5 Molecular Taxonomy and Phylogeny

Byron J. Adams[1], Adler R. Dillman[2] and Camille Finlinson[1]
[1]*Brigham Young University, Provo, Utah, USA;* [2]*California Institute of Technology, Pasadena, California, USA*

5.1. Introduction

The genus *Meloidogyne* contains over 90 described species and each of these species typically has an extremely broad host range (as many as 3000 plant species; Trudgill and Blok, 2001). In addition to their host-range diversity, they also exhibit tremendous cytogenetic variation (aneuploidy and polyploidy) and mode of reproduction (from obligatory amphimixis to meiotic and mitotic parthenogenesis) (Triantaphyllou, 1985; see Chitwood and Perry, Chapter 8, this volume). In current practice, identification of species is based primarily on the morphological features of females, males and second-stage juveniles (Eisenback and Triantaphyllou, 1991; see Hunt and Handoo, Chapter 3, this volume), as well as esterase and malate dehydrogenase isozyme profiles derived from single females by polyacrylamide gel elecrophoresis (Esbenshade and Triantaphyllou, 1985, 1990; Carneiro *et al.*, 2000) and DNA-based barcodes (Powers and Harris, 1993; Powers, 2004; Powers *et al.*, 2005; see Blok and Powers, Chapter 4, this volume). Historically, the diagnostic features deemed most valuable for identification commonly preceded their use as important characters for taxonomic statements (such as new species descriptions) and, subsequently, phylogenetic analyses. As molecular markers increasingly demonstrated improved resolving power, they became more commonplace as diagnostic tools, eventually becoming more prominent as parts of formal taxonomic statements (including descriptions of new species) and phylogenetic analyses (e.g. Castillo *et al.*, 2003; Landa *et al.*, 2008). With the incorporation of molecular sources of characters and refinements to phylogenetic theory, the fields of taxonomy and evolutionary biology have now become more completely integrated as a discipline, such that the terms molecular taxonomy and phylogenetics are (or should be) subsumed as a single research programme (systematics). In this chapter we present a summary of early and contemporary research on the molecular systematics of *Meloidogyne*.

5.2. The History of Reconstructing *Meloidogyne* Phylogenetic History

Some of the earliest work on evolutionary relationships among species of *Meloidogyne* was based on morphological characteristics and relied heavily on many of the characters used for identification

(Eisenback and Triantaphyllou, 1991). Subsequent efforts involved cytogenetics (Triantaphyllou, 1966, 1985), producing evidence that supported hypotheses consistent with the idea that mitotic parthenogens evolved from meiotic parthenogenetic ancestors, following suppression of meiotic processes and establishing various ploidy levels. Triantaphyllou also hypothesized that the amphimictic species in the genus, such as *Meloidogyne exigua*, are highly specialized parasites and should not be considered as ancestral forms (Triantaphyllou, 1985). Cytogenetic studies, followed by protein and DNA analyses, implied a unique origin of, and monophyly among, the ameiotic species (Dickson *et al.*, 1971; Dalmasso and Bergé, 1978; Esbenshade and Triantaphyllou, 1987; Castagnone-Sereno *et al.*, 1993; Baum *et al.*, 1994; van der Beek *et al.*, 1998). Later, studies based on mitochondrial genes soon revealed that mitochondrial genes can be hypervariable, both in patterns of sequence substitution and in gene content and arrangement (Powers and Sandall, 1988; Powers *et al.*, 1993). These properties are desirable for diagnostic or population genetic markers, or for resolving phylogenetic relationships among closely related species (see Blok and Powers, Chapter 4, this volume). More recent phylogenetic analyses have utilized small ribosomal subunit (18S) rDNA sequences (De Ley *et al.*, 2002), large subunit (28S) rDNA (Castillo *et al.*, 2003), and mitochondrial DNA (mtDNA) sequences (Tigano *et al.*, 2005). Lunt (2008) performed separate analyses of four genes – dystrophin, elongation factor 1-alpha, major sperm protein, and RNA polymerase 2 – as part of a clever study to elucidate whether the origin of the asexual *Meloidogyne* lineages was ancient or recent (Adams and Powers, 1996; Hugall *et al.*, 1999). A phenomenon that emerges from each of these studies is the close relationships among the three major mitotic parthenogenetic species: *Meloidogyne arenaria*, *Meloidogyne javanica* and *Meloidogyne incognita*. Regardless of the type of phylogenetic analysis performed, or the genetic locus examined, the preponderance of evidence from single gene analyses suggests that the mitotic parthenogens are clearly evolutionarily distinct from either the meiotic or obligatory amphimictic species. However, gene trees are not always concordant with the evolutionary history of independently evolving species, and the discordance between the two different histories can confound phylogenetic inference. Discordance between gene trees and species trees is most commonly explained as lineage sorting among mitochondrial haplotypes, but, in addition, paralogous genes, such as would be expected for the rDNA tandem array if intraspecific concerted evolution were non-uniform, or non-orthologous genes go undetected (Maddison, 1997; Maddison and Knowles, 2006). Phylogenomic analyses hold the promise of resolving problematic phylogenies by swamping the data sets with signal, despite high noise, by including character information from numerous loci (Eisen, 1998; Eisen and Fraser, 2003), but see Longhorn *et al.* (2007). Although phylogenomic analyses that could exhaustively sample all *Meloidogyne* are premature (if not unnecessary), in a preliminary effort of this kind Scholl and Bird (2005) identified numerous putative homologues and used them to generate a phylogeny for a subset of *Meloidogyne* species. Although this effort was based on a small sample of taxa, the major contribution was the elucidation of relationships among three mitotic parthenogens (*M. hapla*, *M. incognita*, *M. javanica*) that had been poorly resolved in previous phylogenetic analyses. Subsequent refinements to *Meloidogyne* phylogeny have consisted primarily of analyses that have added new or previously unsampled taxa to existing databases (Castillo *et al.*, 2003; Landa *et al.*, 2008).

5.3. Molecular Phylogenetics: Genetic Markers and Evolutionary Relationships

5.3.1 Nuclear Ribosomal DNA Sequences

Nuclear ribosomal DNA is currently the most extensively employed molecular marker for *Meloidogyne* molecular systematics. Variation in mutation rates observed among different genes and spacers within an rDNA transcription unit results in regions of adjacent DNA segments in the cistron that are useful across a wide range of taxonomic hierarchical levels (Hillis and Dixon, 1991). This includes conserved and variable regions of the 18S and 28S subunits, and the more highly variable ITS region. These three rRNA gene regions are the most commonly used genetic markers for nematode molecular systematics, and each of these regions has been employed for *Meloidogyne* phylogenetics (Landa *et al.*, 2008).

Whilst rRNA genes may provide optimal levels of variation for investigating *Meloidogyne* phylogeny, they are not without significant theoretical and analytical drawbacks. The single biggest obstacle in using rRNA genes is that, unlike protein-coding genes, they are not constrained to maintain codon fidelity or even an open reading frame. Whereas the length and composition of protein-coding genes are generally subject to selection by codon usage, rRNA genes are not. Thus, for some rDNA regions, insertion and deletion events (indels) can be as frequent as transitions and transversions, often involving blocks of multiples of nucleotides (Powers, 2004; Powers *et al.*, 1997). Indel events can result in substantial rDNA size differences between sequences (taxa), which complicates the process of generating multiple sequence alignments and reduces confidence in the homology statements for each nucleotide in the multiple sequence alignment. In our experience, and as shown by others, there is usually more variation in tree topology due to differences in the multiple sequence alignment than there is among the different methods used to generate the trees (i.e. parsimony, maximum likelihood, Bayesian and distance methods) (Morrison and Ellis, 1997). Approaches to addressing this problem require thoughtful consideration of the mechanics of how multiple sequence alignments (homology statements) are constructed. These involve the nuts and bolts of how computer algorithms generate multiple sequence alignments, removing the alignment-ambiguous regions based on an a priori metric (i.e. remove ambiguous indels that lie between a predetermined number of invariant nucleotides (Nguyen *et al.*, 2001)), direct optimization (Terry and Whiting, 2005), comparison of secondary structure based on minimum energy models (Subbotin *et al.*, 2006), and minimum posterior probabilities among alternative placements of nucleotides (characters) in the alignment (Loytynoja and Milinkovitch, 2003). These problems are not unique to rDNA, as alignment ambiguity can also arise where protein-coding genes have undergone tremendous divergence, or for other non-coding sequences (such as non-coding regions of mitochondrial DNA; see section 5.3.3).

5.3.1.1. 18S (small ribosomal subunit)

Although 18s sequences for several *Meloidogyne* species had long been available in public databases, De Ley *et al.* (2002) were the first to use this locus as part of a rigorous reconstruction of *Meloidogyne* phylogeny. Their analysis included 12 species of *Meloidogyne* and four outgroup taxa subject to phylogenetic analyses generated from three different multiple sequence alignment methodologies and three different tree-building optimality criteria (distance, parsimony and maximum likelihood). Calculations of phylogenetic signal (skewness of tree length distribution) were high and intraspecific sequence polymorphism low, suggesting that the locus was appropriately robust for resolving relationships among the sampled species but with nodal support strongest at the deeper nodes. Their analysis showed strong support for three clades, which they designated: clade I, conscribing the mitotic parthenogens (*M. incognita*, *M. arenaria* and *M. javanica*); clade II, including the obligatory amphimictic, meiotic and mitotic parthenogens (*Meloidogyne hapla* races A and B, respectively) as well as *Meloidogyne duytsi* and *Meloidogyne maritima*; and clade III, containing three meiotic parthenogens (*Meloidogyne exigua*, *Meloidogyne graminicola* and *Meloidogyne chitwoodi*). As with earlier phylogenetic efforts (Triantaphyllou, 1985; Castagnone-Sereno *et al.*, 1993), their analysis supported the location of the ameiotic species as distantly related to either the obligate amphimictic or meiotic species (Fig. 5.1).

