used underlies the entire meta-analysis and is therefore a decision of critical importance (Osenberg *et al.*, 1999). Details on the calculations and properties of effect sizes can be found in Cooper & Hedges (1994), and Borenstein *et al.* (2009).

Once we obtain a measure of the effect size from each study, the analysis can be performed. Both continuous and categorical approaches for meta-analysis have been developed and can rely upon fixed, random, or mixed models, considering the characteristics of meta-analytic data. Both approaches depend on weighting by the inverse of the sampling variance of each data point (e.g., each effect size). By doing this, meta-analysis takes into account the fact that all studies are not equally reliable, by giving estimates of each study different weights based on their sample size (Osenberg *et al.*, 1999; Gurevitch & Hedges, 2001; Borenstein *et al.*, 2009). Such weighting makes meta-analysis procedures not subject to the problems of vote-counting (Gurevitch & Hedges, 1999; 2001). An updated and complete description of statistical models for meta-analysis can be found in the recent book of Borenstein *et al.* (2009).

Statistical analyses of effect sizes, using tests of significance with parametric or randomization techniques, can then be constructed to estimate the average magnitude of the effect across all studies, to

test whether that effect is significantly different from zero, and to examine potentially causative differences in the effect among studies. In other words, we can answer a variety of questions such as: how large is the effect overall? Is it positive or negative? Is it reliably different than zero? Are the results consistent across studies? If the answer to this last question is negative, are there differences in the magnitude of effect among biologically meaningful categories of studies (life-history traits, trophic levels, natural systems, etc.)?

PUBLICATION BIAS: THE “FILE-DRAWER PROBLEM”

Literature reviews may be subjected to the possibility of publication bias (Müller & Jennions, 2001; Jennions & Müller, 2002), which refers to the potential tendency of scientific journals to publish more frequently studies with statistically significant results (Rosenthal, 1979; Cooper & Hedges, 1994). The lack of non-significant published studies has been termed the “file-drawer problem”, to represent the unknown number of studies that will remain in the researchers’ drawers and will never appear reported in printed format (Rosenthal, 1979). If severe publication bias exists in a quantitative review it can substantially overestimate the real treatment effect observed. This intrinsic problem of any research synthesis can be accounted for in the meta-analytical framework in several ways (Cooper & Hedges, 1994; Rosenberg, 2005).

One type of methods for detecting publication bias is based on the examination of the relation between standard error and effect size (Gurevitch & Hedges, 1999). These include graphical methods (the “Funnel plot”, cf. Palmer, 1999) and formal tests of correlation between standard error (or sample size) and effect size (Begg, 1994). The principle underlying these methods is that, since the effect parameters are consistent across studies, any relation with sample size (or standard error) must reflect publication bias (Gurevitch & Hedges, 1999).

A simpler and quick method to estimate whether publication bias is likely to be a problem is the calculation of a fail-safe number (Rosenberg, 2005). This fail-safe number specifies the number of non-significant, unpublished (or missing) studies that would need to be added to a meta-analysis to nullify its overall effect size (Rosenberg, 2005). If this

number is larger relative to the number of observed studies (larger than $5n + 10$, where n is the number of studies), then one can be fairly confident in the summary conclusions (i.e., the results are robust regardless of publication bias; Rosenthal, 1979; Rosenberg, 2005).

AN EXAMPLE OF HISTORICAL AND CURRENT RESEARCH SYNTHESIS APPROACHES

The loss and fragmentation of natural habitats are ubiquitous phenomena across the earth and their effects on plant population dynamics have been consistently studied over the last two decades (Hobbs & Yates, 2003). Specifically, studies have largely focused on animal pollination and plant reproduction in fragmented habitats. Theoretical arguments suggested that the compatibility systems and the degree of pollination specialization of plants might be useful traits to determine reproductive vulnerability to habitat fragmentation (e.g., Bond, 1994; Hobbs & Yates, 2003). Self-incompatible plants (SI) are obligate outbreeders and, thus, need both conspecific individuals and pollination vectors to reproduce successfully. Therefore, SI plants are expected to be reproductively more vulnerable to fragmentation effects than self-compatible plants (SC), which have the advantage of producing seeds with both outcross and self pollen, making them only facultative dependent on pollinators (Richards, 1997). On the other hand, pollination-specialist plants (S), which are typically pollinated by one or a few taxonomically related pollinator species, may also be reproductively more vulnerable to fragmentation compared to pollination-generalist plants (G), which interact with a higher number and diversity of pollinator species and can therefore buffer the potential loss of some of their pollinators in fragmented habitats (e.g., Morris, 2003).

