Analysis of parasite host-switching: limitations on the use of phylogenies

J. A. JACKSON*

School of Biological Sciences, University of Bristol, Bristol BS8 1UG, UK

SUMMARY

Even the most generalist parasites usually occur in only a subset of potential host species, a tendency which reflects overriding environmental constraints on their distributions in nature. The periodic shifting of these limitations represented by host-switches may have been an important process in the evolution of many host-parasite assemblages. To study such events, however, it must first be established where and when they have occurred. Past host-switches within a group of parasites are usually inferred from a comparison of the parasite phylogeny with that of the hosts. Congruence between the phylogenies is often attributed to a history of association by descent with cospeciation, and incongruence to host-switching or extinction in ‘duplicated’ parasite lineages (which diverged without a corresponding branching of the host tree). The inference of host-switching from incongruent patterns is discussed. Difficulties arise because incongruence can frequently be explained by different combinations of biologically distinct events whose relative probabilities are uncertain. Also, the models of host–parasite relationships implicit in historical reconstructions may often not allow for plausible sources of incongruence other than host-switching or duplication/extinction, or for the possibility that colonization could, in some circumstances, be disguised by ‘false’ congruence.

Key words: Host-switching, phylogenies, cospeciation.

INTRODUCTION

The factors which contribute to host-specificity, or the tendency to infect a restricted group of hosts (Poulin, 1998), represent a key constraint on the occurrence of parasitic organisms in nature. As outlined by Combes (1991) this constraint may have two aspects. One is the possession by the parasite of adaptations that allow it to deal with the morphological, physiological and immunological characteristics of particular hosts only. The other is the accessibility of such compatible hosts in the environment. These limitations may not be static. Host-specificity can change over evolutionary time, with new host lineages being acquired and others lost. Such host-shifts could arise either because environmental changes bring parasites into contact with new host organisms to which they are preadapted, or because genetic changes in host or parasite lineages alter the way different hosts are utilized. It has been suggested that, due to the cost of host-specific adaptations, there may be a trade-off between the ability to exploit one host efficiently and the ability to exploit others (Poulin, 1998). Host-switching might also be linked to speciation, given that parasite populations dispersing to a new host could become isolated and/or be exposed to novel selection pressures. Changes of host by parasites are, therefore, events of intrinsic evolutionary interest and may have been important determinants of the distributions of present-day parasite groups. However, the factors involved in host-switching can only be studied by comparing lineages that have undergone a history of colonization with those which have not. For this to be possible, instances of host-switching must first be identified. As events that have occurred in the past, colonizations of host lineages by parasites are detected by some form of historical reconstruction. Because parasites have no fossil record, this reconstruction relies primarily on present-day host–parasite distributions and the phylogenies of the host and parasite groups involved. Although parasitologists have long been interested in such analyses (Klassen, 1992), numerical methods have only been developed in the last two decades (Brooks, 1981, 1987, 1988, 1990; Brooks & Bandoni, 1988; Page, 1990a, 1991, 1993, 1994, 1996; Page & Charleston, 1998; Ronquist & Nylin, 1990; Ronquist, 1995, 1998; Siddall, 1996; Charleston, 1998). The motivation for such studies is clear (Brooks, 1988; Paterson & Gray, 1997; Poulin, 1998). Knowledge of the history of a host–parasite association would allow a wide range of inferences about the way in which contemporary characteristics of that association evolved, or the way in which historical events themselves may have been influenced by biological characteristics of the hosts and parasites. It has become increasingly apparent (Brooks, 1980; Poulin 1998; Vickery & Poulin, 1998) that historical host associations may be important determinants of parasite communities, which have often only been

* Tel. +0117 928 7478. Fax: +0117 925 7374. E-mail: Joe.Jackson@bristol.ac.uk
considered in the context of present-day ecological factors. Reconstruction of past host colonizations might also provide information on the evolution of important pathogens. For example, analyses of the phylogenies of immune deficiency viruses and their primate hosts have led to contradictory hypotheses on the origins of HIV infections in humans (Mindell, Shultz & Ewald, 1995; Siddall, 1997).

