

Molecular markers for the phylogenetics of mites and ticks

ROBERT H. CRUICKSHANK

Ecology and Entomology Group, Soil, Plant and Ecological Sciences Division, Lincoln University, PO Box 84, Lincoln, Canterbury, New Zealand; cruicksr@lincoln.ac.nz

Abstract

Molecular methods are becoming increasingly important in systematic acarology. In this review I describe the properties of the ideal molecular marker and compare these with genes that have been used for phylogenetic studies of mites and ticks. The second internal transcribed spacer of the nuclear ribosomal gene cluster (ITS2) and the mitochondrial protein-coding gene cytochrome oxidase I (COI) together provide a powerful tool for phylogenetics at low taxonomic levels. The nuclear ribosomal genes 18S and 28S rDNA are equally powerful tool for phylogenetics at the deepest levels within the Acari. What appear to be lacking are markers that are useful at intermediate levels. The mitochondrial ribosomal genes 12S and 16S rDNA have not lived up to their initial promise. I suggest some nuclear protein-coding genes that may be suitable alternatives. Methods for collection and storage of mites for molecular work, DNA extraction and phylogenetic analysis are also briefly discussed.

Key words: Acari, molecular systematics, phylogenetics

Introduction

Molecular methods have revolutionised insect systematics (Roderick, 1996; Caterino *et al.*, 2000) and are increasingly being applied to mites and ticks (Navajas & Fenton, 2000). To date, most phylogenetic studies of the Acari have focussed on economically important groups such as the Ixodida, Psoroptidae, Tetranychoida, Eriophyoidea, Phytoseiidae and *Varroa*. However, phylogenetic studies of groups such as the Parasitengona, Avenzoariidae and *Steganacarus*, which address more basic evolutionary questions, have also recently begun to appear. The vast diversity of mites makes them particularly suitable for these kinds of studies and they represent a unique opportunity to test many evolutionary hypotheses. Such studies rely on robust phylogenies, but these are lacking for most groups of mites. This may be one reason why the mites have been overlooked by the wider community of evolutionary biologists.

One problem that has arisen in the field of insect molecular systematics is that entomologists working on different groups have used different molecular markers for their phylogenies. This has made it difficult to compare and combine results from different studies and has lead one author to describe the field as a “thriving Tower of Babel” (Caterino *et al.*, 2000). This has yet to happen in acarology, and hopefully the more tight knit nature of the acarological community will continue to prevent it. A toolkit of molecular markers suitable for use on mites and ticks is beginning to emerge. The purpose of this paper is to review the development of this toolkit and to suggest how it can be deployed at different taxonomic levels. I will not attempt to review every case in which molecular data have been used to discriminate between species, or those that take a purely population genetics

approach. These have been reviewed elsewhere (Mozes-Koch & Gerson, 1998; Navajas & Fenton, 2000). I will focus instead on studies in which molecular markers have been used to reconstruct phylogenies, including intraspecific phylogenies. DNA sequencing technology has evolved to a point where user friendly kits are available for the entire process and it is no longer necessary to have specialised molecular expertise in order to produce molecular data for phylogenetic analysis. I hope this review will encourage the wider use of molecular techniques in systematic acarology.

Collection and storage of specimens for molecular work

Evidence from insects suggests that the most effective method for preserving specimens for molecular work is ultracold freezing (-80°C) of live specimens (Post *et al.*, 1993; Reiss *et al.*, 1995; Dillon *et al.*, 1996), however, this is not always practicable. An acceptable alternative is storage in 100% ethanol. Mites from other media should be transferred to 100% ethanol and stored at -20°C as soon as possible. Lower concentrations of ethanol give variable results and should be avoided if at all possible, however, mites stored in ethanol as low as 70% for many years can give good yields of DNA. Drying is not generally good for DNA and therefore parasitic mites collected from museum specimens may not be suitable for molecular work.

DNA extraction

Klompen (2000) found it necessary to use 2-5 individuals in each extraction whereas Anderson & Trueman (2000) used only leg tissue dissected from individual mites. If possible it is preferable to extract DNA from single individuals to prevent mixing of distinct genotypes, particularly if there is any question about the identity of the mites. This should be possible using all but the very smallest of mites. Extracted DNA should be stored at -80°C.

When extracting DNA it is usual to slide mount a few individuals as voucher specimens so that they are available for examination subsequent to DNA sequencing (Ruedas *et al.*, 2000). Ideally the voucher specimen should be the individual that is sequenced but this individual is usually completely destroyed during the extraction process. However, non-destructive methods of DNA extraction are available (Rose *et al.*, 1994; Phillips & Simon, 1995). Cruickshank *et al.* (2001) have adapted the DNeasy Tissue Kit from Qiagen (<http://www.qiagen.com/>) for non-destructive DNA extraction of lice by increasing the duration of the lysis step. This method has also been applied to midges (*Culicoides* sp.) and specimens have been successfully identified to species after DNA has been extracted (unpublished results). The method has also been tested on mites and complete grinding of the specimen was only found necessary for the very small mites, although there was some loss of yield in larger mites (unpublished results).