In 2002, Aizen and colleagues reviewed the published literature to seek whether compatibility systems and pollination specialization in plants effectively defined differential vulnerability to fragmentation. They gathered information on the effects of habitat fragmentation on reproductive success and pollination of 46 plant species and conducted a qualitative review by using the vote-counting approach. They statistically compared the relative frequency of species showing negative

fragmentation effects between SI and SC species (Fig. 1A) and between G and S species (Fig. 1B) as denoted by the significant *P* values of the original studies. They found no significant differences between these frequencies (Fig. 1A, B), implying that no generalization could be made based on these ecological traits. Therefore, they concluded that neither the compatibility systems nor the degree of pollination specialization determined differential reproductive vulnerability to habitat fragmentation (Aizen *et al.*, 2002).

A few years later, Aguilar *et al.* (2006) re-tested these two hypotheses by means of a quantitative review analysis of the published literature. We were able to gather information about the effects of fragmentation on pollination and reproductive success from 50 and 89 unique plant species, respectively. A few of these species were studied twice by different authors in different systems, so we conducted meta-analyses on 52 and 93 data points to test effects on pollination and reproduction, respectively. From each published article, we obtained mean values, standard deviations and sample sizes of reproductive output (fruit-set or seed-set) and pollination (visit frequency, pollen loads or pollen tubes counts) variables in both conditions: fragmented habitats and continuous forests. For each species, we calculated the standardized unbiased mean difference (Hedge's *d*) between pollination and reproduction in continuous and fragmented habitats as a measure of effect size. We were able to calculate the overall weighted mean effect size of habitat fragmentation on pollination and plant reproduction. Results were quite unambiguous: fragmentation strongly and negatively affects the pollination process and the reproductive output of plants (Fig. 2). In contrast to the previous review, we found significant differences in the effects of fragmentation on pollination and reproduction between SI and SC species (Fig. 2A, B). In agreement to the hypothesis stated, SI species were more strongly negatively affected by fragmentation than SC species. No significant difference was observed between G and S species (Fig. 2A, B), which disagrees with the original hypothesis in relation to this ecological trait (see Ashworth *et al.*, 2004 for a theoretical elucidation on this response pattern), coinciding with the results of the previous review.

This example illustrates how vote-counting and meta-analysis differ in their statistical power to find

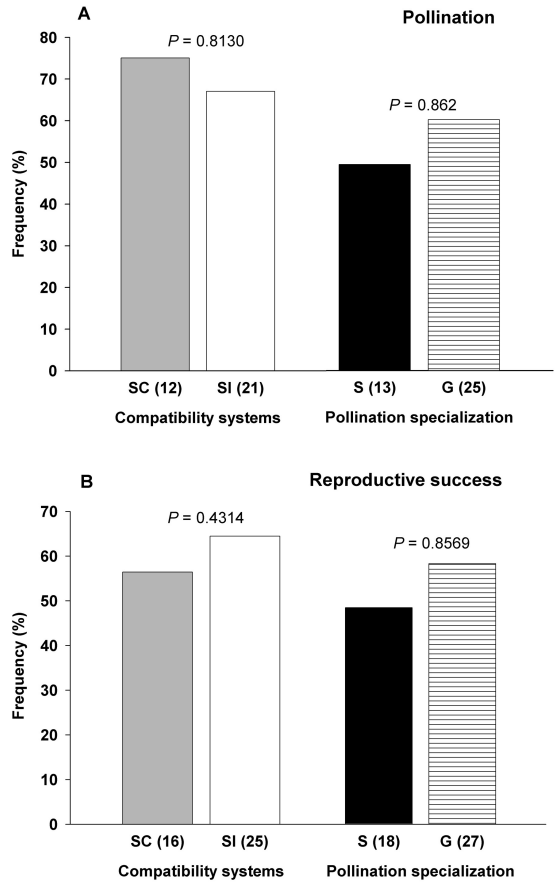


Fig. 1. Relative frequency of self-compatible (SC) vs self-incompatible (SI) and pollination-generalist (G) vs pollination-specialist (S) plant species showing negative fragmentation effects on (A) pollination and (B) reproductive success. Sample sizes are given in parentheses. *P*-values associated with a one tailed Fisher's Exact test. Re-drawn from Aizen *et al.* (2002).

differences in treatment effects, which is particularly important to arrive to valid generalizations.