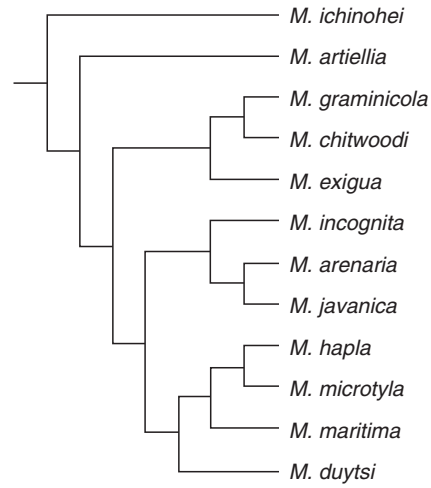

Fig. 5.1. 18s rDNA phylogeny of De Ley *et al.* (2002). The tree was generated from a secondary-structure-based multiple sequence alignment and resolved via maximum parsimony. (Adapted from De Ley *et al.*, 2002.)

Subsequent 18s analyses of Tigano *et al.* (2005) included 19 additional sequences, representing 12 nominal species, as well as several unknown isolates from disparate geographic locations, in a effort that revealed interesting comparisons between the 18s rDNA sequences and other tools for diagnosing species, including morphological and isozyme phenotypes. Their 18s rDNA analysis proceeded from sequences profile-aligned to the optimal sequence alignment of De Ley *et al.* (2002) by distance, parsimony and maximum likelihood tree-building algorithms, all of which produced congruent topologies (Fig. 5.2). Their results revealed polymorphisms between isolates

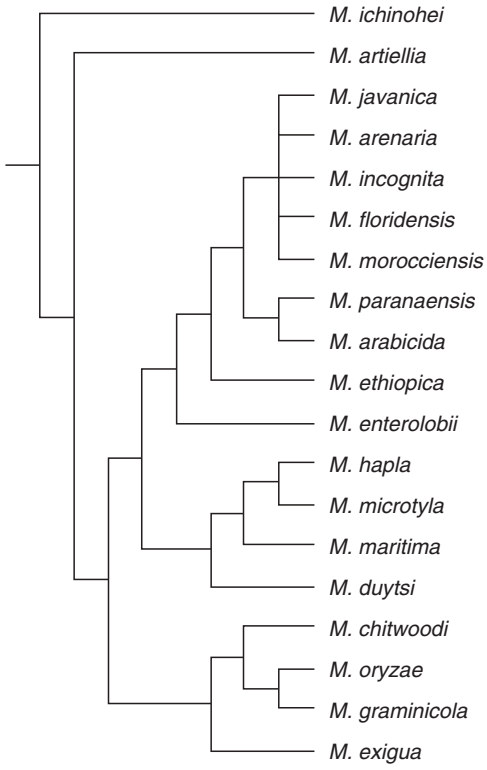

Fig. 5.2. 18s rDNA phylogeny of Tigano *et al.* (2005). The tree was generated from sequences profile-aligned to the secondary-structure-based alignment of De Ley *et al.* (2002) and resolved via maximum likelihood. (Adapted from Tigano *et al.*, 2005.) Only nominal taxa were retained. Taxa with multiple representative sequences were represented by a single semaphoront and relationships collapsed to their most inclusive clade.

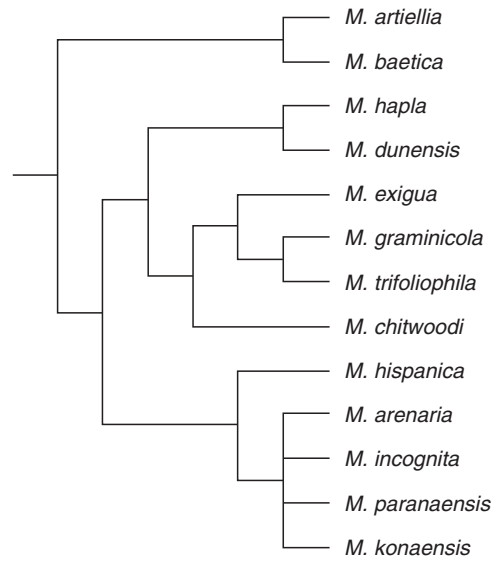

Fig. 5.3. 18s rDNA phylogeny of Landa *et al.* (2008). The tree was generated from sequences aligned using Bionumerics software ver. 4.5 (Applied Maths, Kortrijk, Belgium) and resolved via maximum parsimony (also by Bionumerics). (Adapted from Landa *et al.*, 2008.)

of the same species (as per morphological and isozyme diagnosis), including *M. arenaria*, *M. incognita*, *M. javanica*, *M. exigua* and *M. hapla*. Some of the differences were slight, but several were substantial: *M. arenaria* of De Ley *et al.* (2002) differed from the Tigano *et al.* (2002) sequences by two substitutions and four insertions; two *M. incognita* isolates differed by one insertion and ten substitution events; two *M. javanica* sequences differed by 24 substitutions, 18 ambiguities and four insertions. However, 18s rDNA sequences can also be highly conserved between species. For example, *Meloidogyne hispanica* and *Meloidogyne ethiopica* yield identical 18s sequences, so only *M. hispanica* is listed in the analysis of Landa *et al.* (2008) (Fig. 5.3).

5.3.1.2. 28S (large ribosomal subunit)

The first effort to use the 28s rDNA region to resolve phylogenetic relationships among *Meloidogyne* spp. was that of Tenente *et al.* (2004). Their analysis included nine species represented by 12 sequences. Their most robust phylogenetic analyses included exhaustive, unrooted maximum

parsimony searches and maximum likelihood searches based on various ClustalX alignments. The data for each of the generated alignments revealed strong phylogenetic signal as based on g1 statistics, and most nodes, particularly for the deeper clades, were strongly supported by bootstrap resampling estimates. Their analysis found fairly strong support for two monophyletic clades that are compatible with clades I and III of the 18s analysis of De Ley *et al.* (2002) (but did not include any of the clade II taxa) (Fig. 5.4). The weakest supported nodes, and the ones that differed the most by alignment and tree search strategy, involved the relationships among the mitotic parthenogenetic species *M. arenaria*, *M. incognita*, *Meloidogyne konaensis*, and *Meloidogyne paranaensis*. They conclude that the D2/D3 region of this marker, which showed high degrees of variation between two species of *Acrobleoides* that are morphologically virtually indistinguishable (De Ley *et al.*, 1999), were 'simply too conserved for the phylogenetic analysis of mitotic parthenogenetic *Meloidogyne* species' (Tenente *et al.*, 2004). A contemporary analysis by Castillo *et al.* (2003) also reflected these sentiments. Their analysis, which included the Tenente *et al.* sequences, also included several other unpublished *Meloidogyne* and outgroup (*Pratylenchus*) sequences, and was based on simple ClustalX default alignment parameters under the maximum parsimony optimality criterion. Their results are consistent with Tenente *et al.* (2004), whereby monophyletic groups I and III form a clade with the more ancestral lineage comprising *Meloidogyne artiellia* and *Meloidogyne baetica* (representatives of clade II were not included in the analysis). Like Tenente *et al.* (2004), Castillo *et al.* (2003) and Landa *et al.* (2008) were unable to resolve relationships among species of clade I but resolved deeper nodes with much greater support (Figs 5.5 and 5.6).

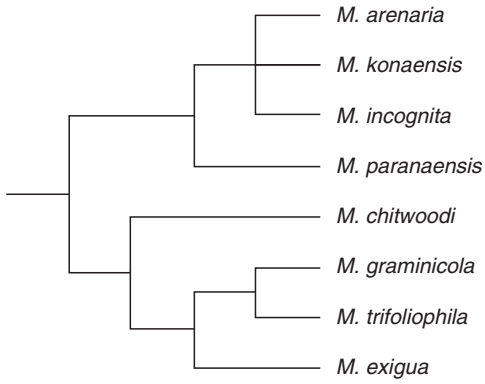

Fig. 5.4. 28s rDNA phylogeny of Tenente *et al.* (2004). The tree was generated from sequences aligned using ClustalX (Thompson *et al.*, 1997) default parameters and resolved using maximum parsimony. (Adapted from Tenente *et al.*, 2004.)

5.3.1.3. ITS (internally transcribed spacer region)

Early work by Hugall *et al.* (1999) on *Meloidogyne* ITS rDNA sequences revealed a somewhat surprising phenomenon: although amphimictic species exhibited only a single ITS lineage, the ameiotic species *M. hapla*, *M. arenaria* and *M. incognita* exhibited numerous lineages, even within individuals. In fact, Hugall *et al.* (1999) showed that up to 90% of the total ITS diversity could be found within an individual nematode, which contained as many as 9–13 different sequence variants. Such variation, and the way it is partitioned across the genus, poses challenges to using it to infer phylogenetic relationships but does provide strong evidence for the hybrid origins of *M. hapla*, *M. arenaria* and *M. incognita* (Hugall *et al.*, 1999).

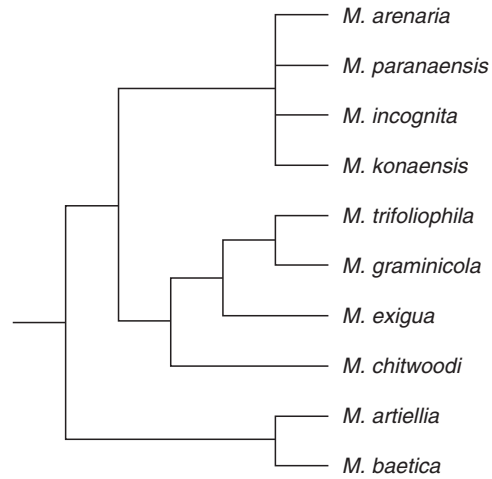

Fig. 5.5. 28s rDNA phylogeny of Castillo *et al.* (2003). The tree is a strict consensus of five equally parsimonious trees generated from sequences aligned using ClustalX default parameters (Thompson *et al.*, 1997) and resolved via maximum parsimony (Swofford, 2002). (Adapted from Castillo *et al.*, 2003.)