Properties of the ideal molecular marker

Single-copy (or multiple homogenous copies)

The ideal molecular marker should be present as a single copy in the haploid genome. If more than one copy is present it is possible that sequences obtained from different individuals in the phylogeny will be from different (paralogous) copies of the gene. The obvious solution is to use only single-copy genes, however, it is not always obvious in advance which genes are present as a single copy, and single-copy genes may be difficult to amplify. Multiple-copy genes may also be useful if

all of the copies have the same sequence so that it does not matter which copy is sequenced. Such genes are generally easier to amplify and give a higher yield. There are two groups of genes that fall into this category, mitochondrial genes and nuclear ribosomal genes, and these are by far the most commonly used genes for phylogenetics. Mitochondrial genes occur in large numbers in each cell but usually all of these copies have the same sequence due to the population bottleneck which mitochondria pass through when they are transmitted to the offspring via the egg (Simon *et al.*, 1994). Other mechanisms, collectively known as concerted evolution, are responsible for keeping all of the copies of the nuclear ribosomal gene cluster the same (Hillis & Dixon, 1991), although heterogeneity of the second internal transcribed spacer region of the nuclear ribosomal gene cluster has been reported in ticks (Rich *et al.*, 1997).

Easy to align

Due to deletion and duplication of small regions of DNA the length of a gene can vary between taxa and sequences must be aligned prior to phylogenetic analysis. For protein-coding genes this is usually relatively straightforward since gaps generally occur in groups of three, corresponding to whole codons, otherwise there would be a frameshift mutation. For ribosomal genes, which are not translated into proteins and do not have this three base-pair codon structure, alignment can be much more difficult. Uncertainty in the alignment will lead to uncertainty in the phylogeny. There are a number of ways to deal with problematic alignments including removal of regions with ambiguous alignments (Gatesy *et al.*, 1993) or the use of secondary structure information for aligning the sequences (Kjer, 1995).

All sites are equally free to vary

Due to the codon structure of protein-coding genes only about a third of the sites are free to vary without altering the sequence of the protein. Mutations that alter the protein are likely to have a deleterious effect on its function and therefore such changes, which are called non-synonymous substitutions, are far less likely. For this reason only a third of the sites sequenced are likely to be phylogenetically informative, at least for closely related taxa. Although maximum likelihood methods can account to some extent for different rates of substitution at different sites, it can be difficult to reconstruct the phylogeny of a group if the rates are very different or at least not easily partitioned into different rate classes (e.g. the first, second and third positions within a codon).

Substitution rate is high enough to provide sufficient number of informative sites

Different markers evolve at different rates and it is important to choose a marker with a substitution rate appropriate for the use to which it is being put. For closely related species a gene with a high substitution rate will be needed to ensure that enough mutations have accumulated during the short time over which the taxa have been evolving.

Substitution rate is low enough to avoid excessive numbers of multiple substitutions

If too fast a gene is used then there will be sites at which two or more substitutions have occurred. Since subsequent substitutions mask previous ones it will become increasingly difficult for most methods of phylogenetic analysis to reconstruct the tree. This is called saturation. This problem can be exacerbated by base composition bias since this makes it more likely that the second mutation at a particular site will be a reversion to the original state (e.g. A→T followed by T→A). For protein-coding genes it may be the case that the synonymous substitution rate is too high even though very few non-synonymous substitutions have occurred.

Roughly equal base compositions

If base compositions are biased then the chances of the same mutation occurring more than once on the tree will increase. This will increase the amount of homoplasy in the data set. This can be a problem for mitochondrial genes, which tend to be AT rich. For example, Navajas *et al.* (1996a) found that the mitochondrial gene cytochrome oxidase I (COI) in Tetranychidae has an AT content of 75% (95% at third codon positions).

No variation in base compositions among taxa

Variation in base composition among taxa can be a problem for some methods of phylogenetic analyses that may group taxa together due to similarity of base composition rather than genuine shared ancestry (Galtier & Gouy, 1995). Navajas *et al.* (1996a) reported significant differences in base composition for COI in different tetranychid genera.

Universal primers are available

For practical reasons it is much easier to use primers which work on a wide range of taxa than ones with a much more restricted range. Primers with a wide taxonomic range are likely to be highly conserved and there is much less chance of a mutation at the primer site, which may prevent the amplification of some taxa in the study, although this can still occur (Fenton *et al.*, 1997; Yli-Mattila *et al.*, 2000). However, primers should not be too universal since this could lead to amplification of gut contents, symbionts, parasites or environmental contamination (Yli-Mattila *et al.*, 2000). While it should be obvious from the sequence when this has happened it is preferable to use primers which are at least arthropod specific when amplifying DNA from mites.

Widely used for maximum combinability of results

Using genes that have been used widely in other studies adds to the value of the data since they can be used in conjunction with sequences from other studies to construct broader scale phylogenies or in comparative studies of molecular evolution.

How do real genes measure up?

No single molecular marker can satisfy all of the conditions listed above, and different markers are more suitable for different purposes. For example, ribosomal genes may be difficult to align but they are likely to have more informative sites than protein-coding genes. Even if the unalignable regions are removed this may still leave more informative sites than are present in a protein-coding gene of the same length. A number of different markers have been used in mites and ticks, and all of them have their problems, but by choosing the marker most appropriate for the task these problems can be minimised before any sequencing is done, saving both time and money.

Non-coding regions

Non-coding regions include the mitochondrial control region (Zhang & Hewitt, 1997), introns of nuclear genes (Friesen, 2000), and the internal transcribed spacer regions of nuclear ribosomal genes. Since these regions are under very little selection pressure they tend to evolve very quickly and are usually only of use within species or for closely related species. Due to their high substitution rate it is common to find heterozygotes for these genes and it is often necessary to clone PCR products prior to sequencing.