META - ANALYSES IN PLANT SCIENCES: STATE OF THE ART

To determine the published existent records of meta-analyses in plant sciences I conducted a web search in the *Scopus* database (www.scopus.com), which contains published material in all areas of science from 1945 to 2009. I used the following keyword combination: meta-analysis AND plant* OR veget*, and limited the time span from 1970 to the present. The results of this search were followed

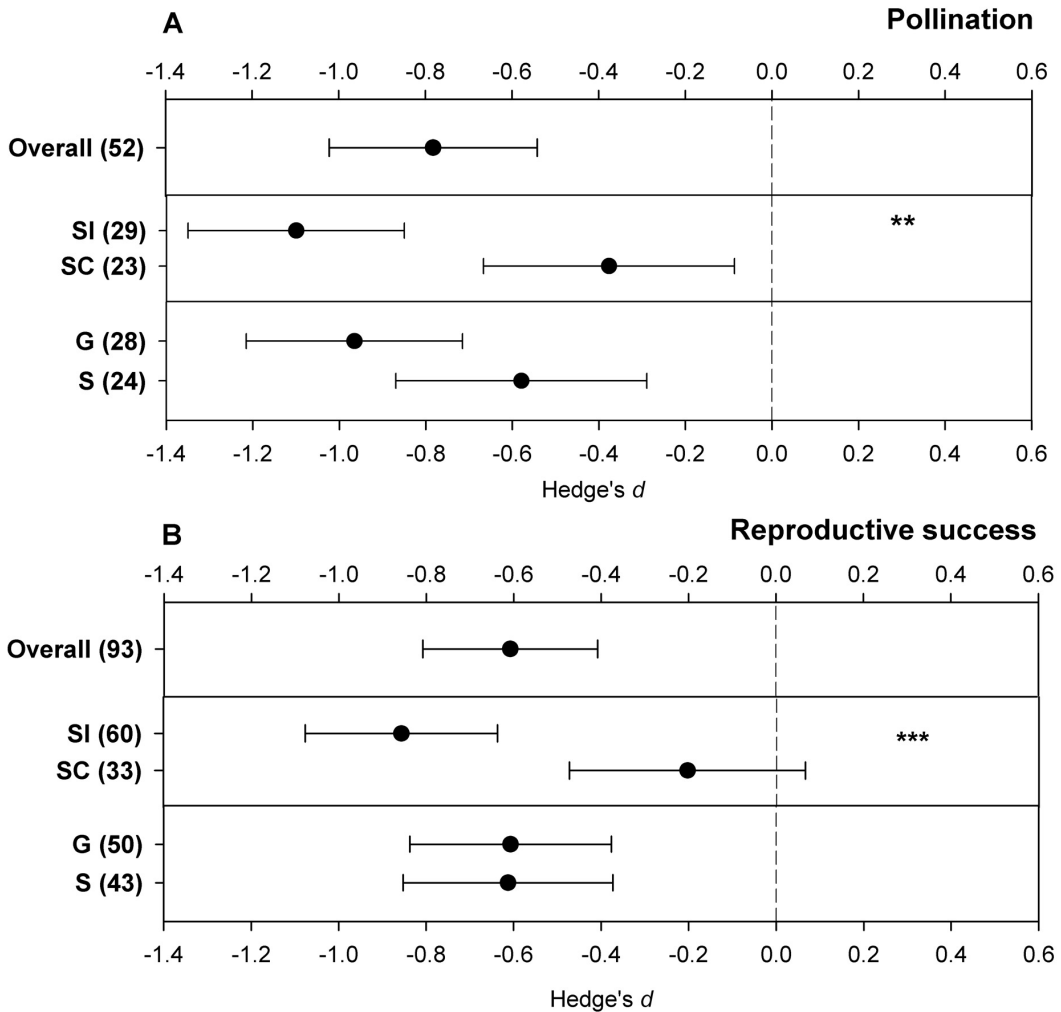


Fig. 2. Weighted-mean effect sizes and 95% bias-corrected confidence intervals of habitat fragmentation on (A) pollination and (B) plant reproductive success. Sample sizes are given in parentheses; dotted line shows effect size Hedge's $d = 0$. SI: self-incompatible, SC: self-compatible, G: generalist, S: specialist. ** $P < 0.01$, *** $P < 0.001$. Redrawn from Aguilar *et al.* (2006)

by a close examination of each record. I restricted the search to review articles synthesizing research in any field involving plants (vascular and non-vascular) as study targets.