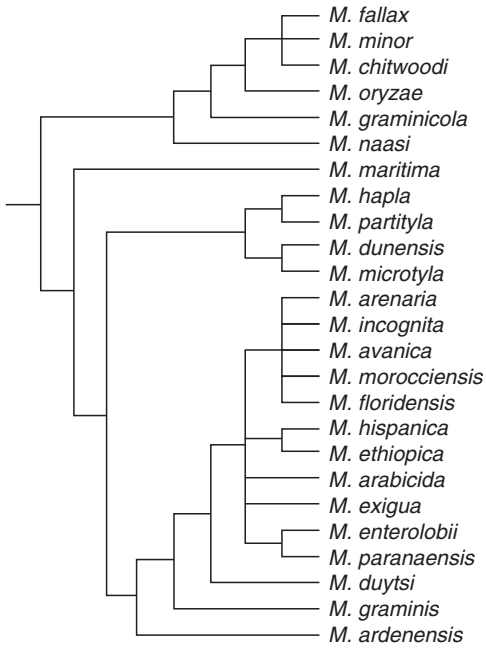

Fig. 5.6. 28 s rDNA phylogeny of Landa *et al.* (2008). The tree was generated from sequences aligned using Bionumerics software ver. 4.5 (Applied Maths, Kortrijk, Belgium) and resolved via maximum parsimony (also by Bionumerics). (Adapted from Landa *et al.*, 2008.)

One of the by-products of exploring ITS sequence variation in search of diagnostic markers is a large, publicly available database of sequences. In an early effort, Castillo *et al.* (2003) generated an unrooted maximum parsimony tree from ITS sequences for nine species. Subsequently, Landa *et al.* (2008) used these and additional sequences to perform a phylogenetic analysis of 16 species from 29 different isolates. Although details of their sequence alignment process is not explicit, the maximum parsimony tree they generated depicts a completely resolved tree but with varying levels of support, particularly through the intermediate nodes. Clades III and I are monophyletic. Clade II is depicted as paraphyletic, but the discordant nodes are also those that are most weakly supported. Also evident in the resulting tree is the apparent paraphyletic nature of the ITS lineages within and among the ameiotic species (Figs 5.7 and 5.8).

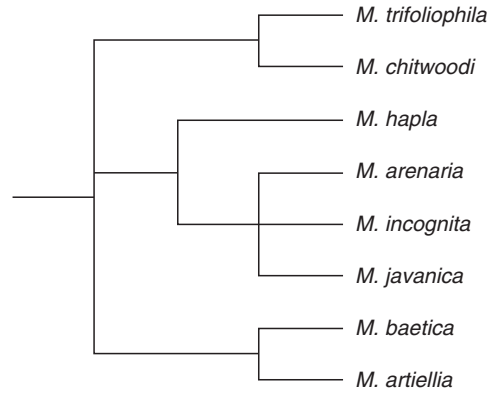

Fig. 5.7. ITS rDNA phylogeny of Castillo *et al.* (2003). The tree was generated from sequences aligned using ClustalX default parameters (Thompson *et al.*, 1997) and resolved via maximum parsimony (Swofford, 2002). The original solution was presented as unrooted, we root it here with *Meloidogyne artiellia* and *Meloidogyne baetica*. (Adapted from Castillo *et al.*, 2003.)

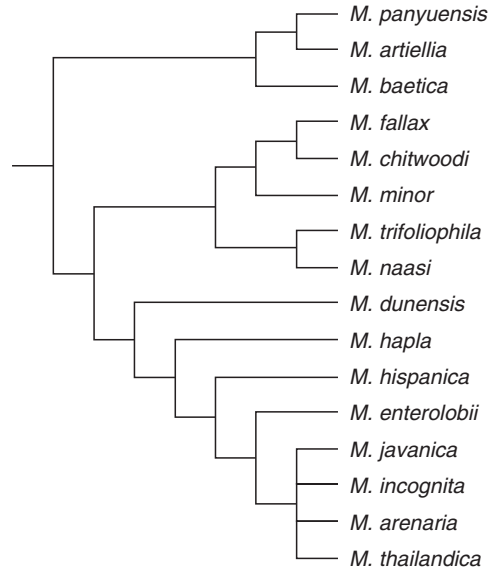

Fig. 5.8. ITS rDNA phylogeny of Landa *et al.* (2008). The tree was generated from sequences aligned using Bionumerics software ver. 4.5 (Applied Maths, Kortrijk, Belgium) and resolved via maximum parsimony (also by Bionumerics). (Adapted from Landa *et al.*, 2008.)

5.3.2. Orthologous Nuclear Genes

As with gene families, concerted evolution does not always work fast enough to homogenize all copy variants in the rDNA cistron. The result is that paralogous gene genealogies can contradict actual phylogenetic relationships of species. This is evident by lineage sorting events for mtDNA (discussed below) and also by species that arose via hybridization events, such as the mitotic parthenogenetic species of *Meloidogyne* (Hugall *et al.*, 1999). Thus, an optimal phylogenetic analysis will include as many nuclear, single-copy genes in the transformation series as possible. In an effort to distinguish the relative timing of the hybridization events involved in the origin of the ameiotic lineages, Lunt (2008) explored the evolution of four different single-copy nuclear genes. Prior to this effort, Scholl and Bird (2005) undertook a phylogenomic approach in order to tease out the relationships between the mitotic parthenogenetic *Meloidogyne* species. Each of these efforts is discussed below.

5.3.2.1. Dystrophin

Dystrophin is a muscle protein that connects the cytoskeleton of a muscle fibre to the surrounding extracellular matrix. It is the longest gene in the human genome but exists in single-copy form in invertebrates (Roberts *et al.*, 1995; Roberts and Bobrow, 1998). Lunt (2008) designed primers that amplified a 670–770 bp product of three exons and two introns from multiple populations of five *Meloidogyne* species (*M. incognita*, *M. javanica*, *M. arenaria*, *Meloidogyne enterolobii* (=*Meloidogyne mayaguensis*) and *M. hapla*) and an outgroup taxon (*Globodera pallida*) for phylogenetic analyses using maximum likelihood. The results for this gene are consistent with the clade designations of De Ley *et al.* (2002) but with failure to resolve unambiguously relationships among the members of clade I (*M. enterolobii*, *M. arenaria*, *M. incognita*, *M. javanica*). However, none of the paraphyletic nodes is well supported by approximate likelihood ratio tests (Fig. 5.9).

5.3.2.2. Major sperm protein (msp)

Major sperm protein is the most abundant protein in nematode sperm and is responsible for the cell's motility (Roberts, 2005). Lunt (2008) studied *Meloidogyne* msp gene genealogies in order to see if the gene underwent an accelerated rate of mutation after the evolution of mitotic parthenogenesis and, therefore, putatively no longer under selection pressure to maintain its function; it did not. Maximum parsimony analysis from multiple populations of five *Meloidogyne* species (*M. incognita*, *M. javanica*, *M. arenaria*, *M. enterolobii*, *M. hapla*) and an outgroup taxon (*G. pallida*) yielded a phylogenetic tree that is consistent with the three clade designation of De Ley *et al.* (2003), except that the positions of clade II and III are reversed relative to clade I, but the node involved in this reversal is not well supported (Fig. 5.10).

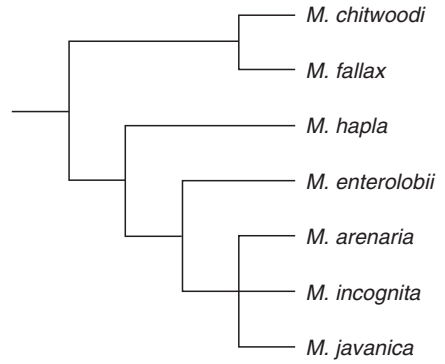

Fig. 5.9. Dystrophin exon gene genealogy of Lunt (2008). The tree was generated from exon sequences aligned with Clustal X (Thompson *et al.*, 1997) (unambiguous alignment due to codon structure) and resolved using maximum likelihood as implemented in PhyML (Guindon and Gascuel, 2003).

5.3.2.3. Elongation factor 1-alpha (EF1-α)

Elongation factor 1-alpha (EF-1α), the GTP binding protein involved in catalysing the binding of the aminoacyl-transfer RNAs to the ribosome, is an essential component of eukaryotic translation (Watson, 2008). Lunt (2008) pursued this gene as an independent estimate of phylogenetic relationships but, upon inspection of the resulting phylogenetic tree, suspected that instead of being single copy, a gene duplication event might have occurred, resulting in paralogous loci. In support of this, Lunt identified two copies in

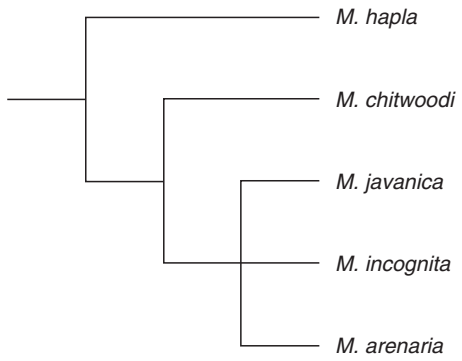

Fig. 5.10. Major sperm protein exon gene genealogy of Lunt (2008). The tree was generated from exon sequences aligned with Clustal X (Thompson *et al.*, 1997) (unambiguous alignment due to codon structure) and resolved using maximum likelihood as implemented in PhyML (Guindon and Gascuel, 2003).

the *Caenorhabditis elegans* genome, yet the maximum likelihood solution among his sampled species was still congruent with the clade designations of De Ley *et al.* (2002). However, the EF-1α gene genealogy among the ameiotic species is paraphyletic, poorly supported and too conserved to distinguish *Meloidogyne fallax* from *M. chitwoodi* (Fig. 5.11).

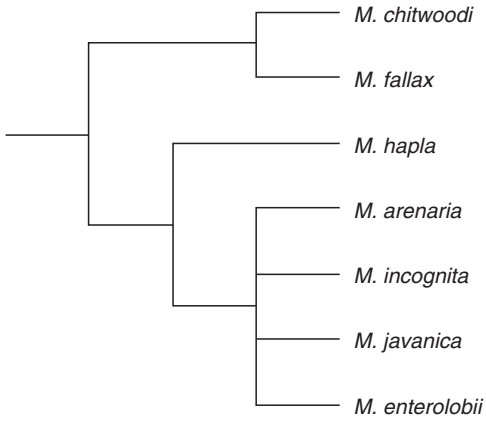

Fig. 5.11. Elongation factor 1-alpha exon gene genealogy of Lunt (2008). The tree was generated from exon sequences aligned with Clustal X (Thompson *et al.*, 1997) (unambiguous alignment due to codon structure) and resolved using maximum likelihood as implemented in PhyML (Guindon and Gascuel, 2003).