Both intuitive and numerical reconstructions assume that congruent host-parasite trees result from a continuous association between host and parasite lineages (with parasites being inherited by hosts from their ancestors), where parasite speciation events mirror those in the host tree. Incongruent trees are assumed to be the result of host-switching or certain patterns of parasite speciation and extinction. It has, to some extent, been recognized that limitations exist on how much historical information is recoverable by these analyses (e.g. Ronquist, 1995). However, recently developed numerical techniques are being increasingly used and advocated as rigorous methods to investigate historical host-parasite associations (e.g. Brooks, 1988, 1990; Page, Clayton & Paterson, 1996; Paterson & Gray, 1997; Siddall, 1997; Page & Charleston, 1998). Some systems of associating organisms, for example certain bacterial endosymbionts and their invertebrate hosts (Peek et al. 1998; Clark et al. 2000), have been demonstrated to show highly congruent phylogenies. Such instances might reasonably be considered strong evidence for a dominant history of association by descent and cospeciation. However, it is almost axiomatic that host and parasite phylogenies will usually show some degree of incongruence (Hafner & Nadler, 1990; Brooks & McLennan, 1991). While not questioning that a high degree of congruence is likely to be a signal of cospeciation (Hafner & Nadler, 1990), this study aims to examine critically the basis upon which reconstruction methods differentiate host-switching from other historical processes.

TERMINOLOGY

Various terminology has been applied to historical processes in host–parasite assemblages (Brooks, 1979; Brooks & McLennan, 1991; Page, 1993; Ronquist, 1995; Hoberg, Brooks & Siegel-Causey, 1997; Paterson & Gray, 1997; Page & Charleston, 1998). In the following account (see also Fig. 1) the key terms are defined as follows. Association by descent (Brooks & McLennan, 1991): where host lineages inherit parasites from their ancestors. Co-speciation (Brooks, 1979): where parasite speciation is coincidentally linked to host speciation, with each daughter host species coming to be infected by one of the daughter parasite species. Duplication (Page, 1993, 1994) (= 'intra-host speciation'): where a parasite lineage speciates without a corresponding speciation in the host lineage. Host-switch: where a parasite invades a new host lineage, leading to association by colonization (Brooks & McLennan, 1991). Host-switches may or may not be associated with parasite speciation. Extinction: where a parasite lineage dies out from a particular host lineage (irrespective of whether it survives in other lineages). Page (1993, 1994) used 'sorting event' to include extinctions, loss of a parasite lineage from a host population during the early stages of speciation (= 'missing the boat'), or sampling error. Here sampling error will be disregarded, whilst 'missing the boat' is considered a special case of extinction, distinguished by its occurrence in a small founder population (a biologically equivalent event might occur in a single host lineage undergoing a population bottleneck).

CONGRUENCE AND HISTORY

Reconstructions of the history of host–parasite assemblages are based on the degree of congruence or incongruence of the host and parasite phylogenies. It is assumed that, under a restricted set of circumstances, a parasite undergoing association by descent with a host group will show a phylogeny congruent with that of its hosts (Fig. 2). Departures from congruence might be explained by host-switches. However, a major problem is that patterns of parasite speciation and extinction can also produce incongruence ('false' incongruence) despite a history of association by descent (see Fig. 2). The nature of these two sources of incongruence (host-switching or speciation/extinction), as they relate to an analysis of historical host–parasite relationships, are now considered.
Fig. 2. The effects of cospeciation, duplication and parasite extinction on the congruence of host and parasite phylogenies. A–C. Parasite evolutionary sequence superimposed on host phylogeny (bold) and to its right. Parasite speciation events indicated by closed circles and extinctions by closed squares. A. Congruence due to cospeciation: parasites speciate contemporaneously with their host, with one daughter parasite species being inherited by each daughter host species, producing congruent phylogenies. B. Incongruence produced by one parasite duplication (♀♂) and three extinctions; parasite phylogeny also shown above with host–parasite associations indicated by stippled lines. C. Colonization sequence producing identical parasite phylogeny to scenario B (colonizations indicated by horizontal stippled lines). (Note ages of nodes. In scenario A (cospeciation) host and parasite nodes are contemporaneous. In scenario B (duplication/extinction) the single duplication (♀♂) is more ancient than any nodes in the host tree. In scenario C (colonization) speciations of parasite lineages switching between host lineages are always younger than the divergence of the hosts. Note, also, that if lineage 1 in scenario B were to switch to any other host and become extinct on its existing host, then it would still be derived from the ancient node (♀♂): in some combinations the roles of host-switching, duplication and extinction may not, then, be distinguished by comparison of host and parasite node ages.