ITS2

The non-coding region which has been most widely used in mites is the second internal transcribed spacer region of the nuclear ribosomal gene cluster (ITS2). This cluster consists of three genes (18S rDNA, 5.8S rDNA and 28S rDNA) which are transcribed into RNA but not translated into protein. These RNA molecules form parts of the ribosome, a cellular structure involved in making proteins. The three rDNA genes are transcribed as a single transcript of RNA transcript separated by two regions (ITS1 and ITS2) which are subsequently spliced out and serve no further purpose. Since they have no other function these ITS sequences are under very little selection pressure and can accumulate substitutions very quickly. This can be very useful for distinguishing between closely related species. By designing primers within the slowly evolving genes that span the rapidly evolving spacers, primers have been developed which work well in a wide variety of taxa.

In what was perhaps the first ever use of molecular phylogenetics in mites, Navajas *et al.* (1992) used primers in the 5.8S and 28S rDNA genes to amplify the ITS2 region in several species of tetranychid mites and showed that this region is suitable for examining phylogenetic relationships within genera, but that differences between genera were too great to enable correct alignment of the sequences. The relationships within the genus *Tetranychus* agreed closely with morphology but sequences from two sympatric species of *Eotetranychus* from different hosts, which appeared morphologically very similar except for a slight difference in the shape of the peritreme, showed substantial genetic divergence. Navajas *et al.* (1998) used ITS2 sequences in combination with COI to investigate intraspecific variation in *Tetranychus urticae* and found species-wide homogeneity of ITS2 sequences but extensive COI polymorphism (reviewed in Navajas (1998)). Navajas *et al.* (2000) subsequently added more taxa to this data set and these revealed very low levels of variation at the ITS2 locus in this species. Navajas *et al.* (1994) found a similar pattern of variation in the cassava green mite (*Mononychellus progresivus*) with intraspecific diversity being low for both ITS2 and COI, but lower in ITS2. Essig *et al.* (1999) used ITS2 to investigate relationships within the genus *Chorioptes*. Their sequences fell into two distinct groups that correlated well with morphology but not with host suggesting that the sequences in their study represent two species both with low host specificity. Zahler *et al.* (1999) used ITS2 sequences to investigate the relationships within the genus *Sarcoptes*. Although sequence variation was found, this did not correlate well with morphology, host or geography suggesting that all of the sequences in their study belong to a single polymorphic species. Zahler *et al.* (1998) performed a similar analysis for the genus *Psoroptes* but with less conclusive results.

ITS1

While not as widely used as ITS2, ITS1 has similar properties to ITS2 and has been used at the population level to study *Ixodes scapularis* (McLain *et al.*, 1995). Navajas *et al.* (1999) compared the sequences of ITS1, ITS2 and 5.8S rDNA in several species of phytoseiids and concluded that ITS1 is more variable than ITS2 and also more difficult to align. ITS1 was also the only region that showed polymorphism within species, although this was low. 5.8S rDNA was substantially more conserved than either ITS. In a study of the coevolution of *Cecidophyopsis* mites and their *Ribes* hosts, Fenton *et al.* (2000) used primers in 18S and 28S rDNA to amplify a 1629bp fragment of the nuclear ribosomal gene cluster spanning both of the internal transcribed spacers and 5.8S. They also sequenced the homologous region in the hosts and found that the phylogeny of the mites does not closely follow the phylogeny of their hosts.

The problems associated with internal transcribed spacers are those typical of rapidly evolving genes. Failure to amplify some species with universal ITS2 primers has been reported for phytoseiids (Yli-Mattila *et al.*, 2000) and eriophyiids (Fenton *et al.*, 1997) and heterogeneity of ITS2 sequences within individuals has been reported in ticks (Rich *et al.*, 1997). In order to minimise these problems,

internal transcribed spacers should be used only for studies of intraspecific variation or phylogenies of very closely related species. It may be necessary to clone PCR products and sequences should be treated as haplotypes rather than individuals.

Mitochondrial genes

Mitochondrial genes have been widely used in molecular systematics. Their high copy number makes them much more easy to work with than single-copy nuclear genes and their strictly maternal inheritance has been particularly useful at the infraspecific level. Mitochondrial genes fall into two categories; ribosomal genes and protein-coding genes. There are two mitochondrial ribosomal genes, 12S and 16S rDNA, which are not separated by internal transcribed spacers. In general the third positions of mitochondrial protein-coding genes evolve more quickly than the ribosomal genes but it may be possible to use the first and second positions of the protein-coding genes at deeper taxonomic levels. One of the potential problems with mitochondrial genes is that genes can be transferred from the mitochondrion to the nucleus. These genes, called numts (nuclear copies of mitochondrial genes) or mitochondrial pseudogenes, can be amplified and sequenced by mistake and this can lead to erroneous phylogenies. While numts have yet to be reported from mites or ticks they are ubiquitous in other arthropods and molecular systematists working on mites should be aware of them. Strategies for identifying and dealing with numts are reviewed by Bensasson (2001).