5.3.2.4. RNA polymerase 2

RNA polymerase 2 is the eukaryotic enzyme responsible for synthesis of mRNA during transcription (Kornberg, 2007). Lunt (2008) designed primers that amplified an approximately 710 bp fragment that included coding sequence from two exons and one intron. Using the exon sequences only resulted in the maximum likelihood solution in Fig. 5.12. The tree is consistent with the clade designations of De Ley *et al.* (2002) but fails to resolve *M. chitwoodi* and *M. fallax* lineages within Clade III, and where they are resolved, the *M. javanica*, *M. incognita* or *M. arenaria* relationships within Clade I are paraphyletic.

5.3.3. Mitochondrial DNA

Despite apparent low genetic diversity (or because of it; see Blok and Powers, Chapter 4, this volume) among populations of *M. arenaria*, *M. javanica* and *M. incognita* (Hugall *et al.*, 1994, 1997; Stanton *et al.*, 1997), mtDNA sequences spanning the COII through lRNA genes have been intensively studied for *Meloidogyne* molecular diagnostics (Blok *et al.*, 2002; Powers, 2004; Brito *et al.*, 2004; Powers *et al.*, 2005). The diagnostic utility of the marker persuaded Tigano *et al.* (2005) to explore

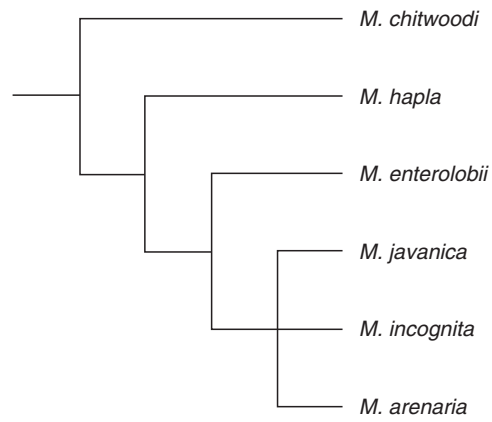

Fig. 5.12. RNA polymerase 2 exon gene genealogy of Lunt (2008). The tree was generated from exon sequences aligned with Clustal X (Thompson *et al.*, 1997) (unambiguous alignment due to codon structure) and resolved using maximum likelihood as implemented in PhyML (Guindon and Gascuel, 2003).

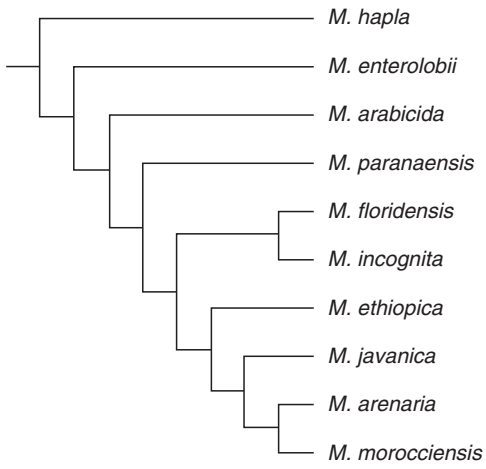

Fig. 5.13. Mitochondrial DNA phylogeny of Tigano *et al.* (2005). The tree was generated from DNA sequences that span the COII through lRNA region, including the complete sequence for tRNA-His, and the AT-rich region. Sequences were aligned using Clustal X (Thompson *et al.*, 1997), adjusted by eye using MacClade (Maddison and Maddison, 2002), and resolved by maximum parsimony as implemented in PAUP* (Swofford, 2002).

its ability to resolve phylogenetic relationships among the very closely related Clade I taxa. By parsimony, distance and maximum likelihood approaches they analysed the complete region, including partial COII and partial lRNA sequence, the complete sequence for tRNA-His and the AT-rich region, and obtained a single optimal solution. The concatenation of the different gene regions was justified by failure to reject shared evolutionary histories by way of an incongruence length difference test. Although there is some discordance between this tree and their rDNA solution, the relationship among the ameiotic species is congruent, if poorly supported, with the 18s rDNA analyses of Tigano *et al.* (2005) and De Ley *et al.* (2002) (Fig. 5.13).

5.3.4. Phylogenomics

Optimally, phylogenetic construction proceeds from consideration of as many independently evolving, heritable characters as possible (Farris, 1983; Kluge, 1997). In order to address the possible problems associated with gene tree/species tree discordance, a reasonable expectation might be that if one simply looks at enough characters, or in this case enough DNA sequence from enough genes, eventually the phylogenetic signal will swamp noise. Such is the idea of using genomic data to resolve phylogenetic relationships, where it has been shown that using small numbers of genes can produce support for incorrect phylogenies, but that support and resolution become optimized at about 20 genes. Taking this into consideration, Scholl and Bird (2005) sampled 47 orthologous genes from several tylenchid nematodes, including *M. chitwoodi*, *M. hapla*, *M. arenaria*, *M. javanica* and *M. incognita*, with the goal of resolving the relationships between the apomicts (*M. arenaria*, *M. javanica* and *M. incognita*) where previous analyses appeared to show conflicting relationships: based on mtDNA, Powers and Sandall (1988) suggested the relationship to be (*M. arenaria* (*M. javanica* + *M. incognita*), whereas the 18s best estimate of De Ley *et al.* (2002) was (*M. incognita* (*M. javanica* + *M. arenaria*)). To construct their phylogenetic trees, Scholl and Bird (2005) first performed rigorous screens to identify orthologous genes from EST sequence databases that had a homologue in *C. elegans* (and thus conserved across large phylogenetic distance, reflecting evolutionary constraint). The orthologues were aligned based on their inferred amino acid sequence and then back-translated to their DNA sequence to maintain open reading frame fidelity. Phylogenies were constructed using multiple alignments of the individual genes and concatenated full-length data sets containing all the genes, or different subsets of genes, via Bayesian inference, maximum likelihood and minimum evolution (via neighbour joining). The optimal solution based on all of the analyses favoured the hypothesis of Powers and Sandall (1988) (Fig. 5.14).

5.4. A *Meloidogyne* Supertree Analysis

It would be excellent if we could just take all the DNA sequences for all of the *Meloidogyne* taxa ever generated, compile them into a single giant multiple sequence alignment, crunch it through some tree-building algorithms and confidently

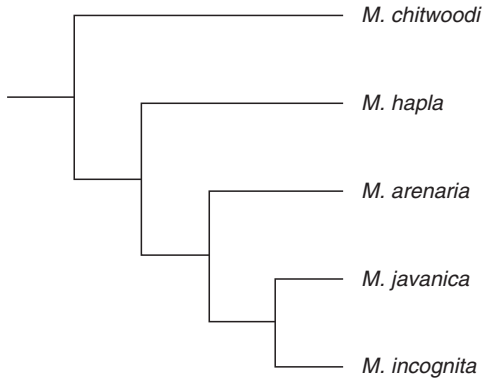

Fig. 5.14. Phylogeny of Scholl and Bird (2005). The tree was generated from a concatenated matrix of 47 putative orthologous genes aligned using Clustal W (Thompson *et al.*, 1994) and resolved using Bayesian analysis as implemented in MRBAYES (Huelsenbeck and Ronquist, 2001).

report the one true tree of *Meloidogyne* spp. evolutionary history. While perhaps optimal, such an exercise is obviated by the range of sequence variation that spans the gamut of evolutionary rates of change on the loci under study. For example, for many of the mtDNA and ITS sequences, the divergence and length differences prohibit making unambiguous multiple sequence alignments among all taxa in the genus. Similarly, not all taxa are present in all of the data sets, and the inclusion of large amounts of missing characters can result in spurious phylogenetic inference (Maddison, 1993; Wiens, 1998). One solution to this problem is to generate a supertree – an evolutionary tree that is assembled from a bunch of smaller trees that share some, but not necessarily all, common taxa (Bininda-Emonds, 2004).

Although it may sound straightforward to take several phylogenetic trees and spin them into a single tree that is the sum of all the parts, it is not. What should not be surprising is that the quality of the obtained supertree is a function of the quality of the phylogenies used to build it. Even if the phylogenies used to build the supertree are robust, there are still several important aspects of building a supertree to consider; we will only touch on a few here that are relevant to *Meloidogyne* (but see Bininda-Emonds *et al.*, 2004). In the case for building a *Meloidogyne* supertree one must consider not only the quality of the different phylogenies used to build the tree but also whether the phylogenies are independent. As opposed to using morphological or molecular characters to build phylogenies, the raw data for generating supertrees are lifted from the topological arrangements of two or more partially overlapping phylogenetic trees. So just as using the character 'male stylet length' five times in a character matrix would be redundant, using three different 18s trees to build a supertree that was to include seven other genetic loci could artificially bias the overall supertree topology in favour of the 18s topology. In other words, the three 18s trees are not independent estimates of phylogenetic relationships and must be dealt with somehow so as not to bias the analysis unfairly. Similarly, what about different trees that are generated by the same author? If Lunt (2008) used the same general methodology, from alignment to tree-building strategy, to construct all of his phylogenetic trees, is there an element of non-independence among his different gene trees? In fact, the majority of *Meloidogyne* phylogenetic analyses to date employed several different approaches in the same publication, from alignment strategy to tree-building algorithms and optimality criteria. More often than not, these different approaches generated several different hypotheses of relationships for each data set. Of these, which do we choose to use as source trees to generate a supertree?

With the above caveats in mind, we generated a supertree based on the best estimates of phylogenetic relationships from the most robust and comprehensive molecular phylogenetic analyses we could find in the literature. In order to account for duplication of gene trees (three 18s, three 28s, and two ITS gene trees) we generated 'mini-supertrees' from each set of redundant gene trees and then used the 'mini-supertree' as the source tree for the main analysis (Bininda-Emonds *et al.*, 1999). In choosing among the alternative phylogenetic arrangements published in most papers (i.e. the maximum parsimony tree versus the maximum likelihood tree), we did what any good taxonomist would do – appeal to authority. We simply chose the hypothesis that was most favoured by the authors.

