Parasitol Res (2003) 91: 497–499
DOI 10.1007/s00436-003-0987-2

ORIGINAL PAPER

David A. Morrison · Erland L. Ljunggren · Jens G. Mattsson

The origin of *Sarcoptes scabiei* in wombats

Received: 21 May 2003 / Accepted: 30 July 2003 / Published online: 7 October 2003

Abstract In 2002, Skerratt et al. phylogenetically analysed sequence data for several haplotypes of the parasitic mite *Sarcoptes scabiei* from wombat, human and dog hosts in Australia, to test scenarios concerning the origin and diversification of the scabies infections in wombats. Here we note that their substantive conclusions can be called into question by the choice of model used in their phylogenetic analysis, the lack of a root for their phylogenetic trees, and their interpretation of the evolutionary scenario.

Skerratt et al. (2002) report partial sequence data for the mitochondrial small-subunit (12S) ribosomal RNA gene for 23 specimens, representing nine haplotypes, of the parasitic mite *Sarcoptes scabiei* (Acari: Sarcoptidae) from wombat, human and dog hosts in Australia. They analyse these data phylogenetically to test scenarios concerning the origin and diversification of the scabies infections in wombats. There are three points that can be made about their analysis and interpretation that seriously call into question some of their substantive conclusions.

First, Skerratt et al. (2002) analyse their data using the maximum parsimony criterion in order to reconstruct the phylogeny of the haplotypes. This form of analysis assumes that the underlying model for the evolutionary history of the organisms is a dichotomously branching tree (Morrison 1996). While this is the traditional paradigm for modelling the relationships between species, it is highly questionable whether this is an appropriate model for the relationships between individuals within a species. If nothing else, the possibility of reticulate evolution should be explicitly considered when dealing with intra-specific variation.

Under these circumstances, a more appropriate analysis method for their data would be to use the parsimony-splits method of Bandelt and Dress (1993), as implemented in the computer program of Huson (1998). This produces a form of median network (Bandelt 1994) that accurately represents all of the character compatibilities and conflicts in the data. If there is no conflict among the characters then the network will form a dichotomous tree but otherwise there will be reticulation among the branches. The amount of reticulation thus reflects the degree to which the data are or are not tree-like.

The unrooted network shown in Fig. 1 is a complete visual representation of the data shown in Table 1 of Skerratt et al. (2002). There are ten sets of edges in the network, representing the ten patterns of nucleotide variation observed among the haplotypes, with two sets of edges having double length because those patterns occur twice in the data set. The distances along the edges between the haplotypes thus represent the nucleotide distances between them, each unit of distance requiring one nucleotide change. Each path through the network is supported by one set of compatible nucleotide characters, and if there is more than one path connecting two haplotypes then there are several (incompatible) sets of characters indicating alternative relationships among those haplotypes.

Clearly, this network is not particularly tree-like. There is as much conflict for any one set of relationships as there is support, and there will thus be several most-parsimonious trees with each parsimony tree having considerable homoplasy. It is therefore questionable whether these data are suitable for constructing a traditional dichotomous phylogenetic tree. Moreover, the network highlights a point that is obscured in the tree shown by Skerratt et al. (2002), which is that *none* of the internal branches of a dichotomous tree will have

D. A. Morrison (✉) · E. L. Ljunggren · J. G. Mattsson
Department of Parasitology (SWEPAR),
National Veterinary Institute and Swedish University
of Agricultural Sciences, 751 89 Uppsala, Sweden
E-mail: David.Morrison@vmm.slu.se
Tel.: +46-18-674161
Fax: +46-18-309162

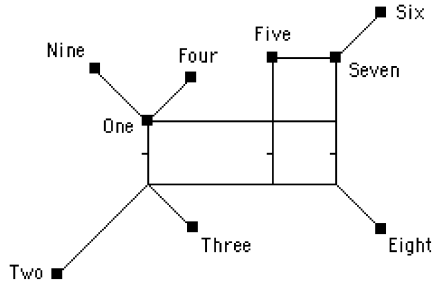

Fig. 1 Parsimony-splits network for the haplotype data of Skerratt et al. (2002). The network is built up of parallelograms forming bands of parallel edges. Each set of parallel edges represents a split (or bipartition) of the nine named haplotypes into two non-overlapping groups, with each edge length being proportional to the number of nucleotide characters that support that particular split. For example, the set of three parallel edges marked with a dash represents the bipartition of the data into the group {Two,Three,Eight} at the bottom of the diagram and the group {One,Four,Five,Six,Seven,Nine} at the top of the diagram, this split being supported by character number 171 of the 12S nucleotide sequence. If the data are truly tree-like then each parallelogram would only have ≤ 2 edges. In order to turn this network into a dichotomous tree, at least one edge from each of the three central boxes needs to be "cut". The various possible combinations of these cuts represent the possible parsimony trees that are compatible with the data (Skerratt et al. 2002 report that there are ten such trees)

unconflicted character support. In particular, the branch that is reported to occur in all of the most-parsimonious trees, which is supported by characters (nucleotides) 59 and 292 (using the numbering of Skerratt et al. 2002), is conflicted by character number 171 of the gene sequence, and this latter character thus requires multiple steps on their tree—this branch may have a high bootstrap value (Skerratt et al. 2002) but it does not have unequivocal support because only two characters support it while one conflicts with it.