16S rDNA

Black *et al.* (1994) used 506 aligned base pairs of 16S rDNA to investigate the phylogeny of ticks at the family level and concluded that this gene is useful for phylogenetics of ticks at or below this level. The bootstrap support for some of their nodes was quite low, indicating that these sequences alone are insufficient to provide full resolution at all nodes in the tree. They concluded that 16S in combination with another gene would give a more fully resolved tree. However, Norris *et al.* (1999) have shown that combining 12S data with these 16S sequences does little to improve the resolution in the tree. Dabert *et al.* (2001) used a 246bp fragment of 16S rDNA, in combination with 41 morphological characters to investigate the phylogeny of avenzoariid feather mites, and obtained an almost completely resolved phylogeny, however, they chose to use decay indices (Bremer, 1988) rather than bootstrap values to assess the degree of support for their nodes, so it is difficult to compare this with other studies. Crosbie *et al.* (1998) were able to use 16S to resolve the relationships among New World *Dermacentor* species and it seems that 16S rDNA is more suited to use at this taxonomic level.

12S rDNA

Norris *et al.* (1999) sequenced domain III of the 12S rDNA gene for 51 species of tick and compared this with previously published (Black & Piesman, 1994) and newly sequenced 16S sequences. They used secondary structure to align both the 12S and 16S sequences and analysed the data sets with and without hypervariable regions removed. They found that removal of the hypervariable regions had no effect on the tree topology. 12S evolved slightly faster than 16S but both genes, either individually or combined were inferior to 18S rDNA (see below). 12S and 16S gave poorly resolved trees with low bootstrap values “portions of which are essentially random” (Norris *et al.*, 1999). This suggests that the initial promise of 16S rDNA was ill founded and that 18S represents a better choice of gene for resolving the relationships of ticks at this level. Norris *et al.* (1999) put this lack of resolution down to the high AT content of the mitochondrial genome causing a large amount of homoplasy and suggest that these genes may be more suited to resolving relationships at the infraspecific level or among closely related taxa. This was borne out by Norris *et al.* (1996) who have used 12S and 16S to investigate intraspecific relationships in *Ixodes scapularis*

and obtained a fully resolved haplotype phylogeny despite a high A↔T bias. Murrell *et al.* (1999) successfully used 12S in combination with COI to resolve relationships among rhipicephaline genera.

COI

COI has a similar range of uses to ITS2 but appears to evolve slightly faster. Navajas *et al.* (1998) found extensive COI polymorphism (but species-wide homogeneity of ITS2) in *Tetranychus urticae*. Anderson and Trueman (2000) used COI to show that *Varroa jacobsoni* is a complex of two morphologically indistinguishable species infesting *Apis cerana*, only one of which (*Varroa destructor*) has been transferred to *Apis mellifera*. Toda (2000) used a 546bp portion of COI to investigate the phylogenetic relationships of Japanese *Panonychus* species and Salomone *et al.* (2002) used a 488bp portion of this gene to examine the relationships between the different members of the oribatid genus *Steganacarus* on the Canary Islands.

Despite its high substitution rate it appears that the utility of COI is not confined to closely related species. Navajas *et al.* (1996b) used this gene to resolve relationships between genera in two subfamilies of the Tetranychidae. This is despite a very high AT content. This contrasts with the poor resolution of 12S and 16S (which are similarly AT rich) at this level in ticks. COI is likely to be on its limit of resolution here.

An alternative approach to using COI for deeper level phylogeny is to eliminate the hypervariable third codon positions from the analysis. This was the approach taken by Söller *et al.* (2001) who used only the first and second codon positions of a 408bp fragment of COI to investigate the phylogeny of the Parasitengona, however, they found that even after removal of third codon positions considerable homoplasy remained and this led to an incompletely resolved phylogeny with low bootstrap values. They also had trouble amplifying all of their taxa, presumably due to mutations at the primer site. This suggests that COI is more suited to use at lower taxonomic levels.

COIII

Kain *et al.* (1999) used a 355bp fragment of the cytochrome oxidase III gene to investigate intraspecific variation in *Ixodes pacificus* and found widespread haplotype sharing between populations (including an island population) with the exception of an apparently genetically distinct population in Utah. It would be interesting to compare the rates of COI, COIII and other mitochondrial genes in the same set of taxa.

Nuclear ribosomal genes

The two nuclear ribosomal genes have been used extensively in molecular systematics and are often used at very deep levels. Within the Acari these genes probably represent the most promising candidates for resolving the deepest relationships within the group.

18S rDNA

Crampton *et al.* (1996) sequenced 260bp of the 28S rDNA gene and 245bp of the 18S rDNA gene of several species of ticks and confirmed that these regions are both more conserved than 16S rDNA and are suitable for phylogenies at the family and subfamily levels. Since these sequences were rather short, Black *et al.* (1997) sequenced the entire 18S gene (1846bp) for a smaller number of taxa. These studies were reviewed by Klompen *et al.* (2000). Norris *et al.* (1999) concluded that 18S alone is superior to 12S and 16S combined for resolution at this taxonomic level.

28S rDNA

Cruickshank and Thomas (1999) used a total of 751bp from two regions of the 28S rDNA gene to examine the phylogenetic relationships within the Dermanyssina. Although they obtained a fully resolved phylogeny, the bootstrap values for many of their nodes were low. However, homoplasy was also low in this data set and this lack of resolution is likely to be due to an insufficient number of informative sites (only 8% of sites were informative according to the parsimony criterion). This suggests that 28S rDNA may be even more useful at a deeper phylogenetic level. Taken together with the evidence from ticks this suggests that, for the Parasitiformes at least, 18S and 28S rDNA are likely to provide a powerful tool for resolving deeper level phylogenies. The phylogenetic utility of these genes has been established in several other groups of terrestrial arthropods including centipedes (Giribet *et al.* 1999a), harvestmen (Giribet *et al.* 1999b) and insects (Whiting *et al.* 1997) and it is reasonable to assume that this will extend to the mites and ticks.