To construct the supertree we mined the optimal topologies from the following publications: De Ley *et al.* (2002), Castillo *et al.* (2003), Tenente *et al.* (2004), Scholl and Bird (2005), Tigano *et al.* (2005), Landa *et al.* (2008) and Lunt

(2008). To account for non-independent estimates we generated mini-supertrees from the data sets that were represented more than once (18s, 28s and ITS). The mtDNA data set was treated as a single estimate, even though it comprised a concatenation of several genes. While potentially misleading, we justify this because Tigano *et al.* (2005) tested for shared evolutionary history and common inheritance of each of the individual genes. The phylogenomic analysis of Scholl and Bird (2005) presents an interesting case because their analysis consisted of 47 concatenated genes. In theory, for the present analysis we should be analysing each of the 47 gene tree topologies independently and then using each one of those topologies as an independent estimate of relationships (source tree) to construct the supertree. When Scholl and Bird performed their analyses on the concatenated data set of 47 genes, the data matrix was treated as if it were a single gene, assuming a single model of evolution (general time-reversible with four categories of gamma-distributed rate heterogeneity). In reality, since it is unlikely that all 47 genes evolved under the same model of sequence evolution, it would have been more appropriate to partition the concatenated data set by gene, each partition with its own most appropriate model of evolution, but that approach was still under development at the time of their analyses (Huelsenbeck *et al.*, 2008). However, Scholl and Bird did do independent analyses of subsets of single genes and reported that they were congruent with the overall concatenated gene phylogeny. Thus, as with the mtDNA topology, for our supertree analysis we used the Scholl and Bird (2005) topology as a single, independent estimate of relationships, acknowledging that were we to use all 47 gene topologies (which were identical) that there would be overwhelming support for a (*M. arenaria* (*M. javanica* + *M. incognita*)) clade in the present analysis. Similarly, by using the mtDNA data as a single source tree, we are probably underestimating the overall support for its topology in the supertree.

We used two supertree construction methods: matrix representation of parsimony (Purvis, 1995) with the matrix generated in RadCon 1.1.6 (Thorley and Page, 2000) and implemented in PAUP* (Swofford, 2002), and the most similar supertree method (dfit) as implemented in CLANN ver. 3.0.0 (Creevey and McInerney, 2005) (Fig. 5.15). Even though these reconstructions involved the analysis of only nine different topologies, their sum of possible unrooted solutions is 2.92156×10^{40}, presenting a fairly computationally intensive effort. Each of the two supertree analyses yielded multiple equally parsimonious (MRP) or costly (dfit) trees (1736 and 7, respectively), which we represent here using two different consensus approaches (combinable components and majority rule) (Fig. 5.15). For both types of supertree construction methods, the consensus trees differed only in terms of resolution (combinable components being more conservative and less resolved). Overall the obtained topologies among the different supertree methods are quite similar. Both approaches identified *Meloidogyne ichinohei* as the lineage that shares a most recent common ancestry with the remaining members of the genus, with the next lineage to branch being the monophyletic clade of (*M. baetica* (*Meloidogyne panyuensis* + *M. artiellia*)). Membership in clade III is congruent between the two analyses, with the exception that the dfit solution suggests sister relationships for (*Meloidogyne oryzae* + *M. graminicola*) and (*Meloidogyne trifoliophila* + *Meloidogyne naasi*). Clade II membership is monophyletic in the dfit topology, differing from the MRP solution, which suggests that *Meloidogyne dunensis* and *Meloidogyne microtyla* are sister taxa. Also, the MRP topology differs radically from the dfit solution as *Meloidogyne graminis* and *Meloidogyne ardenensis* nest within Clade II, but in the MRP solution they are ancestral to Clades I, II and III. Clade I membership is congruent between both analyses, and lack of resolution is completely understandable given that these relationships were poorly resolved in most of the source trees. Both approaches favour *Meloidogyne floridensis* and *M. incognita* as sister taxa (Fig. 5.15B,C), but beyond this relationship there is only discord. The most fundamental difference is that of the relationship between *M. arenaria*, *M. incognita* and *M. javanica*. The MRP solution favours (*M. incognita* (*M. arenaria* + *M. javanica*)), whereas the dfit solution is (*M. arenaria* (*M. javanica* + *M. incognita*)). Because the actual number of 'characters' (tree topologies) is so small ($n = 9$), bootstrap support for any of the relationships is virtually nonexistent (data not shown). However, phylogenetic signal was significantly better than random (permutation tail probability test; $P < 0.01$).

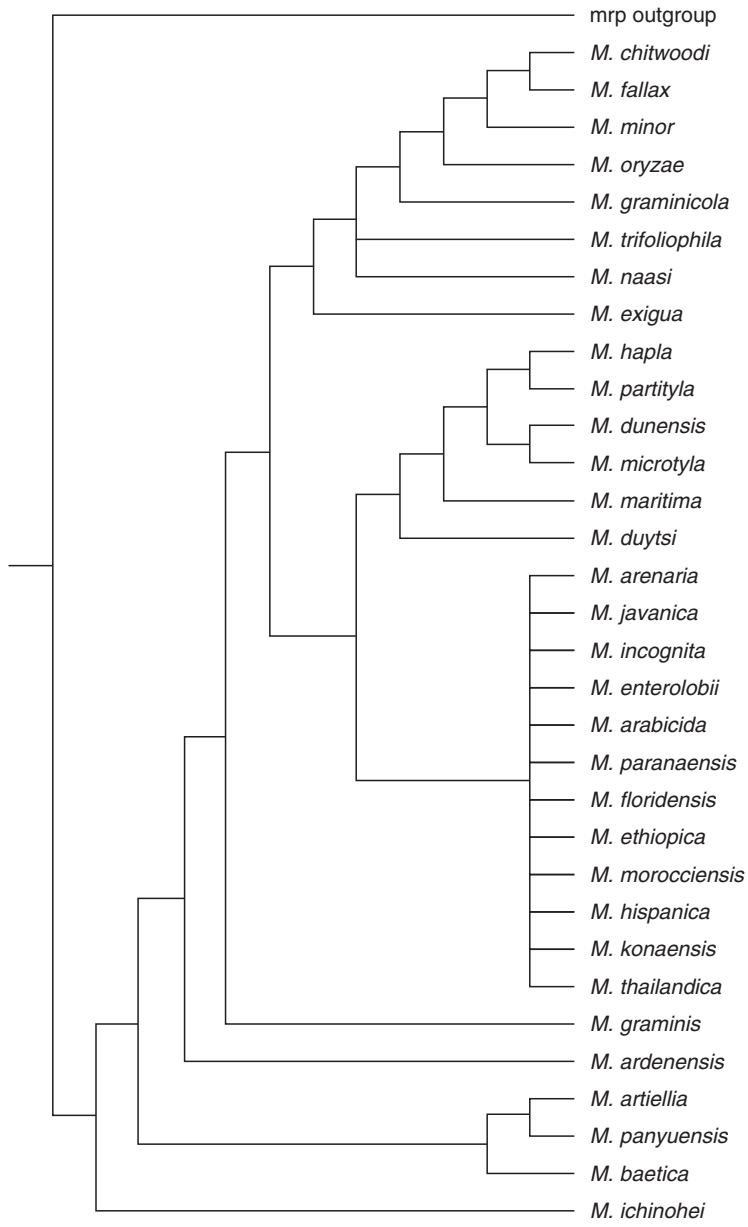

Fig. 5.15. Supertree solutions for seven independent *Meloidogyne* source trees. A: MRP combinable components consensus;

5.5. Conclusions and Future Directions

Overall, the topologies of the phylogenies we used as source trees are remarkably similar. Following the clade designations of De Ley *et al.* (2002), clade I includes the mitotic parthenogens, *M. arenaria*, *M. incognita* and *M. javanica*, as well as *M. enterolobii*, *Meloidogyne morociensis*, *M. ethiopica*, *M. hispanica*, *M. konaensis*, *M. paranaensis*, *Meloidogyne*

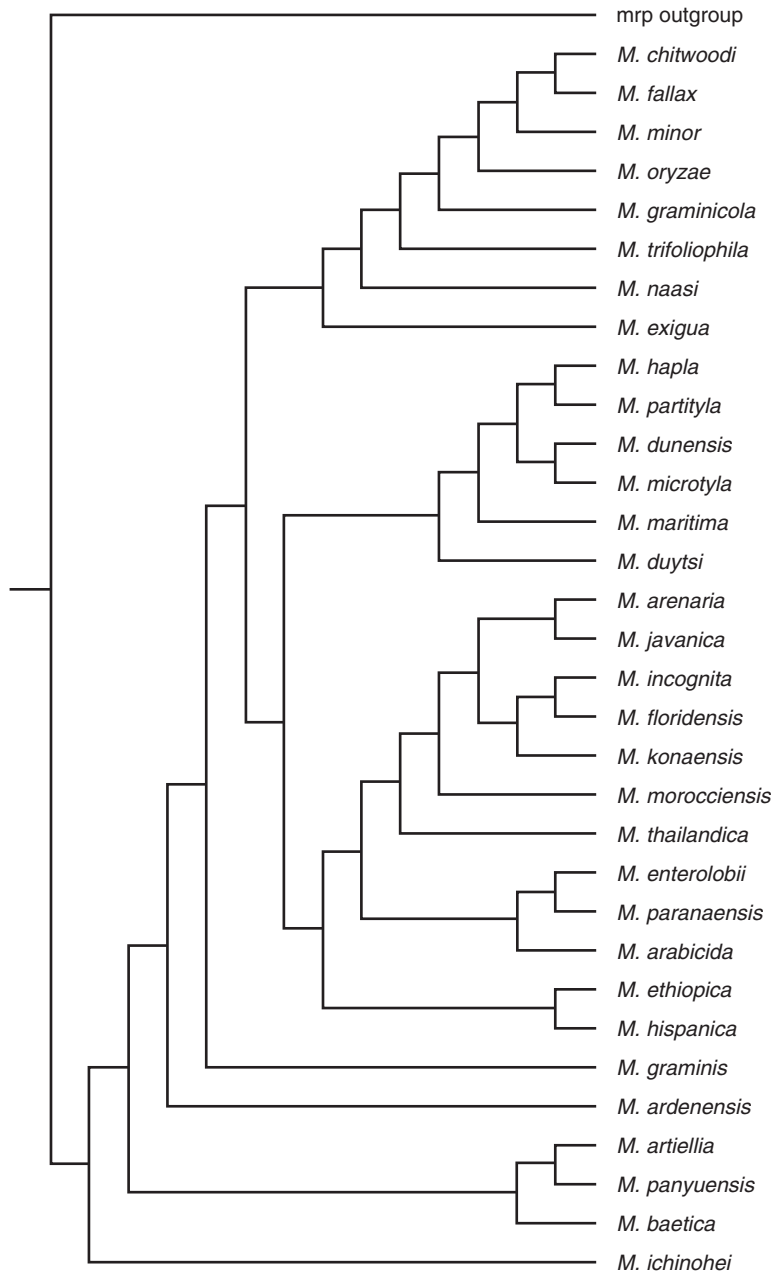

Fig. 5.15. (Continued) B: MRP majority rule consensus tree;

