PRACTICE-ORIENTED CONTROVERSIES AND BORROWED EPISTEMIC CREDIBILITY IN CURRENT EVOLUTIONARY BIOLOGY: PHYLOGEOGRAPHY AS A CASE STUDY.

ALFONSO ARROYO-SANTOS[1,2]
MARK E. OLSON[3]
FRANCISCO VERGARA-SILVA[3]

1. FACULTAD DE FILOSOFÍA Y LETRAS, UNAM
CIRCUITO INTERIOR CIUDAD UNIVERSITARIA,
S/N 04510, MÉXICO, DF.
agripas@gmail.com
TEL/FAX: (52) (55) 5622 7239

2. CENTRO DE INFORMACIÓN GEOPROSPECTIVA
BERLÍN 29A, DEL CARMEN COYOACÁN
04100, MÉXICO DF.

3. INSTITUTO DE BIOLOGÍA, UNAM
CIRCUITO EXTERIOR CIUDAD UNIVERSITARIA
S/N 04510, MÉXICO DF.

Methodological controversies are an important but often neglected issue in the philosophy of science. Because experimental results often cannot settle controversies, other elements must be incorporated to debates. We introduce the notion of borrowed epistemic credibility to better understand the role that non-empirical elements play in such controversies, illustrating our proposal with a recent controversy in phylogeography. Our analysis shows how scientific controversies that spring from disagreements about methodological issues potentially involve deeper debates regarding what constitutes 'good science' in general, and suggests the re-examination of more general issues, such as the nature of inference, rationality, or objectivity.

1. Introduction

Scientists are increasingly turning to computer simulations and statistical inferences in their research efforts, as entire journals such as *Bioinformatics* attest. Because these approaches are so methods-intensive, when disputes arise, theoretical considerations are not necessarily central to their resolution. For example, the methodology-heavy nature of contemporary scientific research has been noticed by Eric Winsberg, who claims that in computer simulation studies, any underlying theory is just one of many factors involved in interpreting the outcome of simulations and, as a consequence, theory may bear no direct substantive relationship to the knowledge generated by simulations (see Winsberg 2006, 2009, 2010). With regard to statistics, many different statistical approaches (e.g. t-tests, principal components, Bayesian networks, and so on) can be used to address any given question. Because different individuals or communities of investigators often have contrasting perceptions of the relative merits of different approaches, tradition or convention tends to guide the choosing of a particular method rather than theory (e.g. Smith 2009). Given that more and more scientific research is built on simulations and statistical analysis, theory takes a back seat to methodological issues in more and more scientific debates.

After years of discussion regarding ‘the practice turn’ in the philosophy of science, increasing interest from philosophers of biology in practical controversies in current evolutionary biology might have been expected. However, philosophers of biology working on evolutionary matters seemingly continue to follow Ernan McMullin in considering theoretical disputes "the commonest source of controversy in science" (McMullin 1987, p. 66 [e.g. Dietrich and Skipper 2007]). Studies of controversies at the theoretical level are certainly important. However, emphasis on theories obscures the fact that many important current conflicts in evolutionary biology are linked to the generation and interpretation of results obtained by means of modeling tools and similar resources. In short, current controversies often have to do not with theory but with the way in which specialists actually measure phenomena, describe processes, and quantitatively address their explanations.

The study of scientific practice can do much to illuminate philosophical issues, as we illustrate with our case study here. We analyze a recent, heated controversy in a high-profile subdiscipline of evolutionary biology called phylogeography. This field, which has been described as a “bridge linking the study of micro and macroevolutionary processes” (Bermingham and Moritz 1998, p. 367), currently provides an outstanding example of a biological controversy in which statistical inferences and computer simulations have taken center stage. Our analysis of the controversy illustrates how theory plays little part in the debate. Further, our analysis suggests that disagreements regarding the value of statistical methods are only the surface of a deeper conflict related to many issues of importance to philosophy of science such as the nature of scientific inferences, the justification of results, or the generation of knowledge. In other words, our focus on practice highlights a scientific debate regarding the nature of science.

To support the view presented here, we introduce the notion of *borrowed epistemic credibility* (BEC). BEC describes the situation in which, unable to justify a particular claim on empirical grounds, a given researcher (a) looks to justify her stance through an appeal to a set of non-empirical values that correspond to proper scientific standards in her field, by (b) invoking similarities in her research to practices, results or methodologies with a well-established set of values. Once the claim for similarity is granted, then we say it has successfully borrowed epistemic credibility, and consequently, the defending claims should be held as correct because they conform to what members of her community consider as 'good scientific practice'[1]. We claim that our notion of BEC illuminates how the methodological debate is only the visible part of a disciplinary conflict where two competing groups disagree on the nature of their discipline.

Today, it is generally agreed that non-empirical values play an important role in science but the question still remains as to how they do so. Addressing the *how* question will allow us not only to better understand scientific practice and the generation of scientific knowledge, but also to better understand what it is that scientists take for objective, valid, or justified, to name three examples (see for example, Gervais 2013, or Douglas 2009). We believe BEC could be useful to understand how other scientific controversies that spring from disagreements about methodological issues potentially involve deeper debates regarding what constitutes 'acceptable' or 'good science' in general, and is also a contribution to understanding the role of non-epistemic values in science.

In the first part of this paper we introduce briefly the field of phylogeography. We direct interested readers to Arroyo-Santos, Olson, and Vergara-Silva (2013) for a more detailed discussion of phylogeography and the controversy in question. Our interest here is mainly to present the root of the controversy and to show how it is not being settled with empirical

[1] In our usage, 'good scientific practice' means the set of norms, values, practices that scientists in a particular area think scientific work should possess.

arguments. Because empirical arguments do not suffice, we argue that contending sides appeal to what we call epistemic credibility. In the following section we develop the notion of borrowed epistemic credibility and show how this seemingly methods-oriented controversy is in fact part of a larger debate on the nature of phylogeography, evolutionary biology, and science itself.

2. A brief summary of phylogeography

Phylogeography was born in the late 1980s as an attempt to unify the fields of phylogenetics and population genetics, in an explicitly biogeographical context (Avise et al. 1987; Avise 2000, 2009; Bermingham and Moritz 1998; Hickerson et al., 2010; Riddle 2008). Phylogeographic studies take thinking in terms of phylogeny, the branching relationships between species, and apply it within species. They reconstruct within-species "phylogenies" and overlay them on geography to infer the processes that have shaped the current distribution of genetic variation. The association between the distribution of genetic variation with respect to geography allows researchers to assess possible evolutionary scenarios regarding the role of 'recurrent processess' such as continual genetic interchange, or of 'unique events' such as the rising of a mountain range.

Phylogeography has become a field in its own right. In 1998, Avise wrote an article to celebrate the 10th anniversary of the paper that started the field. Avise (1998) documented "more than 130 papers that had employed 'phylogeography' in the title or as an index word", adding that "they represent only the tip of the iceberg because numerous additional studies have dealt with the topic implicitly although not by name" (Avise 1998: 371). Similarly, phylogeographer L. Lacey Knowles recently stated that "the unwavering popularity of phylogeography is indisputable" (Knowles 2009, p. 595) and estimated that 4370 papers in which the word 'phylogeography' was explicitly used had been published up to 2008. The importance and wide acceptance of

phylogeography would seem to make it a useful model for the analysis of practice driven controversies in general. To see why, we briefly examine the root of the controversy.