Colonization, speciation and extinction

Host-switching. Factors that promote host-switches by parasite lineages are not well known but a number of reasonable speculations may be advanced. The presence in the immediate environment of organisms similar to an existing host in terms of ecology, morphology, physiology and immune responses might be important in reducing the ‘adaptive jumps’ to new hosts. While transfers of parasites between ecologically similar but phylogenetically distant hosts are well documented, it is likely that, in general, similarity in such characteristics between existing and potential hosts would often depend on their phylogenetic relatedness. This is supported by studies (e.g. Reed & Hafner, 1997) showing that host relatedness determines the level of parasite performance on unnatural compared to natural hosts. The probable link between host phylogeny and parasite host-switches may mean that colonization is more prevalent in parasites of diverse host groups. Poulin (1992), for example, has shown that host range is positively correlated with host group diversity in fish parasites. A further possibility is that the probability of colonization may be affected by potential competitive interactions between established and invading parasite lineages (Barker, 1991, 1996).

At the start of any analysis, the probability of host-switching is unknown and can only be inferred from the degree of incongruence in the host and parasite phylogenies (see below), taking other sources of incongruence into account. Also, given the variety of factors that could influence host-switches, the likelihood of these events might be expected to vary greatly in time and space. For example, parasite transfers may be common in a young, actively radiating host taxon (where many closely related host species are present in the environment), but become less frequent in older taxa composed of relatively ancient, divergent lineages with low speciation rates.

Speciation and extinction. As parasites lack a fossil record their historical speciation and extinction rates are not easy to estimate, the number of species at the present time being determined by the relative rates of speciation and extinction in the past. Identical present-day diversity could, therefore, occur in parasite clades with a rapid or slow turnover of species. It is a major starting limitation of any historical reconstruction that the parasite speciation and extinction rates are unknown or can only be estimated with a high degree of uncertainty. Nee et al. (1994a, b) proposed modelling methods to derive past (and future) speciation and extinction rates from molecular phylogenetic trees. These are likely to be highly sensitive to starting assumptions about the
Other significant historical processes in host–parasite associations

Historical analyses usually only consider congruence as a result of association by descent and incongruence as a result of host-switching or parasite speciation/extinction patterns (although some authors have considered scenarios in which these assumptions might not hold, e.g. Brooks, 1987, 1988, 1990; Page, 1994). In the following section, the potential significance of other sources of ‘false’ incongruence (incongruence despite a history of association by descent) and the potential for ‘false’ congruence (congruence despite a history of host-switching) are considered.

Colonization and phylogeny. False congruence occurs where a history of parasite colonization (Fig. 3) fortuitously produces a parasite phylogeny congruent with that of the hosts. Such scenarios have been considered relatively unlikely on the assumption that colonization is a random process and would only rarely produce congruent patterns. Brooks (1987) suggested spurious congruence might be most likely where a small number of parasites occur in a small number of distantly related hosts. However, as discussed above, the probabilities of host-switches may often be influenced by the phylogenetic relatedness between existing and potential hosts. If colonization occurred in a sequence, with switches occurring from each host to its closest unoccupied relatives, then terminal regions of the parasite phylogeny might show a higher degree of congruence than if host-switching were random (Fig. 3). In this type of scenario, outcomes would depend on host tree topology and the position of initial colonization (Fig. 3). Sequential colonization (subsequent to divergence of the host lineages) in symmetrical host trees would always produce parasite trees that are incongruent. This is except for the final pair of hosts occupied by each invading parasite lineage, if these are derived from a common ancestor from which no other parasitized lineages arise. In the case of pectinate host phylogenies, an initial invasion of a derived host lineage by the parasite would produce a colonization sequence leading to complete incongruence. However, colonization of terminal sister host lineages from a third, basal lineage could result in a congruent section of parasite phylogeny (see Fig. 3).

It is perhaps significant that parasite switching between a source host and its unoccupied sister lineage is never incongruent (Fig. 4) and might be common, given the close host genetic relationship (see above). The possibility of 'pseudospeciation'
Fig. 5. Conversion of incongruent parasite trees by host extinction (parasite trees above, host trees below; host associations indicated by stippled lines; all nodes represent speciation events). A. Congruent trees: any combination of host extinctions would leave the remaining trees congruent (the same applies to any fully congruent host–parasite phylogeny). B. Incongruent trees: extinction of any one of the terminal three host branches would render the host and parasite trees congruent.