The keyword combination yielded 656 review articles in the web search. More than 50% of these records (350) corresponded to quantitative syntheses in very diverse fields of the Medical sciences from Phytomedicine, Pharmacology, Nutrition, and Toxicology to Oncology, Immunology, Neurology, and Psychology. I did not consider these records to represent true quantitative syntheses in plant sciences, because in all of these

cases the synthesis was aimed to generalize on the therapeutic effects of specific plant species, or active components extracted from them, on different aspects of human health.

I found 306 records that were effectively review articles synthesizing research in different fields of plant sciences (i.e., plants as study targets). The first two meta-analyses in plant sciences were conducted quite recently; only 13 years ago, Britten (1996) made a quantitative generalization about the association between multilocus heterozygosity and fitness, and Curtis (1996) synthesized research on leaf gas exchange and nitrogen in trees grown under

elevated carbon dioxide. From that year on the growth of quantitative review articles in plant sciences grew exponentially, with a top record of 60 articles in 2007 (Fig. 3).

I arbitrarily classified all these articles in 11 subject areas or disciplines of research after reading the title, abstract and keywords. Notably, the majority of them were conducted in the field of Ecology (Fig. 4). Within this broad area I included review articles in Evolutionary Ecology, Molecular Ecology, Conservation Biology, Agroecology, and also a couple of articles on Biogeochemical Ecology. The reviews conducted in Ecophysiology were better self-defined and well represented, so they are shown separately. Reviews in Genetics and Agricultural sciences followed in importance, with 26 and 20 syntheses in each area, respectively (Fig. 4). The rest of the areas showed less than 20 quantitative syntheses on several specific thematic subjects.

It is expected that quantitative generalizations will be mainly conducted in those specific themes or scientific areas considered “hot” or of broad circumstantial interest to the scientific or public community. In some other cases, however, quantitative syntheses are carried out to gain new knowledge about old questions or to re-test classical hypotheses. The level of interest in a particular subject it is usually noticed in a significant increase of published literature. Not surprisingly, ongoing global environmental change (both climatic and land use change) as a consequence of human activities, have prompted an enormous amount of research on different aspects of plant sciences within ecological disciplines. A particular good example is the extensive research on the effects of elevated CO₂ on different physiological and ecological processes of plants, such as growth, leaf chemistry or plant-animal interactions and crop yields, to name a few (e.g., Curtis, 1996; Poorter & Navas, 2003; Ainsworth & Long, 2005). Within all the review papers gathered, more than 10% of them dealt with increased CO₂ effects. Particular important generalizations have been obtained in this topic (summarized in Ainsworth *et al.*, 2007), such as a significant increase in the growth of C₄ species to elevated CO₂ contrary to expectations (Poorter & Navas, 2003) or that crop yields increase was much less than previously anticipated under elevated CO₂ (Ainsworth & Long, 2005). Different aspects of land use change effects on plants (reproduction,

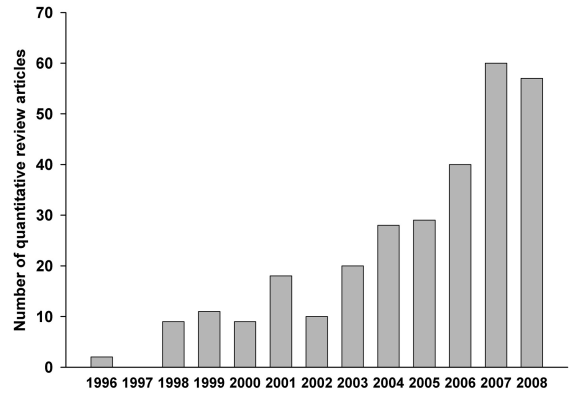


Fig. 3. Temporal trend of published quantitative review articles in plant sciences.