2.2. A common theoretical framework for phylogeography: coalescent theory

A key theoretical element that is common to all perspectives in phylogeography is coalescent theory. The theory of coalescence was presented as an independent mathematical elaboration by Kingman (1982), Hudson (1983), and Tajima (1983), to trace present day genetic lineages back in time to their most recent common ancestors. Coalescence is the reverse of divergence: as we move forward in time, we can think of an individual DNA molecule replicating and siring two new genetic lineages. Such events are known as divergences. Looking backward in time, whenever two genetic lineages merge into the same ancestor, they are said to 'coalesce'.

Coalescent theory translates the intuitive notion stated above into a series of mathematical models whose main goal is to calculate the time elapsed between the most recent common ancestor and the genetic variants found in present day populations. Because there is no single way in which genetic lineages could coalesce, coalescent models are probabilistic in nature, and so their final outcome is a series of evolutionary scenarios that must then be evaluated to infer which is the most likely given certain parameters. It is in this context that statistical methods that apply coalescent theory to phylogeography emerged, and it is in this context that the controversy originated as each method brought along its own particular vision on how to infer evolutionary history from available evidence.

2.3 Phylogeography becomes statistical: nested clade phylogeographic analysis (NCPA)

The first statistical method developed for phylogeography is called Nested Clade Phylogeographic Analysis (NCPA) developed by Alan R. Templeton (originally named NCA; Templeton et al. 1995;

Templeton 1987, 1998, 2001, 2002, 2006, 2008, 2010a; Castelloe and Templeton 1994). NCPA operates by estimating the genealogical relationships between DNA segments sequenced from samples obtained from throughout the range of a given species. The branching diagrams resulting from the computerized NCPA analyses are known as 'haplotype networks' [2]. The 'nested clade' aspect of NCPA refers to a step in which the haplotypes are arranged into a nested hierarchy of groups (Fig. 1). NCPA algorithms then calculate by null-hypothesis testing a pair of indices that reflect expectations regarding how widely haplotypes are distributed, how abundant they are, and how far they have moved historically. Significantly large or small indices are taken as indicating processes such as geographical range expansion or shifts in the center of distribution. To interpret the indices, Templeton et al. (1995) provided an inference key. The inferences drawn with the key were supposed to be deductive consequences of what should obtain given the general expectations of coalescent theory. Accordingly, different evolutionary scenarios, e.g. random mating, gene flow with isolation by distance, sudden range expansion, long-distance dispersal, or events that divide one large population in two smaller ones, should produce differing patterns of significance of the main indices calculated with the associated software (see Templeton 2009a, 2009b).

Figure 1. Haplotype networks and nested clades.

2.4 The 'statistical controversy': NCPA versus model-based approaches

[2] 'Haplotype' denotes any particular DNA variant. The DNA segments used in these studies have numerous positions, each of which may have different states across the individuals sampled. These states can be manifested as substitutions of different bases (G, A, T, or C), or by the presence or absence of a base at a given site (known as an indel, or insertion-deletion event). Different individuals have different combinations of substitutions and indels. Each unique combination of states characterizing a sequenced segment of a DNA molecule is referred to as an haplotype.

NCPA was the undisputed phylogeographic method until 2002, when L. Lacey Knowles and Wayne P. Maddison published a criticism of what they saw as its "primarily non-statistical" methods to make phylogeographic inferences (Knowles and Maddison 2002). In their view, what phylogeography needed were "methods that make both explicit statistical links between process, prediction and test (like the coalescent-based population genetic models) and consider a diverse array of processes of histories" (Knowles and Maddison 2002, p. 2624). They deemed NCPA unsatisfactory, saying that it "does not attempt to distinguish statistically among alternative interpretations, nor does it provide an estimate of the uncertainty in its conclusions." They additionally contended that "for any interpretation derived from Templeton's inference key, we cannot ascertain the confidence limits on the reconstructed history, whether they are so broad as to include many unconsidered alternatives, or if an alternative hypothesis would be almost equally well supported by the data" (*ibidem* p. 2624).

Knowles, Maddison, and other authors (see Knowles 2004, 2008, 2009, Nielsen and Beaumont 2009, and Panchal and Beaumont 2010) contend that phylogeographic inferences are best performed through the use of a family of Bayesian-based statistical models, the most popular of which is called approximate Bayesian computation (ABC; see, for instance, Bertorelle et al. 2010, Beaumont et.al. 2010). According to these authors, "ABC is matching, for the first time in population genetics studies, abundant genetic data and realistic (which usually means complex) evolutionary scenarios" (Bertorelle et al. 2010: 2610), allowing biologists to assess probabilistically a wide array of possible evolutionary hypothesis (Figure 2). Proponents of ABC say that their approach offers a major advantage over NCPA because it is able to measure the "credibility" of their results and to assess more realistic scenarios.

Figure 2. Schematic rationale behind ABC.

As the controversy progressed, model-based methods (as Bayesian based methods came to be known in the context of the controversy), were criticized by Templeton on statistical grounds. For example, he said that his "main objection to ABC was that it can produce posterior 'probabilities' that are not true probabilities" (Templeton 2010b, p. 488), and that "the potential of ABC is currently not realized because of serious statistical and mathematical flaws" (p. 489). Perhaps the most important of these flaws is that ABC models only the evolutionary scenarios programmed by the user, whereas NCPA should discover the footprints of a wide array of processes from salient patterns in the data. Templeton advocates using NCPA to narrow down the space of possible causes for the patterns in the data and then to model them in detail with ABC (Templeton 2009a).

We present this diversity of opinions to highlight that neither NCPA nor the 'model-based' sides in the phylogeography controversy have declared that genetic lineage divergence/coalescence should not be modeled or disagree on any point of coalescent theory. Up to this point, the controversy in phylogeography seems to be about the flaws of two statistical methods. However, we claim that the core of the controversy lies in what different groups of phylogeographers take as good science, and that these notions of science are reflected in scientists' choices of sources of epistemic credibility. To defend this claim we further develop our philosophical framework, and then revisit the controversy.

4. 'Borrowed epistemic credibility' and the phylogeography controversy

Our reading of the phylogeography controversy seems to highlight numerous instances in which scientists do not use empirical arguments to justify their theories, methodologies, or even their results, instead appealing to non-empirical elements such as simplicity, familiarity, or predictive power. These non-empirical elements are known as scientific values or virtues (Longino 1990; Lacey 2005; Kincaid, Wylie and Dupré 2007, Douglas 2009, 2013). Scientific values are usually

invoked in relation to theories or models. For example, in a popular account Larry Laudan (2004) divides values into epistemic (truth indicative) and cognitive (those held by scientists but not truth indicative), but only as applied to theories. Likewise, Hugh Lacey talks of values as "characteristics that scientific theories or hypotheses should possess" (Lacey 2004: 24; see also Douglas 2009, 2013). In this paper, we propose that thinking regarding values can be extended from just the domain of theories and models to that of practice as well. This extension to practice is especially well illustrated in the phylogeography case because quarreling scientists actively seek to bestow certain values upon their work. So, in contrast to previous accounts, in this paper, philosophical analysis is shifted away from theories and the values they may have, to studying the values embraced by particular communities at a given place and time. In what follows, we refer to scientists as defending claims to underscore that scientists are defending particular methods, specific evolutionary scenarios, and the practices that delimit one discipline from another.