Fig. 4. Detectability of host-switches. A–B. Host trees shown below (bold), with parasite speciation/colonization sequence to one side. Resulting parasite trees shown above with host associations indicated by vertical stippled lines. All nodes represent speciation events. A. A parasite clade colonizes (horizontal stippled line) the sister lineage of its original host which is unoccupied due to an earlier extinction. The resulting host and parasite phylogenies are congruent and could be interpreted as evidence of full speciation. B. The same scheme without extinction: the parasite phylogeny is now congruent with that of the host. In an analysis this incongruence might be explained by a host-switch or duplication and extinction events. Note: host-switches onto the unoccupied sister lineage of a source host are never incongruent. If colonization is influenced by host relatedness and potential competitive interactions with other parasites (see text), this class of host-switch might be relatively common.

by this mechanism has previously been suggested by Hafner & Nadler (1988, 1990). Host-switching may also be associated with the extinction of established parasite lineages (Barker, 1991, 1996). For example, colonization of a host might be more likely if its own parasites had gone extinct due to some stochastic process (leaving an 'empty niche' available), or a colonizing parasite might competitively exclude the previous incumbents. It seems possible, then, that the combination of these two putative processes, the tendency for host-switches to occur between the most closely related hosts (sister lineages) and of colonization to be correlated with extinction of established parasite lineages, could produce a significant degree of false congruence (see Fig. 4).

Host extinction (with corresponding extinction of parasite lines) might also have a ratchet-like tendency to maximize congruence over time. Such events could remove whole sections of the parasite tree, whether incongruent or congruent. However, while local extinctions within a congruent section of host and parasite tree cannot make the residual trees incongruent, they could remove branches from incongruent trees to make these congruent (Fig. 5).

Another scenario in which colonization might produce congruent host and parasite phylogenies is where a parasite lineage colonizes a group of hosts and then speciates in a time sequence influenced by the phylogenetic relatedness of host species (Fig. 6) (effectively, the probability of parasite speciation at any one time decreases with increasing phylogenetic relatedness of hosts).

These considerations call into question a basic assumption of historical reconstruction of host–parasite relationships: that in a host and parasite associated by colonization, no more phylogenetic congruence is expected than if the parasite distribution and phylogeny were independent of the host phylogeny. It is worth noting that, given the vacuum of knowledge concerning what processes are important in past host-parasite relationships, any decision to ignore false congruence as being 'unlikely' is an arbitrary one.

Asynchronous cospeciation. It has previously been recognized that, in a cospeciating system, corresponding speciation events in the host and parasite phylogenies are not necessarily contemporaneous (Hafner & Nadler, 1990; Hafner & Page, 1995). An
Fig. 6. Congruent host and parasite phylogenies resulting where the speciation probability of host-switching parasite lineages per unit time depends on the genetic relatedness of the colonized hosts. Below: parasite colonization and speciation sequence shown to right of host tree (bold). Above: resulting parasite phylogeny, with host associations indicated by vertical stippled lines. Closed circles indicate parasite speciations (numbering identifies individual nodes) and horizontal stippled lines colonizations. A hypothetical parasite lineage invades members of a host clade during a short episode of colonization: individual sublineages speciate immediately when switching between relatively distantly related hosts and with a delay when switching between closely related hosts.

underlying assumption in historical reconstructions is that where parasite lineages isolated in diverging host lineages speciate, the speciations will be time-ordered with those of the host. However, it is possible to consider biologically plausible scenarios where this might not be the case. For example (Fig. 7): a single parasite species occurs in a host lineage, which radiates into several lineages. Initially the parasite does not co-speciate, retaining some degree of gene flow between populations on the different hosts. As these host clades diverge phenotypically and ecologically at different rates, their parasite populations become isolated and speciate in a different order from their hosts. Under some sets of conditions, association by descent, with pairwise (but asynchronous) host–parasite cospeciation, could therefore produce incongruent host and parasite phylogenies.

Fig. 7. Asynchronous cospeciation. Below: host tree (bold) with speciation sequence of an associated parasite lineage shown to right (parasite speciation events indicated by closed circles). Above: resulting parasite phylogeny (numbering identifies individual parasite nodes and host associations are shown by stippled lines). Each speciating host lineage inherits the ancestral parasite lineage (association by descent). Parasite populations in individual hosts then speciate in random order, producing incongruent host and parasite phylogenies.