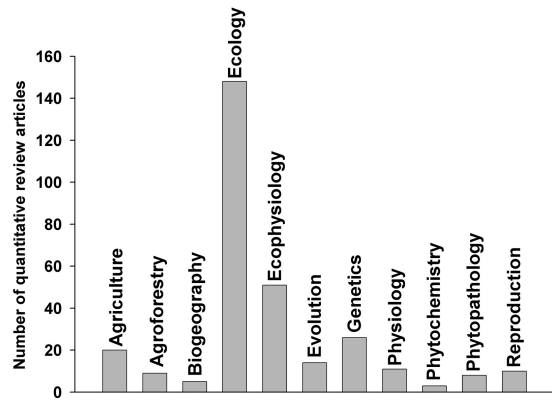


Fig. 4. Incidence of quantitative reviews in different subject areas within plant sciences.

genetic diversity, abundance, richness, etc.) have also been reviewed consistently. In two different meta-analyses we concluded that obligate outcrossing plants, imposed either by temporal (dichogamy) or spatial (herkogamy) barriers or by genetically based self-incompatibility, are significantly more affected in terms of reproduction and genetic diversity by the loss and fragmentation of habitats (Aguilar *et al.* 2006, 2008). Thus, outcrossing self-incompatible species are particularly prone to suffer demographic collapse in fragmented habitats as a consequence of both ecological and genetic factors.

As observed, the increased research interest on different anthropogenic effects on plants may be one of the reasons of finding such a great number of

quantitative reviews within the ecological and ecophysiological areas. There are many areas of research in plant sciences that are only beginning to conduct quantitative reviews. Some long forgotten questions or sleeping theories may well regain interest if approached in a meta-analytical context.

FUTURE PROSPECTS

While current meta-analytical techniques are an important tool to synthesize research in plant sciences, they still have a long way to go through. Some criticisms have been raised against quantitative review approaches in general (e.g., Eysenck, 1978; Bailer, 1997; Egger & Davey, 1998). Several of the flaws criticized are in fact common to any other type of review (e.g., file-drawer problem, research bias, criteria of inclusion, quality of primary papers, etc.), while some others have been based on misunderstandings of meta-analytical techniques (reviewed in Borenstein *et al.*, 2009). A few other flaws, instead, reflect problems in the way that meta-analysis is used in different fields (e.g., inappropriate use of effect sizes), rather than problems in the method itself (Borenstein *et al.*, 2009). In this regard, the medical sciences have experienced an important development and sophistication in meta-analytical techniques whereas in plant sciences there is still much to work on. The challenges ahead involve the definition of new metrics for measuring effects based on particular needs depending on the kind of data and questions raised within the different areas of research (ecology, physiology, genetics, etc.). Accordingly, additional efforts should be made to elaborate specific meta-analytical models and more advanced statistical approaches (e.g., two-way analysis of variance, analysis of covariance, multivariate analysis, and hierarchical nested analyses) and the use of advanced statistical inference methods, such as maximum likelihood and Bayesian meta-analysis (Ainsworth *et al.*, 2007). Ecologists in particular have already begun to seize some of these challenges with great success (e.g., Gurevitch & Hedges 1999; Hedges *et al.* 1999; Osenberg *et al.* 1999). These needs will force meta-analysis in plant sciences to move from lectures of advanced statistical courses into mainstream teaching and research.

On the ambition of finding general patterns in

plant sciences across species, communities, or ecosystems, meta-analyses use the results of independent studies, which many times are the results of individual species (i.e., the replicate data points are the species). For this reason, one important aspect to incorporate in meta-analysis is the effect of phylogenetic relationships within the data in order to account for the lack of independence resulting from shared evolutionary history. Conventional meta-analyses do not take into account such lack of independence across studies. Some researchers have attempted to bridge this gap in meta-analysis through phylogenetic simulations (Ashton & Feldman, 2003; Verdú & Traveset, 2004; 2005). Recently, Adams (2008) developed a more specific model for phylogenetic meta-analysis, which allows summarizing data across studies with their evolutionary history explicitly incorporated. This approach takes advantage of common aspects of linear statistical models used by both, meta-analysis and the phylogenetic comparative method, allowing them to be analytically combined (Adams, 2008). This new method, which shows the increasing research on meta-analytical tools, will prove to be extremely useful in future meta-analyses in plant sciences.