Furthermore, since there are multiple equally parsimonious trees it is unreasonable to show an arbitrarily chosen one of them, as this obscures the conflict that must exist in the data. The parsimony-splits network, on the other hand, displays all of the synapomorphy information simultaneously and is thus a more effective summary of the patterns (i.e. it requires only one diagram instead of several). It seems reasonable to conclude from both the parsimony and parsimony-splits analyses that these data do not represent a dichotomous evolutionary history. Certainly, if these sequences do have such a history then these data are completely inadequate to reconstruct it.

Second, Skerratt et al. (2002) suggest that, because their data analysis produced a set of trees with a common branch, there are two "lineages" among the haplotypes. This conclusion is only justified if the phylogenetic trees have an explicit root (Morrison 1996). Unrooted trees, such as those considered by Skerratt et al. (2002), only represent *potential* relationships among lineages—which of these relationships are real lineages will be determined by where the root of the tree is placed. That is, the interpretation of the possible relationships among the lineages will vary depending on the location of the root.

Thus, by explicitly interpreting the relationship of the lineages Skerratt et al. (2002) implicitly place the root of their tree between haplotypes 1 and 5. There is no evidence presented for this rooting of their tree, nor can there be any such evidence for nucleotide data without one or more outgroup taxa in the data analysis, although it is theoretically possible to determine the root of a phylogenetic network in a maximum-likelihood context (Strimmer and Moulton 2000). There are thus no explicit lineages that can recognised in the data of Skerratt et al. (2002) based on the data presented.

Third, Skerratt et al. (2002) claim that their results are compatible with a particular scenario for the origin and diversification of the scabies infections in wombats. However, without a root for their tree the data must *also* be compatible with exactly the opposite scenario. This follows immediately from the fact that the direction of the evolutionary history on an unrooted phylogenetic tree is reversible until the tree is rooted—it is the root that determines the direction of the evolutionary time-course along the tree.

Thus, the analysis presented by Skerratt et al. (2002) is compatible with *both* the scenario that wombats acquired scabies from humans and dogs and the scenario that humans and dogs acquired scabies from wombats. The current data cannot be used to distinguish between these two incompatible hypotheses. What is needed in order to discriminate between them is comparable data from humans and dogs from outside Australia (given that wombats do not occur naturally outside that continent). These data, in conjunction with a tree root, could then provide a definitive answer to the authors' experimental question.

Since both the second and third points above are related to the rooting of the phylogenetic tree, we attempted to explicitly address this issue by sequencing a potential outgroup taxon, for addition to the parsimony-splits network. DNA was extracted from a single individual of *Chorioptes bovis* (Acari: Psoroptidae) collected from a moose (*Alces alces*), using PrepMan Ultra (Applied Biosystems). The PCR amplification procedures followed those of Skerratt et al. (2002), and a product of the correct size was obtained. This was sequenced using an ABI 3100, also using the strategy of Skerratt et al. (2002). The resulting sequence (DDJB/EMBL accession number AY301010) was easily aligned by eye against the existing *S. scabiei* rRNA alignment.

The *Chorioptes* rRNA gene sequence showed variability with respect to one or more of the *Sarcoptes* sequences at 107 (of 331) aligned nucleotide positions, although only eight of these positions are informative in the context of rooting the parsimony-splits network (or rooting a parsimony tree). Of most direct relevance for our purposes, characters 59 and 292 (numbering still according to Skerratt et al. 2002) support the split of the haplotypes into the two groups {*Chorioptes*,

Five,Six,Seven,Eight} {One,Two,Three,Four,Nine}, which corroborates the rooting proposed by Skerratt et al. (2002). However, character 19 supports the split {*Chorioptes*,Nine} {One,Two,Three,Four,Five,Six,Seven, Eight}, which contradicts this rooting. Thus, the weight of evidence is in favour of the hypothesis proposed by Skerratt et al. (2002) but this evidence is neither large nor unequivocal.

The points raised here have important implications for other phylogenetic studies of within-species genetic variation when explicit hypotheses of evolution are being tested. The possibility of reticulate evolution should always be formally tested by evaluating the tree-like nature of the data, and the requirement for a root to the tree in order to assess the time-course of evolutionary processes needs to be unambiguously addressed.

References

Bandelt H-J (1994) Phylogenetic networks. Verh Naturwiss Ver Hamburg 34:51-–71

Bandelt H-J, Dress AWM (1993) A relational approach to split decomposition. In: Opitz O, Lausen B, Klar R (eds) Information and classification. Springer, Berlin Heidelberg New York, pp 123–131

Huson DH (1998) SplitsTree: analyzing and visualizing evolutionary data. Bioinformatics 14:68–73

Morrison DA (1996) Phylogenetic tree-building. Int J Parasitol 26:589–617

Skerrat LF, Campbell NJH, Murrell A, Walton S, Kemp D, Barker SC (2002) The mitochondrial 12S gene is a suitable marker of populations of *Sarcoptes scabiei* from wombats, dogs and humans in Australia. Parasitol Res 88:376–379

Strimmer K, Moulton V (2000) Likelihood analysis of phylogenetic networks using directed graphical models. Mol Biol Evol 17:875–881