Nuclear protein-coding genes

There has been a move in insect systematics recently towards the use of low copy number nuclear protein-coding genes. These genes have a range of substitution rates and can be applied to a variety of uses. One potential problem with these genes is paralogy. They are often duplicated and it is not always certain that the same copy of the gene has been sequenced in all of the taxa (Cruickshank *et al.* 2001). The only nuclear protein-coding gene that has been used in mites is elongation factor 1 alpha (EF1 α). This gene has been widely used in insects although there has been some lack of overlap between the region sequenced in different studies. Klompen (2000) has assessed the utility of this gene for elucidating relationships among basal Mesostigmata and concluded that it “may be a good molecular marker for Parasitiformes, especially within infraorders, but might be most effective in combination with other molecular markers and/or morphology.”

Since nuclear genes have an intron/exon structure they can also be used at very low taxonomic levels. In a manner similar to the ITS primers described above, highly conserved primers can be designed within exons which amplify sequences which span the highly variable introns. This is known as EPIC (exon primed intron crossing). It has not yet been applied to mites but evidence from the insects suggests that this could be a useful source of highly variable nuclear markers for studies of intraspecific variation.

Other kinds of marker

An alternative approach to sequencing genes is the identification of molecular characters which evolve so rarely that the probability that they have evolved on separate occasions in different lineages is negligible, i.e. they are relatively free of homoplasy (Rokas & Holland, 2000). One such source of characters is mitochondrial gene order. For example, the metastriate ticks all share a particular rearrangement of the mitochondrial gene order (Roehrdanz *et al.*, 2002). This provides very strong evidence for the monophyly of this group since the chances of such a rearrangement occurring twice are almost nil.

Phylogenetic analysis

Molecular methods permit the use of a number of more sophisticated methods of phylogenetic reconstruction than the strictly parsimony based methods which are available for the analysis of morphological data (but see Lewis (2001a) for the application of likelihood models to morphological data). Distance methods, which involve the transformation of sequence data into a matrix of pairwise genetic distances, are often rejected in favour of character-based methods since they involve the loss

of information when constructing of the distance matrix. However, there are often good reasons for using these methods. For example, Navajas *et al.* (1996a) have shown that base composition variation may be a problem for COI sequences in mites. The LogDet distance (Lockhart *et al.*, 1994) can correct for this and may be more reliable than character-based methods for analysing this gene. In addition to maximum parsimony, which will be familiar to morphological systematists, there are several character-based methods for analysing molecular data that are not available to the morphologist. These methods, which include maximum likelihood and Bayesian analysis, rely on statistical models of sequence evolution that describe the rate of substitution at each site. In addition to providing new ways of constructing trees, these methods also allow a statistical approach to testing phylogenetic hypotheses (Huelsenbeck & Crandall, 1997; Huelsenbeck & Rannala, 1997; Goldman *et al.*, 2000; Huelsenbeck *et al.*, 2001; Lewis, 2001b). A detailed discussion of these methods is outside the scope of this review but for more details see Swofford *et al.* (1996), Page & Holmes (1998), Nei & Kumar (2000) and Hall (2001).

Summary

ITS2 and COI together provide a powerful tool for studies of intraspecific variation and phylogenies of closely related species. 18S rDNA and 28S rDNA are equally useful for phylogenetics at the other end of the taxonomic spectrum. What appears to be lacking is a set of markers for use at intermediate taxonomic levels, i.e. between the genus and family. Data from the insects would suggest that 12S and 16S rDNA might be useful for this purpose but evidence to date suggests that these genes are disappointingly poor at resolving phylogenies at this level in the mites, perhaps because of their high AT content (although this is a problem in insects too). The best solution may be nuclear protein-coding genes. With the exception of a single study using EF1 α , these genes have not been used in the mites. Candidate genes that have proved useful in the insects include glucose 6 phosphate dehydrogenase (Sotto-Adames *et al.*, 1994), phosphoenolpyruvate carboxykinase (Freidlander *et al.*, 1996), dopa decarboxylase (Fang *et al.*, 1997) and wingless (Brower & DeSalle, 1998). Perhaps the next step should be a comparative study of the rates of evolution of these and other nuclear protein-coding genes as well as the mitochondrial and ribosomal genes already in use, in a single set of mite taxa spanning a range of taxonomic levels so that the phylogenetic utility of these genes can be assessed.

Acknowledgements

I would like to thank Adrian Paterson and Zhi-Qiang Zhang for their valuable comments on an early draft of this paper. This work was supported by the Marsden Fund of the Royal Society of New Zealand.

References

- Anderson, D.L. & Trueman, J.W.H. (2000) *Varroa jacobsoni* (Acari: Varroidae) is more than one species. *Experimental and Applied Acarology*, 24, 165-189.
- Bensasson, D., Zhang, D.-X., Hartl, D.L. & Hewitt, G.M. (2001) Mitochondrial pseudogenes: evolution's misplaced witnesses. *Trends in Ecology and Evolution*, 16, 314-321.
- Black, W.C. & Piesman, J. (1994) Phylogeny of hard- and soft-tick taxa (Acari: Ixodida) based on mitochondrial 16S rDNA sequences. *Proceedings of the National Academy of Sciences of the United States of America*, 91, 1111-1115.