thailandica and *Meloidogyne arabicida*. The only known exceptional member of the clade is *M. floridensis*, which is described as a meiotic parthenogen (Handoo *et al.*, 2004) but counter-intuitively nests as sister taxon to *M. incognita*. Clade I is sister to clade II, which contains *M. hapla*, *Meloidogyne partityla*, *M. dunensis*, *M. microtyla*, *M. maritima* and *M. duytsi*. The dfit analysis also includes in this group *M. graminis* and *M. ardenensis*. Clades I and II form a clade with respect to clade

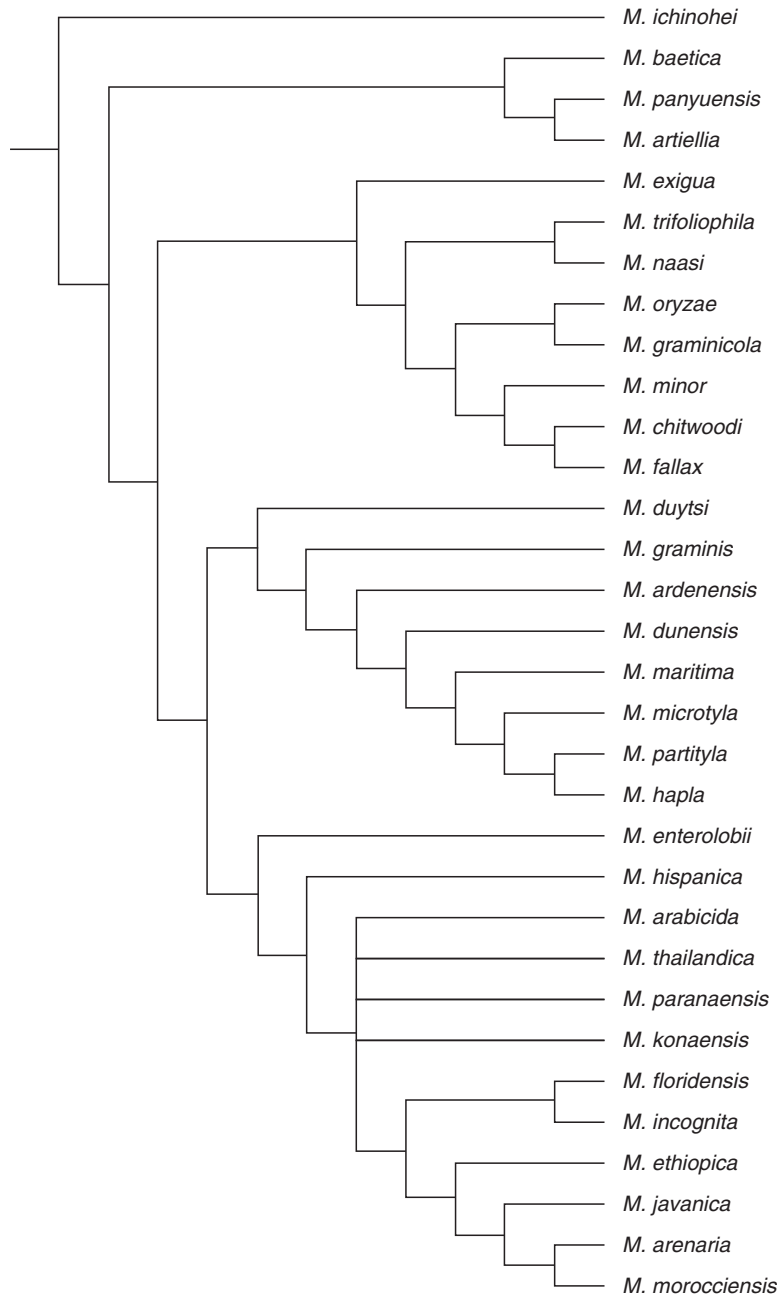

Fig. 5.15. (Continued) C: dfit combinable components consensus;

III, which contains *M. graminicola*, *M. chitwoodi*, *M. exigua*, *M. trifoliophila*, *M. naasi*, *Meloidogyne minor*, *M. fallax* and *M. oryzae*.

Tenente *et al.* (2004) showed that there was considerable variation in the 28s region among their sampled taxa (52 differences between *M. chitwoodi* and its nearest taxon). The analyses of Tigano *et al.* (2005) and De Ley *et al.* (2002) reveal even more amazing variation at the 18s locus. At the conservative end of this spectrum, *M. hispanica*

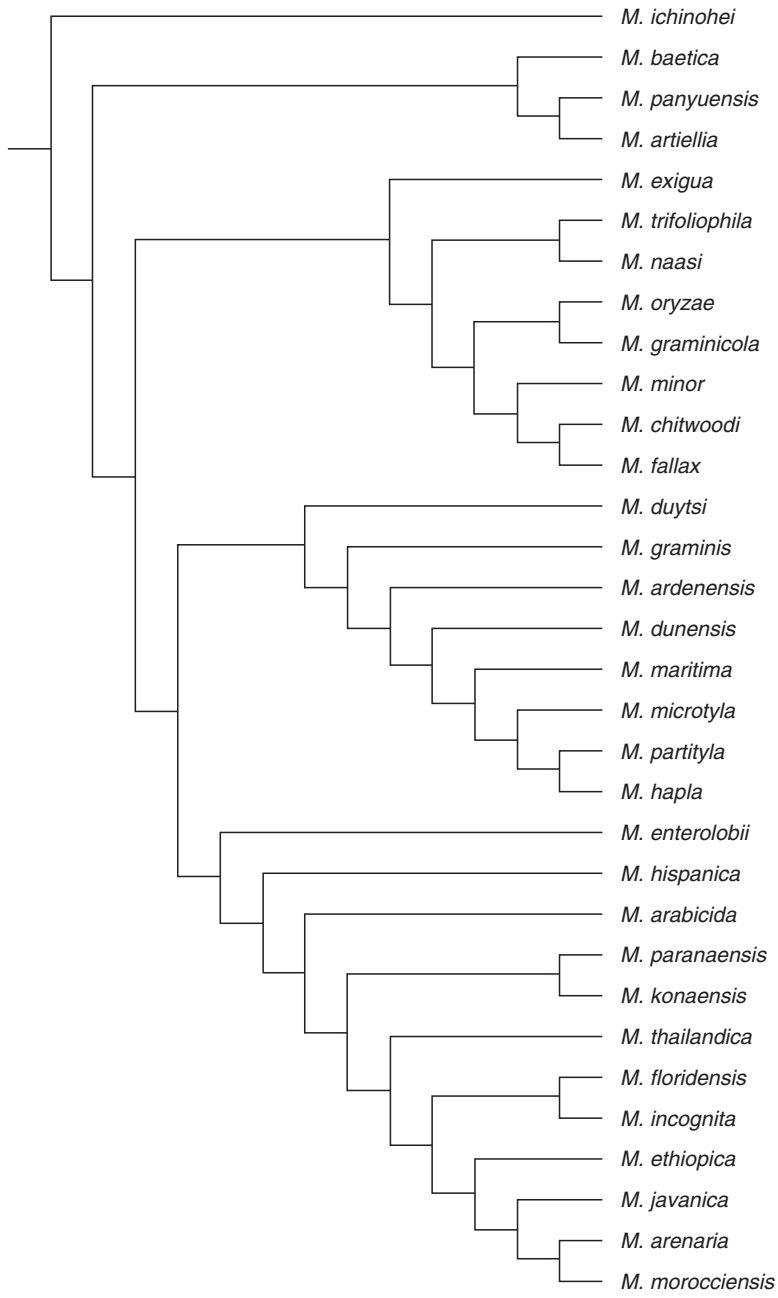

Fig. 5.15. (Continued) D: dfit majority rule consensus tree.

and *M. ethiopica* sequences are identical (Landa *et al.*, 2008). At the other end is the fact that there can be more variation between two species of *Meloidogyne* (say, *M. maritima* and *M. artiellia*) than between a human and a platypus (63 differences). Many of the differences are attributable to base-call ambiguity that could be resolved with more persistent sampling and sequencing effort (probably attributable to intraspecific and intra-individual variation). Still, such discrepancies are

astonishing when one considers that two organisms identified by nematode taxonomists as the same species (i.e. two isolates of *M. javanica*) have far greater 18s rDNA sequence divergence (24 substitutions, 4 indels) than an Australian wombat and a North American possum (3 substitutions, 2 indels). Certainly nematodes are reported to have high rates of sequence divergence relative to morphological evolution (Stein *et al.*, 2003; Sudhaus and Kiontke, 2007), and *Meloidogyne* spp. are part of a rapidly evolving clade (Holterman *et al.*, 2006), but such high divergence at the 18s locus among closely related species, and even among individuals within a species, is sufficient cause for further investigation into the origin(s) of *Meloidogyne* species and the morphological and molecular characters considered informative for diagnosis and species delimitation.