Across disciplines, values as applied to practices can be divided into cognitive and social. Cognitive values are those regarded as being constitutive to science and include, for example, simplicity, or predictive and unificatory power. Social values have been regarded as not being constitutive to science, for example, norms, beliefs, or moral preferences. Traditionally, philosophy of science has considered good scientific practice as involving only cognitive values based on the ideal that the claims of science should be evaluated on empirical grounds alone. Yet, in recent years, work coming especially from the feminist philosophy of science and the philosophy of the social sciences has contributed in showing that the distinction between social and cognitive values is far from clear-cut and that social values play such an important role in scientific matters that they are clearly constitutive to science (e.g. Longino 1990, Kincaid, Wylie and Dupré 2007, Gervais 2013). The interpretation we offer here of the phylogeography controversy helps contribute to this growing recognition of the important interaction of both cognitive and social values in scientific

practice, and of the inseparable links between knowledge (in the form of theories, models, methodologies,…) and the acts of knowing.

The blurry distinction between cognitive and social values is especially salient in discussions regarding methodologies. For example, in the case of phylogeography the methods under discussion could be evaluated by the results they provide but only relative to previously accepted results generated elsewhere. This would imply, for example, appeals to cognitive values such as robustness, the retrieval of known results through the use of different approaches, or familiarity, which refers to approaches that produce results congruent with prevailing views. Robustness and familiarity are cognitive values only if we believe that known results correspond to the actual structure of the world. However, if for any reason we think that a given result does not represent facts of nature but instead an accepted belief shared by a particular community, then they are social values. The borderline is very narrow. For example, consider that just 40 years ago, it was generally accepted that no organism could live at temperatures above 100°C or in very acidic environments such as the human stomach. Going against this belief was one of the main problems faced by Warren and Marshall in showing that *H. pylori* was a causal factor for peptic ulcer disease (Warren and Marshall 1984). In the 1970s any experimental result confirming that no organism could live at certain acidity levels would have been seen as exhibiting cognitive values such as robustness or familiarity, but seen today, the acceptance of those same results simply reflects the beliefs held at the time.

If empirical evidence alone cannot provide conclusive elements to support a given claim, then the claim needs to be supported further by what the particular scientific community considers the standards of good science, a point that has been repeatedly made (e.g. Rudner 1953, Longino 1990, Chang 2004, Douglas 2009). If phylogeography is any indication, then it would seem that the

choice of a given set of standards is not one made idiosyncratically by individual researchers. Instead, entire disciplines construct and use a given set to defend their particular vision of science. While the discussion of how certain cognitive and social values come to be cherished by particular disciplines is beyond the scope of this paper, it is safe to say that it is the result of historical processes having to do with the advent of certain technologies, the value placed on certain forms of reasoning at a given time, or by the agendas pushed by a certain group of individuals for numerous reasons. In our case study, phylogeography is a young discipline trying to find its identity within evolutionary biology. At least two contrasting visions have emerged, represented by the sides involved in the controversy, each exploiting recent advances in molecular biology, genetics, geographic information systems, or computer power. Whenever members of either group appeal to good science, they do it either to endorse or to be sanctioned by a particular group representing certain understanding of what phylogeography ought to be.

4.1. Epistemic credibility in action. The general case

We suggest that, when phylogeographers use epistemic credibility, they do it in one of two ways: a) they *borrow epistemic credibility,* looking for justification by siding with a particular vision of what phylogeography ought to be. In this context, epistemic credibility refers to the set of values, both cognitive and social, that have been cherished by a particular community as the signature of good science. Remember that good science in this context is represented by methodologies, simulations, and practices broadly construed, that exhibit such values. Borrowing epistemic credibility means *constructing an argument in which a researcher connects her claims to the set of good science values by invoking some similarity between her research and previously accepted claims that have a well-established set of values*. Previously accepted claims need not come from the same field as the claims in search of epistemic credibility. For example, assume that a given

scientist is trying to defend a given conclusion based on a set of genetic results. She may borrow epistemic credibility from previously accepted claims in genetics, but she could also borrow it from anthropology, history, sociology or any other field. For example, researchers in genomic medicine have borrowed epistemic credibility from history and anthropology to argue that, out of certain social dynamics, there are at present particular populations genetically more susceptible to suffer from diseases such as type 2 diabetes, or certain types of cancer (see for example, Montoya 2011).

b) Alternatively, scientists can also *lend epistemic credibility* to endorse a given side in a debate. Lending epistemic credibility often occurs when a prestigious scientist highlights similarities between her well-accepted research and that of the people seeking epistemic credibility. In either case, when scientists borrow or lend epistemic credibility, there is an implicit rhetorical element to their stance. This stance is that, if the defended claim resembles in any sense the methodologies, theories, models, or results produced in separate fields, then it should somehow also possess their values. According to our analysis, when scientists have successfully borrowed epistemic credibility for their defended claims, they have also effectively endowed these claims with values that in turn constitute their justification. To illustrate the borrowing of epistemic credibility, we now turn to our case study.

4.2. How, and from where, is epistemic credibility borrowed? The phylogeography controversy as a case study

One of the most popular sources of BEC in phylogeography is what we may call 'the logic of inference', meaning an appeal to what phylogeographers understand as valid or established rules of logical inference. For example, both sides seem to agree on Popperian falsificationism as an ideal of scientific inference. NCPA has been defended as a Popperian approach because it successively subjects null hypotheses to rejection (Templeton 2009a). Templeton charges that, if

the pool of all plausible hypotheses is considered, then successively rejecting competing hypotheses will lead to a "strong" inference (in the sense of Platt 1964, see also Chamberlin 1897, Beard and Kushmerick 2009). In contrast to NCPA, because model-based approaches can only compare the relative fits of a small number of modeled scenarios, it cannot be considered strong scientific inference. The conclusion drawn is that NCPA should be preferred because it conforms to the rules of (some sort of) formal logic, e.g. when Templeton mentions that “the statistics or probabilities used to measure the goodness of fit of the models obey the constraints imposed by formal logic” (2010c: 6376).