- America*, 91, 10034-10038.
- Black, W.C., Klompen, J.S.H., & Keirans, J.E. (1997) Phylogenetic relationships among tick subfamilies based on the 18S nuclear rDNA gene. *Molecular Phylogenetics and Evolution*, 7, 129-144.
- Bremer, K. (1988) The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. *Evolution*, 42, 795-803.
- Brower, A.V.Z. & DeSalle, R. (1998) Patterns of mitochondrial versus nuclear DNA sequence divergence among nymphalid butterflies: the utility of wingless as a source of characters for phylogenetic inference. *Insect Molecular Biology*, 7, 73-82.
- Caterino, M.S., Cho, S. & Sperling, F.A.H. (2000) The current state of insect molecular systematics: a thriving Tower of Babel. *Annual Review of Entomology*, 45, 1-54.
- Crampton, A., McKay, I. & Barker, S.C. (1996) Phylogeny of ticks (Ixodida) inferred from nuclear ribosomal DNA. *International Journal for Parasitology*, 26, 511-517.
- Crosbie, P.R., Boyce, W.M. & Rodwell, T.C.. (1998) DNA sequence variation in *Dermacentor hunteri* and estimated phylogenies of *Dermacentor* spp. (Acari: Ixodidae) in the New World. *Journal of Medical Entomology*, 35, 277-288.
- Cruickshank, R.H., Johnson, K.P., Smith, V.S., Adams, R.J., Clayton, D.H. & Page, R.D.M. (2001) Phylogenetic analysis of partial sequences of elongation factor 1 α identifies major groups of lice (Insecta: Pthiraptera). *Molecular Phylogenetics and Evolution*, 19, 202-215.
- Cruickshank, R.H. & Thomas, R.H. (1999) Evolution of haplodiploidy in dermanyssine mites (Acari: Mesostigmata). *Evolution*, 53, 1796-1803.
- Dabert, J. (2001) Phylogeny of feather mite subfamily Avenzoariinae (Acari: Analgoidea: Avenzoariidae) inferred from combined analyses of molecular and morphological data. *Molecular Phylogenetics and Evolution*, 20, 124-135.
- Dillon, N., Austin, A.D. & Bartowsky, E. (1996) Comparison of preservation techniques for DNA extraction from hymenopterous insects. *Insect Molecular Biology*, 5, 21-24.
- Essig, A., Rinder, H., Gothe, R. & Zahler, M. (1999) Genetic differentiation of mites of the genus *Chorioptes* (Acari: Psoroptidae). *Experimental and Applied Acarology*, 23, 309-318.
- Fang, Q.Q., Cho, S., Regier, J.C., Mitter, C., Matthews, M., Poole, R.W., Freidlander, T.P. & Zhao, S.W. (1997) A new nuclear gene for insect phylogenetics: dopa decarboxylase is informative of relationships within Heliiothinae (Lepidoptera: Noctuidae). *Systematic Biology*, 46, 269-283.
- Fenton, B., Malloch, G. & Moxey, E. (1997) Analysis of eriophyid mite rDNA internal transcribed spacer sequences reveals multiple simple sequence repeats. *Insect Molecular Biology*, 6, 23-32.
- Fenton, B., Birch, A.N.E., Malloch, G., Lanham, P.G. & Brennan, R.M. (2000) Gall mite molecular phylogeny and its relationship to the evolution of plant host specificity. *Experimental and Applied Acarology*, 24, 831-861.
- Freidlander, T.P., Regier, J.C., Mitter, C. & Wagner, D.L. (1996) A nuclear gene for higher level phylogenetics: phosphoenolpyruvate carboxykinase tracks Mesozoic-age divergences within Lepidoptera (Insecta). *Molecular Biology and Evolution*, 13, 594-604.
- Freisen, V. (2000) Introns. In Baker, A.J. (ed.) *Molecular Methods in Ecology*, Blackwell, pp. 274-294
- Galtier, N. & Gouy, M. (1995) Inferring phylogenies from DNA sequences of unequal base compositions. *Proceedings of the National Academy of Sciences of the United States of America*, 92, 11317-11321.
- Gatesy, J., DeSalle, R. & Wheeler, W. (1993) Alignment-ambiguous nucleotide sites and the exclusion of systematic data. *Molecular Phylogenetics and Evolution*, 2, 152-157.
- Giribet, G., Carranza, S., Riutort, M., Baguña, J. & Ribera, C. (1999a.) Internal phylogeny of the Chilopoda (Myriapoda, Arthropoda) using complete 18S rDNA and partial 28S rDNA sequences. *Philosophical Transactions of the Royal Society of London*, B354, 215-222.
- Giribet, G., Rambla, M., Carranza, S., Riutort, M., Baguña, J. & Ribera, C. (1999b) Phylogeny of the arachnid order Opiliones (Arthropoda) inferred from a combined approach of complete 18S, partial 28S ribosomal DNA sequences and morphology. *Molecular Phylogenetics and Evolution*, 11, 296-307.
- Goldman, N., Anderson, J.P. & Rodrigo, A.G. (2000) Likelihood-based tests of topologies in phylogenetics. *Systematic Biology*, 49, 652-670.
- Hall, B.G. (2001) *Phylogenetic Trees Made Easy: A How-To Manual for Molecular Biologists*, Sinauer.
- Hillis, D.M. & Dixon, M.T. (1991) Ribosomal DNA: molecular evolution and phylogenetic inference. *Quarterly Review of Biology*, 66, 411-429.
- Huelsenbeck, J.P. & Crandall, K.A. (1997) Phylogeny estimation and hypothesis testing using maximum likelihood. *Annual Review of Ecology and Systematics*, 28, 437-466.