Widespread parasites. Speciation within a host lineage that occurs without speciation in its parasites (Fig. 8) can result in 'widespread parasites' (Brooks, 1990; Page, 1994, see also Page, 1990b) that are

Fig. 8. Widespread parasite species. Below: parasite speciation sequence superimposed on host phylogeny (bold) and to its right. Parasite speciation events indicated by closed circles. Above: resulting parasite phylogeny (host associations shown by stippled lines). Host speciation where the parasite does not cospeciate (i.e. at the basal host node) results in the same parasite taxon being present in basal and derived host lineages (i.e. lineage 1 retains a common genetic identity). Such a distribution could also be explained by host-switching or duplication/extinction. Note, also, that the nodes in the parasite tree would have the same timing as in the present scheme if lineage 1 had colonized the basal host lineage subsequent to the two speciations.
Fig. 9. Cospeciation with host assemblages. Below: a hypothetical scenario where a single parasite initially occurs in host sister lineages. Both host (bold) and parasite lineages (right) then undergo coincident rounds of speciation due to prevailing ecological/geographical factors (arrows) (e.g., habitat fragmentations which successively isolate descendants of both sister lineages and the parasite). Above: resulting host–parasite phylogenies are partially incongruent (parasite phylogeny shown above host phylogeny; all nodes in phylogenies represent speciation events; host–parasite associations indicated by stippled lines).

inherited by unrelated host species. For example, in Fig. 8 the first branching of the host tree is not mirrored by the parasite. Subsequently two cospeciation events occur with two parasite lineages diverging from another, which retains gene flow and genetic identity with populations in the basal host line. This scheme results in one parasite species being present in both the basal and a derived host, a distribution which might equally be explained by host-switching or duplication/extinction events. In effect, scenarios of this type are a special case of the asynchronous cospeciation described above. Rather than parasite speciation responding to host speciation with a variable time delay, it occurs either contemporaneously or not at all.

Fig. 10. Effects of coalescent host lineages on historical reconstructions. A. A symmetrical host tree (bold) undergoing coalescence of two non-sister lineages; the parasites (tree shown to right) have previously cospeciated with the hosts and the coalescent host lineage inherits taxa from both of its 'parents'. B–C. If the lineage coalescence went unrecognised, then phylogenetic analysis of the hosts would produce two possible trees. These are shown in B and C (host trees shown below in bold, parasite trees above; host–parasite associations indicated by stippled lines). Each host tree would support a different, spurious, historical scenario.

Cospeciation with host assemblages. A number of historical processes (host-switching, extinction, independent host speciation) may lead to individual parasite species being present in different, unrelated host lineages. Indeed host distributions of this nature are indisputably common in the present day. If parasites had a tendency to coevolve with such non-monophyletic host assemblages, rather than individual lineages, this would have the potential to produce complex, mixed patterns of congruence and incongruence without further host-switching or duplication/extinction. Fig. 9 shows an example where a parasite clade tracks assemblages of non-monophyletic hosts as these speciate in response to prevailing geographical/ecological isolating barriers. In such a case subsequent analysis might suggest a pattern of partial incongruence, even though host-switching and parasite extinctions had not occurred (Fig. 9).
Reticulate evolution. While the possibility of coalescence between lineages is often ignored in historical phylogenetic studies (but see Brooks, 1988), interspecies hybridization and gene introgression occurs in many modern-day animal and plant taxa. Speciation involving polyploidization of hybrids (allopolyploidization) is a significant process in plants, and can also occur in certain animal groups (Kobel, 1996; Chenuil, Galtier & Berrebi, 1999). Hybrid host lineages could potentially inherit all, or a subset, of the parasites from their parental lineages (e.g. Tinsley & Jackson, 1998; Jackson, Tinsley & Kigolo, 1998). Similarly, hybrid parasites might remain infecting either all, or a subset, of the hosts of their parental lineages. In such situations phylogenetic analysis of the hybrids (if hybridization had occurred between non-sister groups) and their close relatives would produce spurious cladistic reconstructions, which could be congruent or incongruent with those of the associating host/parasite, but which would result in misleading historical reconstructions (Fig. 10). The same principle might apply to fully coalesced lineages as to those which underwent a degree of gene introgression and then diverged (here spurious phylogenies could be constructed if morphological or molecular characters included those affected by introgressed genes). Unfortunately, hybridization and lineage coalescence would be expected to occur sporadically in time and not be easily detectable after the event.