Supporters of model-based approaches do not refute the view of Popperian falsificationism as a valid approach, and indeed seem to share it with supporters of NCPA. Instead, they accuse NCPA of being inductive, traditionally presented in evolutionary biology as the antithesis to Popper and indeed to science (Mayr 1982, Jaksić 1981), such as when Beaumont and Panchal (2008, p. 2564) say that "Templeton (2008) cites Popper (1959) in support of the NCPA approach against model-based statistical analysis. However, we would suggest that although NCPA consists of a large number of hypothesis tests based on permutation methods, in the end it follows an inductivist paradigm of trying to derive a general explanation directly from the data… By contrast to NCPA, in model-based analysis [such as the ABC approach], one model is pitted against another in the face of the data, and this, surely, is a more valid scientific approach” (Beaumont and Panchal 2008, p.2564). Templeton and collaborators borrow credibility in support of their claims from the long tradition in phylogenetics that has endorsed Popperianism (Platnick and Gaffney 1978, Stamos 1996, Helfenbein and De Salle 2005). In doing so, what Templeton is saying is that NCPA should be preferred because it represents an example of what other phylogeneticists have considered acceptable scientific inference. Likewise, Beaumont, Panchal, and ABC defenders in general

borrow credibility from the recent surge of Bayesianism in evolutionary studies to back their assertions (for a review see for example, Huelsenbeck 2001).

Another source of disagreement is over the value of verbal versus quantitative reasoning. For example, Beaumont and Panchal (2008) charged that "A verbal, reasoned, argument is presented in Templeton et al. (1995) to justify the method, and the inferences it makes, not dissimilar in style and authority to the Corpus Aristotelicum. The authors of 265 papers that have used NCPA are, in a sense, appealing to this authority. One needs to ask: is this science?" (p. 2564). The appeal to what counts as good science is explicit in Beaumont and Panchal's critique of NCPA. In the view of these authors, at issue is not the argument presented by Templeton and colleagues, but the lesser value of verbal reasoning versus quantitative expressions.

Other instances in which phylogeographers search for epistemic credibility involve notions of robustness and familiarity. Familiarity and robustness have long been held as standards of good science. For example, Newton-Smith listed them among his standards of "good scientific theories". He said that a good theory should preserve the observational success of its predecessors, and that theories should have a good track record (Newton-Smith 1981: 226-232). Templeton does not shy away from these long-held values to support NCPA and attack ABC. Examples can be found in Templeton's critique of Fagundes's use of the ABC approach to contrast three different models of human evolution (Templeton 2008, 2009a, 2010). Templeton charges that one of Fagundes's models is contrary to the prevailing notion of isolation by distance between humans living in Eurasia and Africa, noting that "it is patent that the parameter values chosen by Fagundes et al. (2007) are strongly discrepant with the empirical data on autosomal coalescent times" (Templeton 2009a, p. 323). Templeton invokes robustness and familiarity of the results, alluding to other

controversies in evolutionary biology to show that ABC models do not reach the conclusions generated in similar fields.

As a final instance, epistemic credibility is frequently borrowed from different representations of authority. A conspicuous example is Beaumont et al. (2010), in which 22 authors unite in a single paper to express their reservations regarding NCPA and their support of ABC models. The message of this surfeit of authors would seem to be that the endorsements of many scientists against one approach implies that it is incorrect. A similar implication that solitariness is associated with the incorrect position is when Beaumont and Panchal (2008, p. 2563) note that "there is a disagreement between Templeton (2004, 2008), who suggests the method works well, and three independent groups (Knowles & Maddison 2002; Petit & Grivet 2002; Panchal & Beaumont 2007), who believe that they have demonstrated that it does not. As far as we are aware, there are currently no publications other than those of Templeton and co-workers to support the accuracy or efficacy of NCPA."

The preceding instances represent the appeal to the authority implied by consensus among many scientists, but another source of EC can be to appeal to the authority of a single prominent figure. Knowles (2008, p. 2712) exemplifies this when she refers to authors who voice "other concerns over the validity of NCPA's inferences." Among these authors she cites evolutionary geneticist and bioinformatician Joseph Felsenstein, one of the leading developers of the methods used for reconstructing the evolutionary relationships of organisms (see, for example, Felsenstein 1985, 2004, or 2008). Given his prominence in phylogenetics, his verdict against NCPA would naturally have considerable weight. However, the only reference to NCPA in Felsenstein's well-known treatise on phylogenetic inference (Felsenstein 2004) has the following structure: "A more statistical approach was taken by Templeton (1998), using the nested clade analysis tree

reconstruction methods introduced earlier by Templeton et al. (1988). Although well-defined enough to be implemented by computer programs (Clement, Posada, Crandall, 2000; Posada, Crandall, and Templeton, 2000), these methods do not attempt to take into account the uncertainty of the estimate of the tree, and there has been little study of their statistical properties. A notable exception is the paper by Knowles and Maddison (2002). Although the need to use manual steps in the analysis limited the number of replications they could make, they found that the single-tree approach was problematic." (Felsenstein 2004, p. 484). In contrast to the assertion of Knowles (2008), Felsenstein does not express anything that can be construed as 'other concerns' beyond what had already been discussed in the literature. As a result, we can only interpret Knowles's citation of Felsenstein as an attempt to borrow credibility from his authority, and to shore up her position. Finally, Rémy Petit (2008a, p. 1404) appeals also to authority, of institutions in this case, when he concludes that "the results of Panchal & Beaumont (2007) convince me that reputable journals should (i) discourage the use of the NCPA method for single locus data sets (...), and (ii) still be suspicious of NCPA analyses based on multiple loci". Reading between the lines of Petit's argument, any journal publishing NCPA studies should be held in suspicion.

Table 1. Some sources of BEC in the phylogeography controversy.

5. Epistemic credibility, the phylogeography controversy, and the theory-practice divide

Thinking of theory as standing separately from practice has a long tradition in the philosophy of science. As a result, philosophers of science have an extensive battery of conceptual tools at their disposal to think about the theory-practice relation. The vocabulary of laws and axioms, of the distinction between models and theories, syntactic and semantic relations to the world, as well as notions of theories as explanatory devices, all stem from the conception of theory as being in some way distinct from practice. The view of theory as interacting inextricably with practice has less of a tradition, and because of the complexity of this view, much conceptual machinery remains to be built (e.g. Keller 2002, Rheinberger 1997, Pickering 1995, Douglas 2013). Here, we have offered the notion of borrowed epistemic credibility (BEC), which we believe can help understand how theory and practice interact to construct scientific knowledge.

Our analysis of the phylogeography controversy shows how scientific values go beyond the virtues exhibited by particular theories or hypotheses. These values become an important part of scientific practice itself, as scientists actively seek to bestow certain values upon their work. Speaking in terms of epistemic credibility might lead some people to believe we are endorsing a subjective account of science in which, regardless of its real merits (for example, in terms of experimental evidence), any claim endorsed by the right people will come to be regarded as good science. However, note that BEC is restricted to cases in which experimentation does not (or has not yet) provide conclusive evidence. Therefore, claims must be evaluated intersubjectively by the community. In these instances, what is considered objective is the result of a general agreement by the community based on the strength of whatever experimental evidence there is, its set of values, and sometimes, who is vouching for it (see for example, Longino 1990, Douglas 2009, or Gervais 2013). Intersubjective objectivity is by no means a relativist stance and our point is that particular communities have developed strict rules to evaluate what should be considered as good science.