- Huelsenbeck, J.P. & Rannala, B. (1997) Phylogenetic methods come of age: testing hypotheses in a phylogenetic context. *Science*, 276, 227-232.
- 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.
- Kain, D.E., Sperling, F.A.H., Daly, H.V. & Lane, R.S. (1999) Mitochondrial DNA sequence variation in *Ixodes pacificus* (Acari: Ixodidae). *Heredity*, 83, 378-386.
- Kjer, K.M. (1995) Use of rRNA secondary structure in phylogenetic studies to identify homologous positions: an example of alignment and data presentation from the frogs. *Molecular Phylogenetics and Evolution*, 4, 314-330.
- Klompen, H. (2000) A preliminary assessment of the utility of elongation factor-1a in elucidating relationships among basal Mesostigmata. *Experimental and Applied Acarology*, 24, 805-820.
- Klompen, J.S.H., Black, W.C., Keirans, J.E. & Norris, D.E. (2000) Systematics and biogeography of hard ticks, a total evidence approach. *Cladistics*, 16, 70-102.
- Lewis, P.O. (2001a) A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic Biology*, 50, 913-925.
- Lewis, P.O. (2001b) Phylogenetic systematics turns over a new leaf. *Trends in Ecology and Evolution*, 16, 30-37.
- Lockhart, P.J., Steel, M.A., Hendy, M.D. & Penny, D. (1994) Recovering evolutionary trees under a more realistic model of sequence evolution. *Molecular Biology and Evolution*, 11, 605-612.
- McLain, D.K., Wesson, D., Oliver, J.H. & Collins, F. (1995) Variation in ribosomal DNA internal transcribed spacer I among eastern populations of *Ixodes scapularis* (Acari: Ixodidae). *Journal of Medical Entomology*, 32, 351-360.
- Mozes-Koch, R. & Gerson, U. (1998) Separation of acarine species by enzymatic, immunological and molecular methods. *Systematic and Applied Acarology*, 3, 9-18.
- Murrell, A., Campbell, N.J.H. & Barker, S.C. (2000) Mitochondrial 12S rDNA indicates that the Rhipicephalinae (Acari: Ixodida: Ixodidae) is paraphyletic. *Molecular Phylogenetics and Evolution*, 16, 1-7.
- Navajas, M. (1998) Host plant associations in the spider mite *Tetranychus urticae* (Acari: Tetranychidae): insights from molecular phylogeography. *Experimental and Applied Acarology*, 22, 201-214.
- Navajas, M., Cotton, D., Kreiter, S. & Gutierrez, J. (1992) Molecular approach in spider mites (Acari: Tetranychidae): preliminary data on ribosomal DNA sequences. *Experimental and Applied Acarology*, 15, 211-218.
- Navajas, M. & Fenton B. (2000) The application of molecular markers in the study of diversity in acarology: a review. *Experimental and Applied Acarology*, 24, 751-774.
- Navajas, M., Fournier, D., Lagnel, J. & Boursot, P. (1996a) Mitochondrial COI sequences in mites: evidence for variation in base composition. *Insect Molecular Biology*, 5, 1-5.
- Navajas, M., Gutierrez, J., Bonato, O., Bolland, H.R. & Mapangou-Divasse, S. (1994) Intraspecific diversity of the cassava green mite *Mononychellus progresivus* (Acari: Tetranychidae) using comparisons of mitochondrial and nuclear ribosomal DNA sequences and cross-breeding. *Experimental and Applied Acarology*, 18, 351-360.
- Navajas, M., Gutierrez, J., & Lagnel, J. (1996b) Mitochondrial cytochrome oxidase I in tetranychid mites: a comparison between molecular phylogeny and changes of morphological and life history traits. *Bulletin of Entomological Research*, 86, 407-417.
- Navajas, M., Lagnel, J., Fauvel, G. & de Moraes, G. (1999) Sequence variation of ribosomal internal transcribed spacers (ITS) in commercially important Phytoseiidae mite. *Experimental and Applied Acarology*, 23, 851-859.
- Navajas, M., Lagnel, J., Gutierrez, J. & Boursot, P. (1998) Species-wide homogeneity of nuclear ribosomal ITS2 sequences in the spider mite *Tetranychus urticae* contrasts with extensive mitochondrial COI polymorphism. *Heredity*, 80, 742-752.
- Navajas, M., Tsagkarakov, A., Lagnel, J. & Perrot-Minnot, M.J. (2000) Genetic differentiation in *Tetranychus urticae* (Acari: Tetranychidae): polymorphism, host races or sibling species? *Experimental and Applied Acarology*, 24, 365-376.
- Nei, M. and Kumar, S. (2000) *Molecular Evolution and Phylogenetics*, Oxford University Press.
- Norris, D.E., Klompen, J.S.H. & Black, W.C. (1999) Comparison of the mitochondrial 12S and 16S ribosomal DNA genes in resolving phylogenetic relationships among hard ticks (Acari: Ixodidae). *Annals of the Entomological Society of America*, 92, 117-129.
- Norris, D.E., Klompen, J.S.H., Keirans, J.E. & Black, W.C. (1996) Population genetics of *Ixodes scapularis*