HISTORICAL RECONSTRUCTION

Methodologies

A number of different approaches to the historical reconstruction of host-parasite relationships have been adopted. At their simplest, such analyses are limited to an intuitive identification of possible evolutionary scenarios from a visual examination of host-parasite phylogenies (Brooks, 1979). This has the disadvantage of lacking an explicit, repeatable and exhaustive methodology. Brooks Parsimony Analysis (BPA) (Brooks, 1981, 1987, 1988, 1990) was the first numerical technique to be proposed. It involves mapping of the parasite tree, which is coded as a multistate transformation series, onto the host phylogeny. Parallelism might be interpreted as host-switches or either host-switches or speciation/extinction (Brooks, 1988; Paterson & Gray, 1997). BPA has been criticised (Page, 1993; Paterson & Gray, 1997) for producing spurious reconstructions in some circumstances, and for overemphasizing the importance of host-switching. Brooks (1988), however, seems to emphasize that BPA is properly used as a relatively neutral tool to examine instances of incongruence. These may be attributed to host-switching or other historical processes on intuitive grounds, or on the basis of biogeographic and ecological evidence. Other methods were developed from reconciliation analysis (Page, 1993), a technique originally used in vicariance biogeography (Nelson & Platnick, 1981). Here a parasite tree is reconciled with the host tree in a way that maximizes the number of cospeciations (accounting for incongruence with hypothetical speciations and extinctions). Host-switching is not recognized and can only be inferred by ad hoc criteria (Page, 1993). A further development of reconciliation-based methods, TREEmap analysis (Page, 1994), incorporates searches for optimal reconstructions that do involve host-switches. This may postulate mutually incompatible host-switches, where one switch requires that the other had occurred between hosts not existing at the same time (Ronquist, 1995; Page & Charleston, 1998). More recently, analytical methods that solve the later problem have been proposed (Ronquist, 1995, 1998; Page & Charleston, 1998; Charleston, 1998). These attach relative costs to different events (such as host-switching or speciation), and then seek optimal reconstructions minimizing these costs.

Practical application of numerical methods

The range of host-parasite systems in which some form of numerical reconstruction has been attempted is now relatively varied (see, for example, Brooks, 1988; Brooks & O’Grady, 1989; Hafner & Page, 1995; Clayton, Price & Page, 1996; Hoberg et al., 1997; Page et al. 1998; Paterson & Gray, 1997; Siddall, 1997; Paterson & Poulin, 1999) and includes digenean, cestode, crustacean, insect and viral parasites/pathogens in vertebrates. The main reconstruction techniques applied have been BPA and reconciliation or TREEmap analysis, while a few studies have compared both methods. Hoberg et al. (1997) analysed highly incongruent relationships in a cestode-bird system with BPA and reconciliation analysis. In contrast, Paterson & Gray (1997) applied BPA and TREEmap analysis to a mainly congruent louse-bird system with a significant degree of incongruence. Both BPA and reconciliation-based methods indicated the general importance of colonization for the cestode-bird system (Hoberg et al. 1997) and of association by descent for the louse-bird system (Paterson & Gray, 1997). However, alternative reconstruction techniques implied quite different evolutionary scenarios for the same host-parasite systems. Hoberg et al.’s study found that BPA supported a history of recent colonization, whereas reconciliation analysis explained the host-parasite relationships by more ancient host-switching and cospeciation (with the postulation of higher numbers of speciation and extinction events). Paterson & Gray’s TREEmap analysis inferred ancient duplication of parasite lineages, with a relatively large number of subsequent extinction events. Their BPA suggested a general history of
Inference of host-switching
cospeciation, but differed in the way possible instances of host-switching were treated and, the authors noted, made less explicit assertions about the history of parasite lineages.

Different numerical methods may therefore vary markedly in the detail of the scenarios proposed and in the number, age and type of historical processes that are inferred.