We do not claim that scientific controversies closely linked to practice and/or methodology exclusively rest on epistemic credibility. However, we argue that epistemic credibility likely plays an important role in fields in which direct evidence is often difficult to generate, as is the case with anthropology, archeology, or evolutionary studies to name three examples. In these cases, scientists often lack all the evidence that would be necessary to back their assertions and as a result appeal to epistemic credibility. For example, scientists involved in the molecular clock controversy indirectly appealed to epistemic credibility (Dietrich and Skipper 2007). We see a similar analytical context in Winsberg (2006, p. 2), who states that the credibility of a simulation model comes not only from its governing theory, but also "from the antecedently established credentials of the model building techniques employed by the simulationists". In these examples, and similar to what we have seen in phylogeography, scientists have had to look for arguments beyond theory and its empirical consequences to settle a theoretical dispute or to justify particular modeling practices. However, whereas Dietrich and Skipper and Winsberg make reference to extra-theoretical elements in the context of epistemic virtues and/or the social aspects of science surrounding scientific disputes, our framework places BEC at the core of scientific controversies, not at the periphery.

Phylogeography is an excellent field with which to examine this phenomenon, because it is a new discipline forged by the union of phylogenetics, biogeography, and population genetics (Avise 2000, Avise 2009, Hickerson et al. 2010). Judging from the path that phylogeography has followed, it would seem that newly arisen fields do not emerge out of the construction of novel theory but from the pulling together in the laboratory of disparate research traditions. Theory, along with practice generally construed, and the set of values used to justify certain claims, may be inherited from the different fields being brought together, but it is by no means clear why certain elements make the step to the new field, or how theory originated within the new field is constructed.

However, the conceptual tools that we propose here can illuminate some aspects of the forging of new disciplines by exposing the web of distinct elements imported from other fields and how they are used to construct a new discipline. Again, the notion of epistemic credibility is vital as the new discipline imports not only theory and methodologies but, as the phylogeographic debate illustrates, what the members of the new community should take for rational, objective, valid, or scientific.

5.1 Why different groups use different strategies for borrowing epistemic credibility

Practices of borrowing of epistemic credibility demarcate traditions or subdisciplines rather than the idiosyncratic tastes of individual researchers. A glance at the bibliography of recent publications in phylogeography reveals two traditions. The authors and bibliography of Beaumont et al. (2010) map out much of the "model-based" school (see also Bloomquist et al. 2010) whereas the cites of other authors (e.g. Templeton 2010a) trace the NCPA school. Given these schools and their relations between labs and across generations, one can trace back sources of epistemic credibility in the debate to other important values held by different schools within evolutionary biology. To name some examples, different schools defend their own interpretations of statistical inference (see Sarkar 1992), the value of graphical versus mathematical representation (see Petit 2008), or the value of verbal versus mathematical reasoning (Mayr 1982; Beaumont et.al. 2010). The debate over NCPA at least in part would seem to illustrate the social dimension of knowledge. From this point of view, normativity is a rational standard provided by a community (see for example Hacking 1992 or Kusch 2002). Applied to our case, phylogeographers in disagreement represent two communities within evolutionary biology that have quarreled for decades over the correct way to conceptualize their field, evolutionary history, and science itself. Each community defends a particular set of values that in a sense reflect a style of reasoning that as Hacking argues,

becomes “a timeless canon of objectivity, a standard or model of what it is to be reasonable about this or that type of subject matter” (Hacking 1992: 10) (See figure 3). The values defended by each community are made explicit in the sources from which epistemic credibility is borrowed by NCPA and model-based supporters. If this is correct, then it should be possible to reconstruct the intellectual history of evolutionary biology to show how and why different factions emerge and what their differences are. Perhaps such an analysis could help find ways to work out the differences.

Figure 3. Diagrammatic representation of the set of controversies in phylogeography. Phylogeography is a synthetic discipline formed from the union of different disciplines within evolutionary biology. The controversies are represented, first, as a controversy between model-based and non-model based methodologies (central circle). Second, as a controversy about phylogeography itself as competing groups appeal to different sources of BEC based on particular notions of what counts as good science (outer circles), and finally, as a controversy about science itself.

6. Conclusion

The controversies surrounding phylogeography illustrate how a seemingly methodological controversy may in fact reflect a larger debate between two visions of what science should be. This larger debate is evident in the numerous sources of epistemic credibility borrowed by the participants. From a philosophical perspective, we illustrate a case of scientific controversy that takes place largely in the realm of practices. In our analysis, the interplay of theory and practice-based controversies, scientific evaluation using social and cognitive values, and the web of conceptual relationships traced by the borrowing of epistemic credibility, contrasts with the

traditional conception of theory as separate from practice. Instead, our analysis adds support to the increasing recognition of the inseparable reciprocity between concepts and practice in science.

Our analysis also highlights the interplay between scientific issues of fact and value that "have too often been left implicit rather than confronted head on (Kincaid et al. 2007, p. vii). The sources of epistemic credibility invoked in this particular controversy are not only rhetorical devices used to defend a particular methodology, but trace the evolution of the field of genealogical studies by exposing the epistemic, methodological, and theoretical commitments shared by its different communities. Our reconstruction shows how the controversy transcends the particulars of model-based models versus NCPA to deal with numerous fundamental disagreements in terms of the importance of gene trees, the need to automate the inference process, the importance of formal inferences, and in general the sources of epistemic credibility that justify the claims of a given community and give identity to the field. Perhaps, then, this controversy is another chapter in a long debate evolutionary biologists have had to understand better their field, and science in general, and shows how complex practices, in which the distinction between theory and practice is difficult to demarcate, begs revision of notions of the nature of explanation and how scientists understand what is at stake behind notions of good science.

7. References

Arroyo-Santos A., Olson M.E., Vergara-Silva, F. 2013. (in press). "The phylogeography debate and the epistemology of model-based evolutionary biology." *Biology & Philosophy*. DOI 10.1007/s10539-013-9391-1

Avise, J.C. 1992. "Molecular population structure and the biogeographic history of a regional fauna: a case history with lessons for conservation biology." *Oikos* 63:62-76.

Avise, J.C. 1998. “The history and purview of phylogeography: a personal reflection.” *Molecular Ecology* 7: 371-379.

Avise, J.C. 2000. *Phylogeography: The History and Formation of Species*. Harvard University Press: Cambridge.

Avise, J.C. 2009. “Phylogeography: retrospect and prospect.” *Journal of Biogeography* 36: 3–15.

Avise, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T., Neigel, J.E., Reeb, C.A., Saunders, N.C. 1987. “Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics.” *Annual Review of Ecology and Systematics* 18:489-522.

Beard, D.A., Kushmerick, M.J. 2009. “Strong inference for systems biology.” *PLoS Computational Biology* 5 (8): e1000459.

Beaumont, M., Panchal, M. 2008. “On the validity of nested clade phylogeograpphical analysis.” *Molecular Ecology* 17: 2563–2565.