- (Acari: Ixodidae) based on mitochondrial 16S and 12S genes. *Journal of Medical Entomology*, 33, 78-89.
- Page, R.D.M. & Holmes, E.C. (1998) *Molecular Evolution: A Phylogenetic Approach*, Blackwell.
- Phillips, A.J. & Simon, C. (1995) Simple, efficient, and nondestructive DNA extraction protocol for arthropods. *Annals of the Entomological Society of America*, 88, 281-283.
- Post, R.J., Flook, P.K. & Millest, A.L. (1993) Methods for the preservation of insects for DNA studies. *Biochemical Systematics and Ecology*, 21, 85-92.
- Reiss, R.A., Schwartz, D.P. & Ashworth, A.C. (1995) Field preservation of Coleoptera for molecular genetic analyses. *Environmental Entomology*, 24, 716-719.
- Rich, S.M., Rosenthal, B.M., Telford, I.S.R., Spielman, A., Hartl, D.L. & Ayala, F.J. (1997) Heterogeneity of the internal transcribed spacer (ITS2) region within individual deer ticks. *Insect Molecular Biology*, 6, 123-129.
- Roderick, G.K. (1996) Geographic structure of insect populations: gene flow, phylogeography, and their uses. *Annual Review of Entomology*, 41, 325-352.
- Roehrdanz, R.L., Degrugillier, M.E. & Black, W.C. (2002) Novel rearrangements of arthropod mitochondrial DNA detected with long-PCR: applications to arthropod phylogeny and evolution. *Molecular Biology and Evolution*, 19, 841-849.
- Rokas, A. & Holland, P.W.H. (2000) Rare genomic changes as a tool for phylogenetics. *Trends in Ecology and Evolution*, 15, 454-459.
- Rose, O.C., Brookes, M.J. & Mallet, J.L.B. (1994) A quick and simple nonlethal method for extracting DNA from butterfly wings. *Molecular Ecology*, 3, 275.
- Ruedas, L.A., Salazar-Bravo, J., Drago, J.W. & Yates, T.L. (2000) The importance of being earnest: what, if anything, constitutes a specimen examined? *Molecular Phylogenetics and Evolution*, 17, 129-132.
- Salomone, N., Emerson, B.C., Hewitt, G.M. & Bernini, F. (2002) Phylogenetic relationships among the Canary Island Steganacaridae (Acari, Oribatida) inferred from mitochondrial DNA sequence data. *Molecular Ecology*, 11, 79.
- Simon, C., Frati, F., Beckenbach, A.T., Crespi, B., Liu, H. & Flook, P. (1994) Evolution, weighting and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America*, 87, 651-701.
- Söller, R., Wohltmann, A., Witte, H. & Blohm, D. (2001) Phylogenetic relationships within terrestrial mites (Acari: Prostigmata, Parasitengona) inferred from comparative DNA sequence analysis of the mitochondrial cytochrome oxidase subunit I gene. *Molecular Phylogenetics and Evolution*, 18, 47-53.
- Sotto-Adames, F., Robertson, H.M. & Berlocher, S.H. (1994) Phylogenetic utility of partial DNA sequences of G6PDH at different taxonomic levels in Hexapoda with emphasis on Diptera. *Annals of the Entomological Society of America*, 87, 723-736.
- Swofford, D.L., Olsen, G.J., Waddell, P.J. & Hillis, D.M. (1996) Phylogenetic inference. In Hillis, D.M., Moritz, C. & Mable, B.K. (eds.) *Molecular Systematics* (2nd Edition), Sinauer, pp. 407-514.
- Toda, S., Osakabe, M.H. & Komazaki, S. (2000) Interspecific diversity of mitochondrial COI sequences in Japanese *Panonychus* species (Acari: Tranychidae). *Experimental and Applied Acarology*, 24, 821-829.
- Whiting, M.F., Carpenter, J.C., Wheeler Q.D. & Wheeler, W.C. (1997) The Strepsiptera problem: phylogeny of the holometabolous insect orders inferred from 18S and 28S ribosomal DNA sequences and morphology. *Systematic Biology*, 46, 1-68.
- Yli-Mattila, T., Paavanen-Huhtala, S., Fenton, B. & Tuovinen, T. (2000) Species and strain identification of the predatory mite *Euseius finlandicus* by RAPD-PCR and ITS sequences. *Experimental and Applied Acarology*, 24, 863-880.
- Zahler, M., Essig, R., Gothe, R. & Rinder, H. (1998) Genetic evidence suggests that *Psoroptes* isolates of different phenotypes, hosts and geographic origins are conspecific. *International Journal for Parasitology*, 28, 1713-1719.
- Zahler, M., Essig, A., Gothe, R. & Rinder, H. (1999) Molecular analyses suggest monospecificity of the genus *Sarcoptes* (Acari: Sarcoptidae). *International Journal for Parasitology*, 29, 759-766.
- Zhang, D.-X. & Hewitt, G.M. (1997) Insect mitochondrial control region: a review of its structure, evolution and usefulness in evolutionary studies. *Biochemical Systematics and Ecology*, 25, 99-120.

Accepted: 4 June 2002