Meta-analysis was created out of the need to arrive to generalizations by extracting numerical information from the cryptic records of inferential data analyses in reports of research in journals and other printed sources (Glass, 1976, 2000). The growth and scope of future meta-analyses will greatly depend on and benefit from the full availability and completeness of reported data in primary research studies. To achieve this, researches and journals need to standardize the quality and procedures of reporting summary data in published articles, making them readily available for meta-analytical use. More ambitiously, in the future we would wish to be able to access to archives of original raw data of primary published studies allowing the construction of complex data landscapes that depict the relationships among independent, dependent and mediating variables (Glass, 2000). This ultimate enterprise will require a joint effort by implementing network collaboration among researchers around the world to make such raw data available. Clearly, such exiting possibility requires an entirely new conception on the problem of how we integrate research findings in science.

ACKNOWLEDGEMENTS

I truly appreciate the thorough revisions of Drs. Martin Videla & Lorena Ashworth. Their comments and constructive criticisms substantially improved the final version of this Ms. I was supported by the Agencia Nacional de Promoción Científica y Técnica (PICT 00132) and CONICET (PIP 00790). I am currently a member of the Carrera de Investigador from CONICET.

BIBLIOGRAPHY

- ADAMS, D. C. 2008. Phylogenetic meta-analysis. *Evolution* 62: 567-572.
- AGUILAR, R., L. ASHWORTH, L. GALETTO & M.A. AIZEN. 2006. Plant reproductive susceptibility to habitat fragmentation: review and synthesis through a meta-analysis. *Ecol. Lett.* 9: 968-980.
- AGUILAR, R., M. QUESADA, L. ASHWORTH, Y. HERRERIAS-DIEGO & J. LOBO. 2008. Genetic consequences of habitat fragmentation in plant populations: susceptible signals in plant traits and methodological approaches. *Mol. Ecol.* 17: 5177-5188.
- AINSWORTH, E. A. & S. P. LONG. 2005. What have we learned from 15 years of free-air CO₂ enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy properties and plant production to rising CO₂. *New Phytol.* 165: 351-372.
- AINSWORTH, E. A., M.S. ROSENBERG & X. WANG. 2007. Meta-analysis: the past, present and future. *New Phytol.* 176: 742-745.
- AIZEN, M. A., L. ASHWORTH & L. GALETTO. 2002. Reproductive success in fragmented habitats: do compatibility systems and pollination specialization matter? *J. Veg. Sci.* 13: 885-892.
- ARNQVIST, G. & D. WOOSTER, D. 1995. Meta-analysis: synthesizing research findings in ecology and evolution. *Trends Ecol. Evol.* 10: 236-240.
- ASHTON, K. G. & C. R. FELDMAN. 2003. Bergmann's rule in nonavian reptiles: turtles follow it, lizards and snakes reverse it. *Evolution* 57: 1151-1163.
- ASHWORTH, L., R. AGUILAR, L. GALETTO & M. A. AIZEN. 2004. Why do pollination generalist and specialist plant species show similar reproductive susceptibility to habitat fragmentation? *J. Ecol.* 92: 717-719.
- BAILAR, J. C. 1997. The promise and problems of meta-analysis. *New Engl. J. Med.* 337: 559-561.
- BEGG, C. B. 1994. Publication bias. In: COOPER, H. & L. V. HEDGES (eds.), *The handbook of research synthesis*, pp. 399-409. Russell Sage Foundation, New York.
- BOND, W. J. 1994. Do mutualisms matter? Assessing the impact of pollinator and disperser disruption on plant extinction. *Phil. Trans. R. Soc. Lond. B.* 344: 83-90.
- BORENSTEIN, M., L. V. HEDGES, J. P. T. HIGGINS & H. R., ROTHSTEIN, 2009. Introduction to meta-analysis. John Wiley & Sons Ltd. London, UK.
- BRITTEN, H. B. 1996. Meta-analyses of the association between multilocus heterozygosity and fitness. *Evolution* 50: 2158-2164.
- BUNGE, M. 1988. La ciencia, su método y su filosofía. Ediciones Siglo Veinte, Buenos Aires.
- COOPER, H. & L. V. HEDGES. 1994. *The handbook of research synthesis*. The Russell Sage Foundation, New York.
- CURTIS, P. S. 1996. A meta-analysis of leaf gas exchange and nitrogen in trees growing under elevated carbon dioxide. *Plant, Cell & Environm.* 19: 127-137.
- EGGER, M & G. DAVEY SMITH. 1997. Meta-analysis: Potentials and promise. *Brit. Med. J.* 315: 1371-1374.
- EGGER, M & G. DAVEY SMITH. 1998. Bias in location and selection of studies. *Brit. Med. J.* 316: 61-66.
- EYSENCK, H. J. 1978. An exercise in mega-silliness. *Am. Psychol.* 33: 517.
- GLASS, G. V. 1976. Primary, secondary, and meta-analysis of research. *Educational Researcher.* 5: 3-8.
- GLASS, G. V. 2000. Meta-analysis at 25. Tempe, AZ. Arizona State University. Retrieved March 15, 2009 from <http://glass.ed.asu.edu/gene/>
- GUREVITCH, J. & L. V. HEDGES. 1999. Statistical issues in ecological meta-analyses. *Ecology* 80: 1142-1149.
- GUREVITCH, J. & L. V. HEDGES. 2001. Meta-analysis: combining the results of independent experiments. In: Scheiner, S. M. & Gurevitch J. (eds.), *Design and Analysis of Ecological Experiments*, pp. 378-398. 2nd edition. Oxford University Press, NY.
- HEDGES, L. V. & I. OLKIN. 1985. *Statistical methods for Meta-analysis*. Academic Press, New York.
- HEDGES, L. V., J. GUREVITCH & P. CURTIS. 1999. The meta-analysis of response ratios in ecology. *Ecology* 80: 1150-1156.
- HOBBS, R. J. & C. J. YATES. 2003. Impacts of ecosystem fragmentation on plant populations: generalizing the idiosyncratic. *Austr. J. Bot.* 51: 471-488.
- HUNT, M. 1997. *How science takes stock: The story of meta-analysis*. NY: Russell Sage Foundation.
- HUNTER, J. E. & F. L. SCHMIDT. 1990. *Methods of meta-analysis: correcting error and bias in research findings*. SAGE Publications, Newbury Park.
- JENNIONS, M. D. & A. P. MÜLLER. 2002. Publication bias in ecology and evolution: an empirical method using the "trim and fill" method. *Biol. Rev.* 77: 211-222.
- KLIMOVSKY, G. 2001. Las desventuras del conocimiento científico. A.Z. Editora, Buenos Aires.
- MÜLLER, M. D. & A. P. JENNIONS. 2001. Testing and adjusting for publication bias. *Trends Ecol. Evol.* 16: 580-586.
- MORRIS, W. F. 2003. Which mutualists are most essential?: buffering of plant reproduction against the extinction of pollinators. In: KAREIVA, P. & S. LEVIN (eds.), *The Importance of Species: Perspectives on Expendability and Triage*, pp. 260-280. Princeton University Press, Princeton.