Beaumont, M.A., Nielsen, R., Robert, C., Hey, J., Gaggiotti, O., Knowles, L., Estoup, A., Panchal, M., Corander, J., Hickerson, M., Sisson, S.A., Fagundes, N., Chikhi, L., Beerli, P., Vitalis, R., Cornuet, J.M., Huelsenbeck, J., Foll, M., Yang, Z.H., Rousset, F., Balding, D., Excoffier, L. 2010. “In defence of model-based inference in phylogeography.” *Molecular Ecology* 19: 436 – 446.

Beaumont, M.A., Rannala, B. 2004. “The Bayesian revolution in genetics.” *Nature Reviews Genetics* 5: 251-261

Bertorelle, G., Banazzo, A., and Mona, S. 2010. “ABC as a flexible framework to estimate demography over space and time: some cons, many pros.” *Molecular Ecology* 19: 2609-2625.

Bloomquist, E.W., Lemey, P., Suchard, M.A. 2010. “Three roads diverged? Routes to phylogeographic inference.” *Trends in Ecology & Evolution* 25: 626-32.

Castelloe, J., Templeton, R.A. 1994. “Root probabilities for intraspecific gene trees under neutral coaslescent theory.” *Molecular Phylogenetics and Evolution* 3: 102-113.

Chamberlin, T.C. 1897. “The method of multiple working hypotheses.” *Journal of Geology* 5: 837.

Chang, H. 2004. *Inventing Temperature. Measurement and Scientific Progress*. Oxford University Press: New York.

Clement, M., Posada, D., Crandall, K. 2000. “TCS: a computer program to estimate gene genealogies.” *Molecular Ecology* 9: 1657–1659.

Dietrich, M.R., Skipper, R.A. 2007. “Manipulating Underdetermination in Scientific Controversy: The Case of the Molecular Clock.” *Perspectives on Science* 15: 295-326.

Douglas, H.E. 2009. *Science, Policy, and the Value-Free Ideal*. University of Pittsburgh Press: Pittsburgh.

Douglas, H.E. 2013. “The Value of Cognitive Values.” *Philosophy of Science* 80: 796-806.

Fagundes, N.J.R., Ray, N., Beaumont, M., Neuenschwander, S., Salzano, F.M., Bonatto, S.L., Excoffier, L. 2007. “Statistical evaluation of alternative models of human evolution.” *Proceedings of the National Academy of Sciences* 104: 17614–17619.

Felsenstein, J. 1985. Phylogenies and the comparative method. *American Naturalist* 125: 1–15

Felsenstein, J. 2004. *Inferring Phylogenies*. Sinauer Associates: Sunderland, Mass.

Felsenstein, J. 2008. " Comparative methods with sampling error and within-species variation: contrasts revisited and revised." *American Naturalist* 171: 713-725.

Gervais, R. 2013. "Non-cognitive values and objectivity in scientific explanation: egalitarism and the case of the Movius line." *Perpectives on science* 21: 429-452.

Hacking, I. 1992. "'Style' for Historians and Philosophers." *Studies in the History and Philosophy of Science* 23: 1-20.

Helfenbein, K.G., DeSalle, R. 2005. "Falsifications and corroborations: Karl Popper's influence on systematics." *Molecular Phylogenetics and Evolution* 35: 271–280.

Hickerson, M.J., Carstens, B.C., Cavender-Bares, J., Crandall, K.A., Graham, C.H., Johnson, J.B., Rissler, L., Victoriano, P.F., Yoder, A.D. 2010. "Phylogeography's past, present, and future: 10 years after Avise." *Molecular Phylogenetics and Evolution* 54: 291–301.

Hudson, R.R. 1983. "Properties of a neutral allele model with intragenic recombination." *Theoretical Population Biology* 23: 183–201.

Huelsenbeck, J.P., Ronquist, F., Nielsen, R., Bollback, J.P. 2001. "Bayesian inference of phylogeny and its impact on evolutionary biology." *Science* 294: 2310-2314.

Jaksić, F.M. 1981. "Recognition of Morphological Adaptations in Animals: The Hypothetico-Deductive Method." *BioScience* 31: 667-670.

Keller, E.F. 2002. *Making Sense of Life : Explaining Biological Development with Models, Metaphors, and Machines.* Harvard University Press: Cambridge.

Kincaid, H., Dupré, J., Wylie, A. 2007. *Value-free science? Ideals and illusions.* Oxford University Press: New York.

Kingman, J.F.C. 1982. The coalescent. *Stochastic Processes and their Applications* 13: 235–248.

Knowles, L.L. 2004. "The burgeoning field of statistical phylogeography." *Journal of Evolutionary Biology* 17: 1-10.

Knowles, L.L. 2008. "Why does a method that fails continue to be used?" *Evolution* 62: 2713-2717.

Knowles, L.L., Maddison, W.P. 2002. "Statistical phylogeography." *Molecular Ecology* 11: 2623–2635.

Kusch, M. 2002. *Knowledge by Agreement: The Programme of Communitarian Epistemology*. The Clarendon Press: Oxford.

Lacey, H. 2004. "Is there a significant distinction between cognitive and social values?" pp. 24-51 in P. Machamer & G. Wolters (Eds.) *Science, Values, and Objectivity*. Pittsburgh: University of Pittsburgh Press.

Lacey, H. 2005. *Values and Objectivity: The Controversy over Transgenic Crops*. Rowman and Littlefield: Lanham, MD.

Laudan, L. 2004. "The epistemic, the cognitive, and the social." Pp. 14-23 in P. Machamer & G. Wolters (Eds.) *Science, Values, and Objectivity*. Pittsburgh: University of Pittsburgh Press.

Longino, H.E. 1990. *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry.* Princeton University Press: Princeton.

Mayr, E. 1982. *The Growth of Biological Thought: Diversity, Evolution and Inheritance.* Harvard University Press, Cambridge.

McMullin, E. 1987. “Scientific controversy and its termination.” pp. 49–91 in Engelhardt H.T., & Caplan, A.L. (Eds) *Scientific controversies: Case studies in the resolution and closure of disputes in science and technology.* Cambridge University Press, Cambridge.

Montoya, M. 2011. *Making the Mexican Diabetic: Race, Science, and the Genetics of Inequality*. University of California Press: Berkeley.

Newton-Smith, W.H. 1981. *The Rationality of Science.* Routledge and Kegan Paul: London.

Nei, M., Takahata, N. 1993. “Effective population size, genetic diversity, and coalescence time in subdivided populations.” *Journal of Molecular Evolution* 37:240-244.

Nielsen, R., Beaumont, M.A. 2009. “Statistical inferences in phylogeography.” Molecular Ecology 18: 1034-1047.

Panchal, M., Beaumont, M.A. 2007. “The automation of nested clade phylogeographical analysis.” Evolution 61: 1466–1480.

Panchal, M., Beaumont, M.A. 2010. Evaluating nested clade phylogeographic analysis under models of restricted gene flow. Systematic Biology 59:415-432.

Petit, R.J. 2008a. “The coup de grâce for nested clade phylogeographic analysis?” Molecular Ecology 17: 516–518.

Petit, R.J. 2008b. “On the falsifiability of the nested clade phylogeographic analysis method.” Molecular Ecology 17: 1404.