There are several reasons why the published 18s and mtDNA analyses might have returned different topologies for the Clade I mitotic parthenogens. The first, and most obvious, is that they represent discordant evolutionary histories. This notion was first empirically tested by Tigano *et al.* (2005), who performed an incongruence length difference test (Farris *et al.*, 1994, 1995) (but see Hipp *et al.*, 2004; Barker and Lutzoni, 2002; Darlu and Lecointre, 2002) on their mtDNA and 18s rDNA sequences and found significant differences, sufficient to conclude that the two data sets did not share a common evolutionary history. This observation could simply be due to the fundamental nature of the two markers. 18s sequences reside within a tandemly repeated cistron that is subject to mutation and gene conversion through concerted evolution, which could result in paralogous, and not orthologous, gene sequences (Slowinski and Page, 1999). Alternatively, it is possible that lineage sorting of mitochondrial haplotypes, mutation rate heterogeneity or sampling error from the small number of phylogenetically informative nucleotide bases examined could result in discordance between gene and species trees (Maddison, 1997; Funk and Omland, 2003; Avise, 2007).The problem of resolving phylogenetic relationships among the mitotic parthenogenetic species goes far beyond differential lineage sorting and gene conversion, and is most certainly compounded by their probable hybrid origins. It has long been suspected that *M. incognita*, *M. arenaria* and *M. javanica* arose through hybridization events between sexual or meiotic parthenogenic taxa (Triantaphyllou, 1985; Castagnone-Sereno *et al.*, 1993).

What will it take to achieve a fully resolved, robust *Meloidogyne* phylogeny with near complete representation of all its species? First, and perhaps most importantly, it will take a rigorous sampling effort of both genes and taxa. This Herculean step requires thorough field sampling of genetic variation across the globe, collaborative research involving experts in both morphological and molecular identification, and the resources and will to generate enormous amounts of DNA sequence data for each species. Secondly, it must be recognized that because the relative amount of DNA sequence divergence is so varied between taxa (some extremely high, some extremely low), phylogenetic analysis of the whole group will require use of suites of genes that evolve very slowly for deep nodes and very rapidly for shallow nodes of the tree. It is likely that genes appropriate for resolving relationships among closely related species will be inappropriate, if not completely alignment ambiguous, among distantly related species. Thus, phylogenomic and total evidence analyses using concatenated data sets will probably be highly informative for resolving relationships among deeper nodes, but unless they can also sample variation within and between populations, meta-analyses and supertree construction are likely to be required to assemble a tree with the greatest explanatory power.

The observed inability to obtain monophyletic relationships among ITS rDNA sequences (Hugall *et al.*, 1999) and/or 'alleles' of putative single-copy nuclear loci (Lunt, 2008) from hybridogenic lineages is completely consistent with the expected fate of such genes upon phylogenetic analysis. In fact, it is highly likely that even the resolution of these genealogies exhibited by the phylogenomic analysis of Scholl and Bird (2005) would dissolve upon further sampling of 'allelic' variation among additional individuals of these species from disparate populations. Simply put, the evolutionary lineages that comprise these lineages may not have unique evolutionary origins or fates and thus are not only intractable phylogenetically but also ontologically (Ghiselin, 1997; Adams, 2001). Regardless, the resolution of historical relationships among the genes that comprise the 'species' remains the single most powerful tool in the arsenal of comparative methods for understanding the origin and evolution of what are arguably the most perplexing, and vexing, nematodes on earth.

5.6. References

Adams, B.J. (2001) The species delimitation uncertainty principle. *Journal of Nematology* 33**,** 153–160.

Adams, B.J. and Powers, T.O. (1996) An rDNA phylogenetic analysis of the genus *Meloidogyne*: a recent origin of mitotic parthenogenetic species. *Nematropica* 26, 193–194 (abstract).

Avise, J.C. (2007) *On Evolution*. The Johns Hopkins University Press, Baltimore, Maryland.

Barker, F.K. and Lutzoni, F.M. (2002) The utility of the incongruence length difference test. *Systematic Biology* 51, 625–637.

Baum, T.J., Gresshoff, P.M., Lewis, S.A. and Dean, R.A. (1994) Characterization and phylogenetic analysis of 4 root-knot nematode species using DNA amplification fingerprinting and automated polyacrylamide gel electrophoresis. *Molecular Plant-Microbe Interactions* 7, 39–47.

Bininda-Emonds, O.R.P. (2004) *Phylogenetic Supertrees: Combining Information to Reveal the Tree of Life*. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Bininda-Emonds, O.R.P., Gittleman, J.L. and Purvis, A. (1999) Building large trees by combining phylogenetic information: a complete phylogeny of the extant Carnivora (Mammalia). *Biological Reviews* 74, 143–175.

Bininda-Emonds, O.R.P., Jones, K.E., Price, S.A., Cardillo, M.G.R. and Purvis, A. (2004) Garbage in, garbage out: data issues in supertree construction. In: Bininda-Emonds, O.R.P. (ed.) *Phylogenetic Supertrees: Combining Information to Reveal the Tree of Life*. Kluwer Academic Publishers, Dordrecht, the Netherlands, pp. 267–280.

Blok, V.C., Wishart, J., Fargette, M., Berthier, K. and Phillips, M.S. (2002) Mitochondrial DNA differences distinguishing *Meloidogyne mayaguensis* from the major species of tropical root-knot nematodes. *Nematology* 4, 773–781.

Brito, J., Powers, T.O., Mullin, P.G., Inserra, R.N. and Dickson, D.W. (2004) Morphological and molecular characterization of *Meloidogyne mayaguensis* isolates from Florida. *Journal of Nematology* 36, 232–240.

Carneiro, R.M.D.G., Almeida, M.R.A. and Quénéhervé, P. (2000) Enzyme phenotypes of *Meloidogyne* spp. populations. *Nematology* 2, 645–654.

Castagnone-Sereno, P., Piotte, C., Uijthof, J., Abad, P., Wajnberg, E., Vanlerberghe Masutti, F., Bongiovanni, M. and Dalmasso, A. (1993) Phylogenetic relationships between amphimictic and parthenogenetic nematodes of the genus *Meloidogyne* as inferred from repetitive DNA analysis. *Heredity* 70, 195–204.

Castillo, P., Vovlas, N., Subbotin, S. and Troccoli, A. (2003) A new root-knot nematode, *Meloidogyne baetica* n. sp (Nematoda : Heteroderidae), parasitizing wild olive in Southern Spain. *Phytopathology* 93, 1093–1102.

Creevey, C.J. and Mcinerney, J.O. (2005) Clann: investigating phylogenetic information through supertree analyses. *Bioinformatics* 21, 390–392.

Dalmasso, A. and Bergé, J.B. (1978) Molecular polymorphism and phylogenetic relationship in some *Meloidogyne* spp.: application to taxonomy of *Meloidogyne*. *Journal of Nematology* 10, 323–332.

Darlu, P. and Lecointre, G. (2002) When does the incongruence length difference test fail? *Molecular Biology and Evolution* 19, 432–437.

De Ley, I.T., De Ley, P., Vierstraete, A., Karssen, G., Moens, M. and Vanfleteren, J. (2002) Phylogenetic analyses of *Meloidogyne* small subunit rDNA. *Journal of Nematology* 34, 319–327.

De Ley, P., Felix, M.A., Frisse, L.M., Nadler, S.A., Sternberg, P.W. and Thomas, W.K. (1999) Molecular and morphological characterisation of two reproductively isolated species with mirror-image anatomy (Nematoda : Cephalobidae). *Nematology* 1, 591–612.

Dickson, D.W., Huisingh, D. and Sasser, J.N. (1971) Dehydrogenases, acid and alkaline phosphatases, and esterases for chemotaxonomy of selected *Meloidogyne, Ditylenchus, Heterodera,* and *Aphelenchus* spp. *Journal of Nematology* 3, 1–16.

Eisen, J.A. (1998) Phylogenomics: improving functional predictions for uncharacterized genes by evolutionary analysis. *Genome Research* 8, 163–167.

Eisen, J.A. and Fraser, C.M. (2003) Phylogenomics: intersection of evolution and genomics. *Science* 300, 1706–1707.

Eisenback, J.D. and Triantaphyllou, H.H. (1991) Root-knot nematodes: *Meloidogyne* species and races. In: Nickle, W.R. (ed.) *Manual of Agricultural Nematology*. Marcel Dekker, New York, pp. 191–274.

Esbenshade, P.R. and Triantaphyllou, A.C. (1985) Use of enzyme phenotypes for identification of *Meloidogyne* species. *Journal of Nematology* 17, 6–20.

Esbenshade, P.R. and Triantaphyllou, A.C. (1987) Enzymatic relationships and evolution in the genus *Meloidogyne* (Nematoda, Tylenchida). *Journal of Nematology* 19, 8–18.

Esbenshade, P.R. and Triantaphyllou, A.C. (1990) Isozyme phenotypes for the identification of *Meloidogyne* species. *Journal of Nematology* 22, 10–15.

Farris, J.S. (1983) The logical basis of phylogenetic analysis. In: Platnick, N.I. and Funk, V.A. (eds) *Advances in Cladistics*. Columbia University Press, New York, pp. 570–572.

Farris, J.S., Kallersjo, M., Kluge, A.G. and Bult, C. (1994) Testing significance of incongruence. *Cladistics* 10, 315–319.

Farris, J.S., Källersjö, M., Kluge, A.G. and Bult, C. (1995) Constructing a significance test for incongruence. *Systematic Biology* 44, 570–572.

Funk, D.J. and Omland, K.E. (2003) Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology Evolution and Systematics* 34, 397–423.

Ghiselin, M.T. (1997) *Metaphysics and the Origin of Species*. State University of New York Press, Albany, New York.

Guindon, S. and Gascuel, O. (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology* 52, 696–704.

Handoo, Z.A., Nyczepir, A.P., Esmenjaud, D., van der Beek, J.G., Castagnone-Sereno, P., Carta, L.K., Skantar, A.M. and Higgins, J.A. (2004) Morphological, molecular, and differential-host characterization of *Meloidogyne floridensis* n. sp (Nematoda : Meloidogynidae), a root-knot nematode parasitizing peach in Florida. *Journal of Nematology* 36, 20–35.