- OSENBERG, C. W., O. SARNELLE, S. D. COOPER, & R. D. HOLT. 1999. Resolving ecological questions through meta-analysis: goals, metrics, and models. *Ecology* 80: 1105–1117.
- PALMER, A.R. 1999. Detecting publication bias in meta-analyses: a case study of fluctuating asymmetry and sexual selection. *Amer. Nat.* 154: 220-233.
- POORTER, H. & M. L. NAVAS. 2003. Plant growth and competition at elevated CO₂: on winners, losers and functional groups. *New Phytol.* 157: 175-198.
- RICHARDS, A. J. 1997. *Plant breeding systems*. 2nd ed. Chapman and Hall, London.
- ROSENBERG, M.S. 2005. The file-drawer problem revisited: a general weighted method for calculating fail-safe numbers in meta-analysis. *Evolution* 59: 464-468.
- ROSENTHAL, R. 1979. The “file drawer problem” and tolerance for null results. *Psychol. Bull.* 86: 638-641.
- SMITH, M. L. & G. V. GLASS. 1977. Meta-analysis of psychotherapy outcome studies. *Amer. Psychol.* 32: 752-60.
- VERDÚ, M. & A. TRAVESET. 2004. Bridging meta-analysis and the comparative method: a test of seed size effect on germination after frugivores` gut passage. *Oecologia* 138: 414-418.
- VERDÚ, M. & A. TRAVESET. 2005. Early emergence enhances plant fitness: a phylogenetically controlled meta-analysis. *Ecology* 86: 1385-1394.

Recibido el 27 de febrero de 2009, aceptado el 29 de marzo de 2009.