Petit, R.J., Grivet, D. 2002. “Optimal randomization strategies when testing the existence of a phylogeographic structure.” Genetics 161: 469-471.

Pickering, A. 1995. *The mangle of practice: time, agency, and science*. University of Chicago Press: Chicago.

Platnick, N. I., Gaffney, E.S. 1978. "Evolutionary biology: a Popperian perspective." *Systematic Zoology* 27: 138-141.

Platt, J.R. 1964. "Strong inference." Science 146: 347-353.

Posada, D., Crandall, K.A. 2001. "Intraspecific phylogenetics: Trees grafting into networks". *Trends in Ecology and Evolution* 16: 37-45.

Posada, D., Crandall, K.A., Templeton, A.R. 2000. "GeoDis: A program for the Cladistic Nested Analysis of the Geographical Distribution of Genetic Haplotypes." *Molecular Ecology* 9:487-488.

Rheinberger, H.J. 1997. *Toward a history of epistemic things: synthesizing proteins in the test tube.* Stanford University Press: Palo Alto.

Riddle, B.R. 2008. "What is modern biogeography without phylogeography?" *Journal of Biogeography* 36: 1-2

Rieppel, O. 2008. "'Total evidence' in phylogenetic systematics." *Biology & Philosophy* 24: 607-622.

Rudner, R. 1953. "The scientist qua scientist makes value judgments." *Philosophy of Science* 20: 1–6.

Sarkar, S. (Ed). 1992. *The founders of evolutionary genetics (a centenary reappraisal)*. Kluwer Academic Publishers: Dordrecht.

Smith, R.J. 2009. "Use and misuse of the reduced major axis for line-fitting." *American Journal of Physical Anthropology* 140, 476—486.

Stamos, D.N. 1996. “Popper, falsifiability, and evolutionary biology.” *Biology & Philosophy* 11: 161-191.

Tajima, F. 1983. “Evolutionary relationship of DNA sequences in finite populations.” *Genetics* 105: 437–460.

Templeton, A.R. 1987. “A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. I. Basic theory and an analysis of alcohol dehydrogenase activity in Drosophila.” *Genetics* 117: 343-351.

Templeton, A.R. 1998. “Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history.” *Molecular Ecology* 7: 381–39.

Templeton, A.R. 2001. “Using phylogeographic analyses of gene trees to test species status and processes.” *Molecular Ecology* 10: 779-791.

Templeton, A.R. 2002. “"Optimal" randomization strategies when testing the existence of a phylogeographic structure: A reply to Petit and Grivet.” *Genetics* 161: 473-475.

Templeton, A.R. 2006. *Population genetics and microevolutionary theory.* John Wiley & Sons: New Jersey.

Templeton, A.R. 2008. “Nested clade analysis: an extensively validated method for strong phylogeographic inference.” *Molecular Ecology* 17: 516-518.

Templeton, A.R. 2009a. “Statistical hypothesis testing in intraspecific phylogeography: nested clade phylogeographical analysis vs. approximate Bayesian computation.” *Molecular Ecology* 18: 319-331.

Templeton, A.R. 2009b. “Why does a method that fails continue to be used: the answer.” *Evolution* 63: 807–812.

Templeton, A.R. 2010a. “Coalescent-based, maximum likelihood inference in phylogeography.” *Molecular Ecology* 19: 431-435.

Templeton, A.R. 2010b. “Correcting approximate Bayesian computation.” *Trends in Ecology and Evolution* 25:488-9.

Templeton, A.R. 2010c. “Coherent and incoherent inference in phylogeography and human evolution.” *Proceedings of the National Academy of Sciences* 107: 6376-81.

Templeton, A.R., Routman, E., Phillips, C.A. 1995. “Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, Ambystoma tigrinum.” *Genetics* 140: 767-782.

Winsberg, E. 2006. “Models of Success Versus the Success of Models: Reliability Without Truth.” Synthese 152: 1-19.

Winsberg, E. 2009. “Computer Simulation and the Philosophy of Science.” *Philosophy Compass* 4 (5):835-845.

Winsberg, E. 2010. *Science in the Age of Computer Simulation*. The University of Chicago Press: Chicago.

Woolley, S.M., Posada, D., Crandall, K.A. 2008. “A Comparison of Phylogenetic Network Methods Using Computer Simulation.” *PLoS ONE* 3(4): e1913. doi:10.1371/journal.pone.0001913

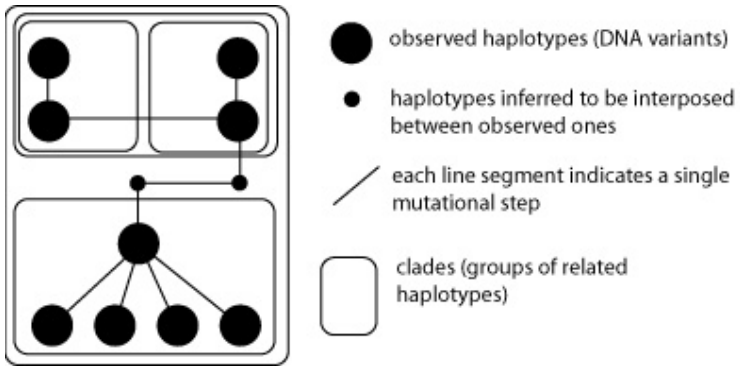

Fig. 1. Haplotype networks and nested clades. Different algorithms are available to infer the way that the DNA variants or haplotypes sequenced are related to one another, and to represent these inferred relationships in a network diagram. Additional algorithms are then used to decide how the network is to be partitioned into a nested hierarchy of clades, or groups of closely related haplotypes. Haplotypes deep in the network are more likely to be ancestral ones than ones that are isolated at the tips, which are more likely to have arisen recently. Ancestral haplotypes have had more time to disperse, so all things being equal, are expected to be more widely distributed geographically.