Hillis, D.M. and Dixon, M.T. (1991) Ribosomal DNA: molecular evolution and phylogenetic inference. *The Quarterly Review of Biology* 66, 411–453.

Hipp, A.L., Hall, J.C. and Sytsma, K.J. (2004) Congruence versus phylogenetic accuracy: revisiting the incongruence length difference test. *Systematic Biology* 53, 81–89.

Holterman, M., van der Wurff, A., van den Elsen, S., van Megen, H., Bongers, T., Holovachov, O., Bakker, J. and Helder, J. (2006) Phylum-wide analysis of SSU rDNA reveals deep phylogenetic relationships among nematodes and accelerated evolution toward crown clades. *Molecular Biology and Evolution* 23, 1792–1800.

Huelsenbeck, J.P. and Ronquist, F. (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17, 754–755.

Huelsenbeck, J.P., Joyce, P., Lakner, C. and Ronquist, F. (2008) Bayesian analysis of amino acid substitution models. *Philosophical Transactions of the Royal Society B – Biological Sciences* 363, 3941–3953.

Hugall, A., Moritz, C., Stanton, J. and Wolstenholme, D.R. (1994) Low, but strongly structured mitochondrial-DNA diversity in root-knot nematodes (*Meloidogyne*). *Genetics* 136, 903–912.

Hugall, A., Stanton, J. and Moritz, C. (1997) Evolution of the AT-rich mitochondrial DNA of the root knot nematode, *Meloidogyne hapla. Molecular Biology and Evolution* 14, 40–48.

Hugall, A., Stanton, J. and Moritz, C. (1999) Reticulate evolution and the origins of ribosomal internal transcribed spacer diversity in apomictic *Meloidogyne. Molecular Biology and Evolution* 16, 157–164.

Kluge, A.G. (1997) Testability and the refutation and corroboration of cladistic hypotheses. *Cladistics – the International Journal of the Willi Hennig Society* 13, 81–96.

Kornberg, R.D. (2007) The molecular basis of eukaryotic transcription. *Proceedings of the National Academy of Sciences of the USA* 104, 12955–12961.

Landa, B.B., Rius, J.E.P., Vovlas, N., Carneiro, R., Maleita, C.M.N., Abrantes, I.M.D. and Castillo, P. (2008) Molecular characterization of *Meloidogyne hispanica* (Nematoda, Meloidogynidae) by phylogenetic analysis of genes within the rDNA in *Meloidogyne* spp. *Plant Disease* 92, 1104–1110.

Longhorn, S.J., Foster, P.G. and Vogler, A.P. (2007) The nematode-arthropod clade revisited: phylogenomic analyses from ribosomal protein genes misled by shared evolutionary biases. *Cladistics* 23, 130–144.

Loytynoja, A. and Milinkovitch, M.C. (2003) A hidden Markov model for progressive multiple alignment. *Bioinformatics* 19, 1505–1513.

Lunt, D.H. (2008) Genetic tests of ancient asexuality in root knot nematodes reveal recent hybrid origins. *Bmc Evolutionary Biology* 8, 194.

Maddison, W.P. (1993) Missing data versus missing characters in phylogenetic analysis. *Systematic Biology* 42, 576–581.

Maddison, W.P. (1997) Gene trees in species trees. *Systematic Biology* 46, 523–536.

Maddison, W.P. and Knowles, L.L. (2006) Inferring phylogeny despite incomplete lineage sorting. *Systematic Biology* 55, 21–30.

Maddison, W.P. and Maddison, D.R. (2002) *MacClade version 4.0.5*. Sinauer, Sunderland, Massachusetts.

Morrison, D.A. and Ellis, J.T. (1997) Effects of nucleotide sequence alignment on phylogeny estimation: a case study of 18S rDNAs of Apicomplexa. *Molecular Biology and Evolution* 14, 428–441.

Nguyen, K.B., Maruniak, J. and Adams, B.J. (2001) Diagnostic and phylogenetic utility of the rDNA internal transcribed spacer sequences of *Steinernema*. *Journal of Nematology* 33, 73–82.

Powers, T.O. (2004) Nematode molecular diagnostics: from bands to barcodes. *Annual Review of Phytopathology* 42, 367–383.

Powers, T.O. and Harris, T.S. (1993) A polymerase chain reaction method for identification of five major *Meloidogyne* species. *Journal of Nematology* 25, 1–6.

Powers, T.O. and Sandall, L.J. (1988) Estimation of genetic divergence in *Meloidogyne* mitochondrial DNA. *Journal of Nematology* 20, 505–511.

Powers, T.O., Harris, T.S. and Hyman, B.C. (1993) Mitochondrial DNA sequence divergence among *Meloidogyne incognita*, *Romanomermis culicivorax*, *Ascaris suum*, and *Caenorhabditis elegans*. *Journal of Nematology* 25, 564–572.

Powers, T.O., Todd, T.C., Burnell, A.M., Murray, P.C.B., Fleming, C.C., Szalanski, A.L., Adams, B.A. and Harris, T.S. (1997) The rDNA internal transcribed spacer region as a taxonomic marker for nematodes. *Journal of Nematology* 29, 441–450.

Powers, T.O., Mullin, P.G., Harris, T.S., Sutton, L.A. and Higgins, R.S. (2005) Incorporating molecular identification of *Meloidogyne* spp. into a large-scale regional nematode survey. *Journal of Nematology* 37**,** 226–235.

Purvis, A. (1995) A modification to Baum and Ragan's method for combining phylogenetic trees. *Systematic Biology* 44, 251–255.

Roberts, R.G. and Bobrow, M. (1998) Dystrophins in vertebrates and invertebrates. *Human Molecular Genetics* 7, 589–595.

Roberts, R.G., Nicholson, L., Bone, Q. and Bobrow, M. (1995) Phylogenetic analysis of vertebrate and invertebrate dystrophins and utrophins. *Journal of Medical Genetics* 32, 142–142.

Roberts, T.M. (2005) Major sperm protein. *Current Biology* 15, R153.

Scholl, E.H. and Bird, D.M. (2005) Resolving tylenchid evolutionary relationships through multiple gene analysis derived from EST data. *Molecular Phylogenetics and Evolution* 36, 536–545.

Slowinski, J.B. and Page, R.D.M. (1999) How should species phylogenies be inferred from sequence data? *Systematic Biology* 48, 814–825.

Stanton, J., Hugall, A. and Moritz, C. (1997) Nucleotide polymorphisms and an improved PCR-based mtDNA diagnostic for parthenogenetic root-knot nematodes (*Meloidogyne* spp.). *Fundamental and Applied Nematology* 20, 261–268.

Stein, L.D., Bao, Z., Blasiar, D. *et al.* (2003) The genome sequence of *Caenorhabditis briggsae*: a platform for comparative genomics. *Public Library of Science – Biology* 1, 166–192.

Subbotin, S.A., Sturhan, D., Vovlas, N., Castillo, P., Tanyi Tambe, L., Moens, M. and Baldwin, J.G. (2006) Application of secondary structure model of rRNA for phylogeny: D2-D3 expansion segments of the LSU gene of plant-parasitic nematodes from the family Hoplolaimidae Filipjev, 1934. *Molecular Phylogenetics and Evolution* 43, 881–890.

Sudhaus, W. and Kiontke, K. (2007) Comparison of the cryptic nematode species *Caenorhabditis brenneri* sp n. and *C. remanei* (Nematoda : Rhabditidae) with the stem species pattern of the *Caenorhabditis elegans* group. *Zootaxa*, 45–62.

Swofford, D.L. (2002) *PAUP˚: Phylogenetic Analysis Using Parsimony (˚and Other Methods)*. Sinauer Associates, Sunderland, Massachusetts.

Tenente, G.C.M.V., De Ley, P., De Ley, I.T., Karssen, G. and Vanfleteren, J.R. (2004) Sequence analysis of the D2/D3 region of the large subunit RDNA from different *Meloidogyne* isolates. *Nematropica* 34, 1–12.

Terry, M.D. and Whiting, M.F. (2005) Behavior of POY versus CLUSTALL across a broad parameter landscape. *Cladistics* 21, 272–281.

Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) Clustal-W – improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22, 4673–4680.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. (1997) The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* 25, 4876–4882.

Thorley, J.L. and Page, R.D M. (2000) RadCon: phylogenetic tree comparison and consensus. *Bioinformatics* 16, 486–487.

Tigano, M.S., Carneiro, R., Jeyaprakash, A., Dickson, D.W. and Adams, B.J. (2005) Phylogeny of *Meloidogyne* spp. based on 18S rDNA and the intergenic region of mitochondrial DNA sequences. *Nematology* 7, 851–862.

Triantaphyllou, A.C. (1966) Polyploidy and reproductive patterns in the root-knot nematode *Meloidogyne hapla. Journal of Morphology* 118, 403–414.

Triantaphyllou, A. (1985) Cytogenetics, cytotaxonomy and phylogeny of root-knot nematodes. In: Sasser, J.N. and Carter, C.C. (eds) *An Advanced Treatise on* Meloidogyne, Volume 1. North Carolina State University Graphics, North Carolina, Raleigh, North Carolina, pp. 113–126.

Trudgill, D.L. and Blok, V.C. (2001) Apomictic, polyphagous root-knot nematodes: exceptionally successful and damaging biotrophic root pathogens. *Annual Review of Phytopathology,* 39, 53–77.

van der Beek, J.G., Folkertsma, R., Zijlstra, C., Van Koert, P.H.G., Poleij, L.M. and Bakker, J. (1998) Genetic variation among parthenogenetic *Meloidogyne* species revealed by AFLPs and 2D-protein electrophoresis contrasted to morphology. *Fundamental and Applied Nematology* 21, 401–411.

Watson, J.D. (2008) *Molecular Biology of the Gene*. Benjamin-Cummings, New Jersey.

Wiens, J.J. (1998) Does adding characters with missing data increase or decrease phylogenetic accuracy? *Systematic Biology* 47, 625–640.