**Host switching versus extinction**

In practice, either parasite host-switching or duplication/extinction is usually invoked to explain incongruent host and parasite phylogenies. However, it has been widely recognized that the patterns produced by these two processes are potentially confounded (Brooks, 1988; Page, 1993, 1994; Barker, 1996). Fig. 2 illustrates how different colonization or duplication/extinction scenarios may result in the same incongruent host–parasite phylogenies and associations. Within the context of BPA or reconciliation-based analyses (e.g. TreeMap, Page, 1994), the aim is to produce an explanation of incongruence with a minimum number of plausible evolutionary events. The types of event that are implicated may reflect the analytical method used. They may also depend on interpretation by the analyst: for example, reconstruction might be selected over another on intuitive grounds, or on the basis of additional information (e.g. geological or biogeographical data). An initial reservation with such approaches is that it is unreasonable to expect that the evolution of host–parasite associations will always have proceeded through the minimum possible number of steps. Whilst a parsimony criterion may be suitable for phylogeography reconstructions from large sets of characters, it is arguably an unsuitable basis for estimating a one-off historical sequence from what is, effectively, a single complex character (the incongruence between two phylogenies). Further, the different events which could be used to explain incongruence result from fundamentally distinct biological processes (host-switches, speciation, extinction) (Ronquist, 1995) whose relative probabilities may be unknown. As a reconciliation method, TreeMap analysis maximizes cospeciation and explains incongruence by parasite duplication/extinction. At the same time it allows searches for optimal host-switches that account for incongruence with reduced numbers of speciations and extinctions. The balance between assigning incongruent patterns to host-switching or to speciation/extinction (i.e. selection of particular reconstructions) must ultimately be determined by the investigator. BPA may, in some senses, be considered freer from a priori assumptions than reconciliation methods. It does not, as its basis, attempt to maximize one process (cospeciation) by the postulation of certain types of event (speciation/extinction). However, there would usually be a working assumption that congruence is due to cospeciation. Also, where instances of incongruence are identified (and no evidence from other sources is available) their cause can only be assigned by intuitive criteria. As discussed above, alternative numerical methods tend to attribute incongruence to deep parasite nodes with a large number of subsequent extinctions (e.g. TreeMap analysis) or to more recent host-switching (e.g. possible interpretations of BPA). Given that departures from phylogenetic congruence arise from different classes of event whose relative probabilities are unknown, the information contained within incongruent patterns is inherently ambiguous. A real danger exists, therefore, that any arbitrary assumptions and/or biases underlying analytical procedures may be reflected directly in the scenarios which are inferred. Some recent studies have attempted to address the problem that distinct historical processes might occur with different probabilities. Ronquist (1995, 1998) and Charleston (1998) have emphasized the assignment of costs (or weights) to particular events (e.g. host-switching, duplication or extinction). These costs relate to the relative probabilities of the events: so that higher costs would be attached to more improbable events and reconstructions would, in some way, minimize the overall cost of an observed pattern in terms of the different event types. However, even if such methodologies are applied, the problem remains as to how the relative costs may be estimated (Charleston, 1998). A possibility is that modelling techniques (Nee et al. 1994a, b) could be used to determine speciation and extinction rates from molecular phylogenies (Page, 1996). Here, any uncertainty in the estimates would be compounded with that arising from the analysis of congruence/incongruence itself. Also (as discussed above), it may not be possible to distinguish the relative probabilities of different types of parasite speciation (cospeciation, duplication and speciation associated with host-switching), each of which is associated with a different historical process. Huelsenbeck, Rannala & Yang (1997) modelled a host-parasite association (the gophers and chewing lice of Hafner et al. 1994) after estimating host speciation and extinction rates with the methods of Nee et al. (1994a, b). They assumed that parasites only speciated and became extinct alongside their hosts, and allowed only one extant parasite lineage per host. The actual host-switching rate was estimated by comparing the observed deviance of the host and parasite trees to that produced in simulations using different host-switching rates. However, this approach (and that of Huelsenbeck, Rannala & Larget, 2000) was undertaken at the expense of ignoring parasite duplication and extinction as a source of incongruence.

A further complication is that in cospeciating, non-symmetrical host and parasite trees, the phylo-
model of potential host–parasite association and that other plausible confounding processes are often not taken into account. In practice, only two sources of incongruence are usually considered: host-switching and parasite speciation and extinction (e.g. Page, 1993; Hoberg et al. 1997; Charleston, 1998). It is implicit that parasite cospeciation is time-ordered with host speciation (e.g. Brooks, 1979) and occurs with a monophyletic host lineage. Further, it is assumed that colonization produces congruent host and parasite trees no more often than if host-switches were random and occurred with equal probability between any pair of contemporaneous hosts in an analysis (i.e. the phylogeny and host associations of a colonizing parasite are expected to be independent of the host phylogeny). There is thus a tendency for congruent patterns to be attributed only to association by descent. As outlined above, many of these assumptions are questionable. A number of effects have been identified which might lead to spurious congruence between host and parasite phylogenies after a history of colonization. False incongruence can result from several processes (asynchronous cospeciation; parasite cospeciation with host assemblages; independent host speciation; reticulate evolution scenarios) whose effects may be indistinguishable from each other and from the consequences of host-switching and duplication/extinction.