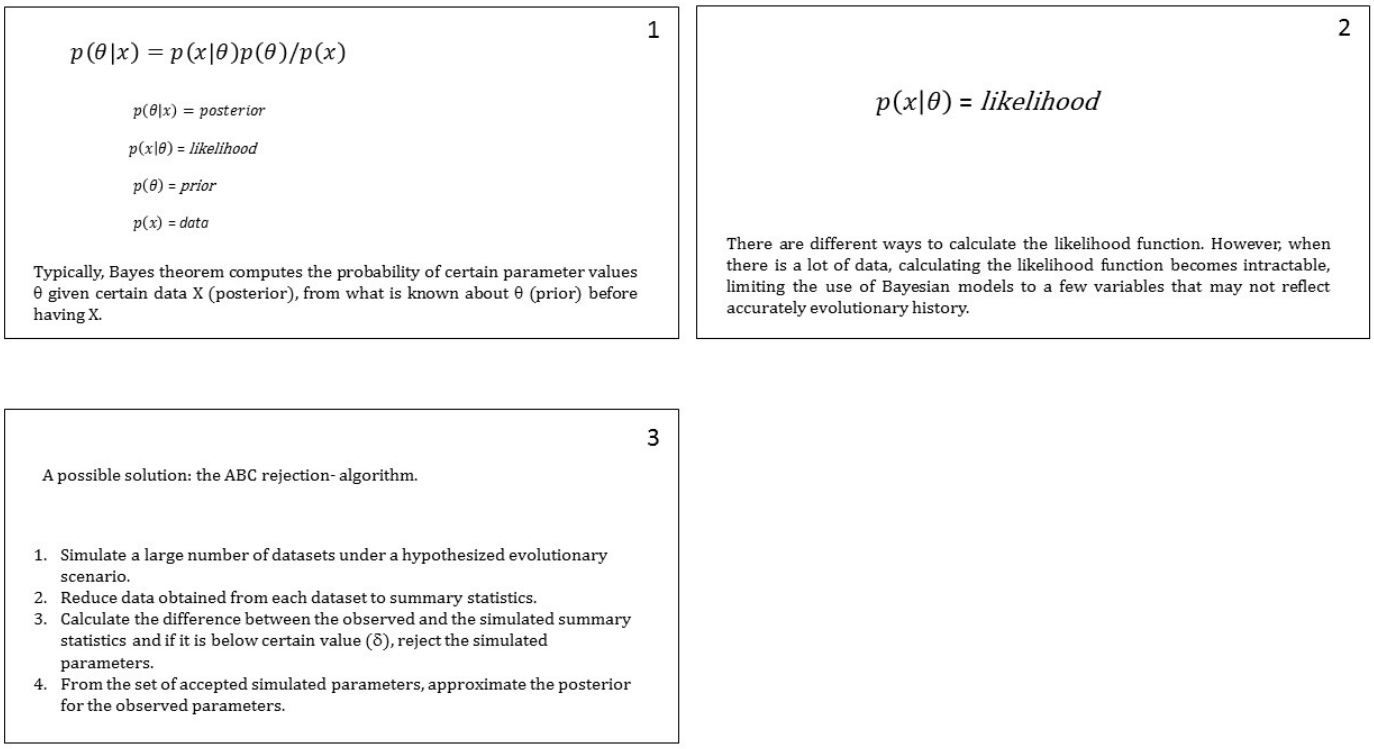

Figure 2. Schematic rationale behind ABC. Bayesian models are limited by the difficulty of calculating the likelihood function as available data increases (boxes 1 and 2). However, constructing models that provide better insights into evolutionary history require considerable data. Methods such as ABC (box 3) circumvent the problem of calculating the likelihood function, and so are becoming a very important tool in phylogeography.

Sources for BEC.	How is that good science?
Robustness	Claims are similar to those of previous works.
Bootstrapping	Claims resemble those that would be expected in the field
Logical inference	Claims conform to the preferred notion of "scientific" inference.
Quantification	Claims rest on numerical values and therefore are more objective than qualitative results
Authority	Claims have been endorsed by prominent scientist or resemble those that have been previously endorsed by prominent figures.

Table 1. **Some sources of BEC. The framework can likely be applied to other controversies, either within evolutionary biology or elsewhere in the life sciences.**

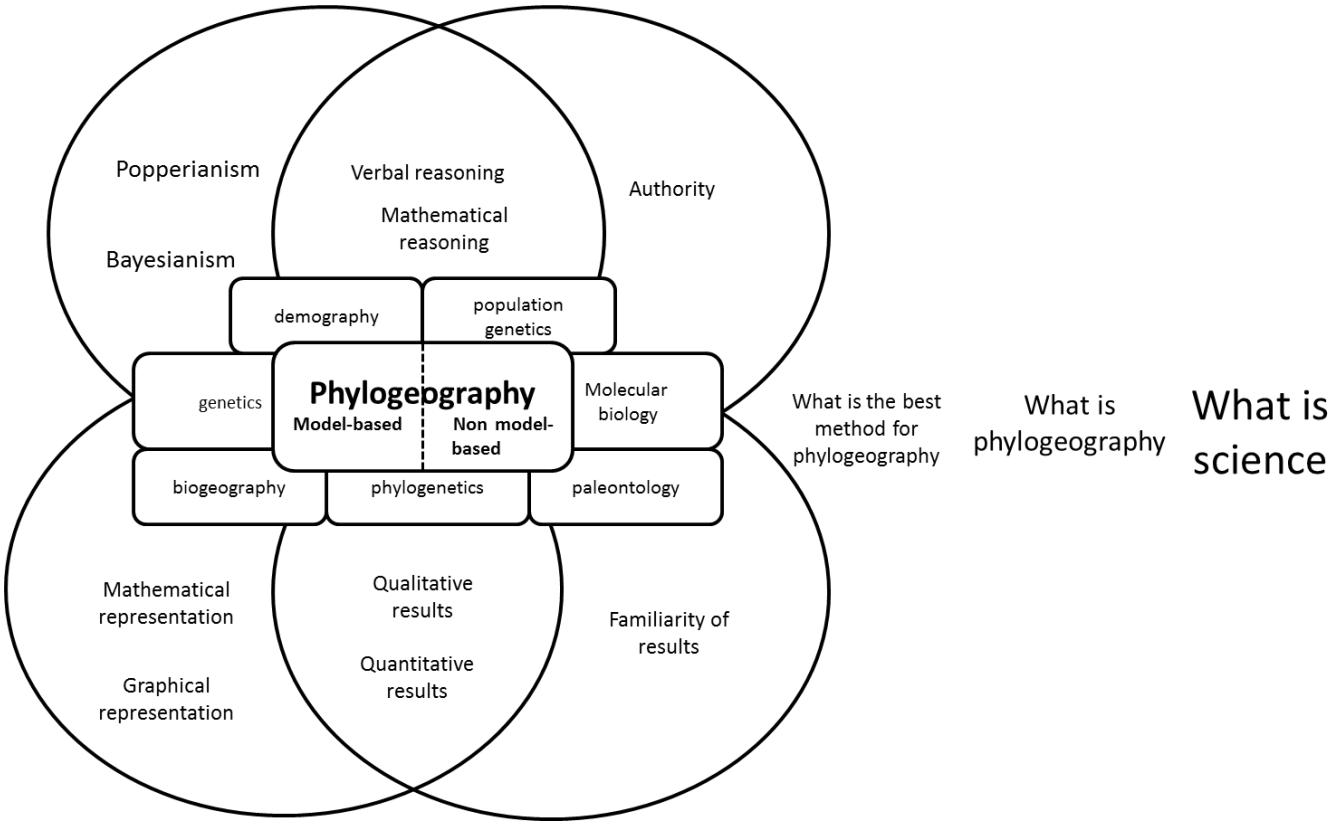

Figure 3. Diagrammatic representation of the set of controversies in phylogeography. Phylogeography is a synthetic discipline formed from the union of different disciplines within evolutionary biology. The controversies are represented, first, as a controversy between model-based and non-model based methodologies (central square). Second, as a controversy about phylogeography itself as competing groups appeal to different sources of BEC based on particular notions of what counts as good science (outer circles), and finally, as a controversy about science itself.