Timing

The preceding discussions primarily relate to situations in which the ages of branching points within host and parasite phylogenies cannot be considered. Brooks (1987) suggested that instances of spurious congruence could be identified by evidence for the age of host–parasite relationships derived from historical geology. Comparison of the timing of nodes in both host and parasite trees using molecular techniques (Hafner & Nadler, 1990; Hafner & Page, 1995; Page, 1996; Page et al. 1998) might also eliminate many otherwise plausible scenarios from a historical reconstruction. The potential difficulties of inferring lineage age where there might have been local variations in evolutionary rate (perhaps associated with host-switching in the parasite) are beyond the scope of this review. However, if relative branching times can be established this may allow some colonization and association by descent scenarios to be distinguished by the ages of corresponding nodes in the host and parasite phylogenies. Synchronous divergence might be interpreted as evidence for cospeciation (Page et al. 1998). Host-switches, on the other hand, would often be associated with more recent parasite nodes, and duplication/extinction scenarios with more ancient parasite nodes (Page et al. 1998; see also Fig. 2). Procedures for assessing the timing of nodes in host
and parasite trees have been considered by Hafner & Nadler (1990), Hafner & Page (1995), Page (1991, 1996) and Huelsenbeck et al. (1997). However, even if a temporal comparison of branching events can be achieved, some potential problems remain. For example, colonizations by parasites against a background of duplication and extinction, might, in some circumstances, be associated with nodes more ancient than the corresponding host node (see Fig. 2). Knowledge of node timings cannot be used to distinguish incongruence due to host-switching from false incongruence due to some of the processes described above (e.g. asynchronous cospeciation, coalescent evolution scenarios). It would also not resolve some of the ambiguities which might arise from the presence of 'widespread' parasites (see Fig. 8).

CONCLUSIONS

Numerical reconstructions of historical host-parasite associations usually allow for up to four types of event in parasite lineages: host-switching, cospeciation, duplication and extinction. They assume that association by descent occurs with monophyletic host lineages and that, during episodes of cospeciation, parasite speciation takes place in the same temporal sequence as those of their hosts. In practice, host-switching or parasite duplication and extinction are usually invoked to explain incongruence, and significant congruence is attributed to cospeciation. This model may be unrealistically narrow for several reasons: (1) Some effects, including the influence of host phylogenetic relationships on parasite colonization and speciation, may produce spurious congruence. (2) Incongruence may occur despite association by descent if parasite lineages speciate in a different temporal sequence to their hosts (i.e. asynchronous cospeciation). (3) Host speciation without corresponding parasite speciation may result in the same parasite species being inherited by distantly related hosts. (4) The cospeciation of parasite lineages with non-monophyletic host assemblies (without host-switching or duplication/extinction) could be a common source of mixed congruent/incongruent patterns, given the frequency of such associations in nature. (5) Historical coalescence of host or parasite lineages, if undetected, would result in spurious reconstructions.

Even if a limited model of host–parasite historical association is accepted, and incongruence can only be produced by host-switching or duplications and extinctions, these processes are confounded. Numerical methods, which, within the context of the technique applied, aim to produce a reconstruction involving the least number of plausible evolutionary events, do not deal with the problem that different events (e.g. host-switching, speciation associated with host-switching, cospeciation, speciation within a host lineage, or extinction) would occur with different probabilities. The advantage of methods seeking to assign relative 'costs' to distinct events (which are then minimized in a historical reconstruction) is limited by the difficulty of estimating such quantities. Comparison of node timings in both parasite and host trees might distinguish between different historical scenarios involving cospeciation, colonization or duplication/extinction, but some sources of incongruence would still be potentially confounded with others.

The inference of parasite host-switches from incongruence between host and parasite phylogenies, then, is seriously compromised by alternative explanations for incongruence and by the possibility that colonization does not always produce incongruence.

ACKNOWLEDGEMENTS

I am very grateful to Professor R. C. Tinsley and to three referees for suggestions which improved the original manuscript and for support from BBSRC research grant 7/S12169.

REFERENCES


