Phylogenetic Pattern, Evolutionary Processes and Species Delimitation in the Genus *Echinococcus*

Alan J. Lymbery, Freshwater Fish Group and Fish Health Unit, School of Veterinary and Life Sciences, Murdoch University, Murdoch 6150, Western Australia, Australia

Contents

1. Introduction 3
2. Species of *Echinococcus* 3
3. Species concepts and species delimitation 7
4. Phylogenetic pattern 9
5. Evolutionary processes 11
   5.1. Adult worms reproduce by both self-fertilisation and cross-fertilisation 12
   5.2. Genetic diversity within populations reflects time since colonisation 14
   …5.3 Gene flow prevents differentiation among populations in the absence of geographical or ecological barriers 16
6. Are currently described species of *Echinococcus* evolutionarily independent? 17
   6.1. *Echinococcus oligarthra* and *E. vogeli* 18
   6.2. *Echinococcus granulosus* and *E. felidis* 18
   6.3. *Echinococcus equinus* 19
   6.4. *Echinococcus multilocularis* and *E. shiquicus* 19
   6.5. *Echinococcus canadensis* 20
   6.6. *Echinococcus ortleppi* 23
7. A coalescent-based approach to species delimitation 23
8. Biogeography and speciation 26
9. Phenotypic consequences of speciation 29
10. Conclusions 32
Abstract

An accurate and stable alpha taxonomy requires a clear conception of what constitutes a species and agreed criteria for delimiting different species. An evolutionary or general lineage concept defines a species as a single lineage of organisms with a common evolutionary trajectory, distinguishable from other such lineages. Delimiting evolutionary species is a two-step process. In the first step, phylogenetic reconstruction identifies putative species as groups of organisms that are monophyletic (share a common ancestor) and exclusive (more closely related to each other than to organisms outside the group). The second step is to assess whether members of the group possess genetic exchangeability (where cohesion is maintained by gene flow among populations) or ecological exchangeability (where cohesion is maintained because populations occupy the same ecological niche). Recent taxonomic reviews have recognised nine species within the genus *Echinococcus*. Phylogenetic reconstructions of the relationships between these putative species using mtDNA and nuclear gene sequences show that for the most part these nine species are monophyletic, although there are important incongruences that need to be resolved. Applying the criteria of genetic and ecological exchangeability suggests that seven of the currently recognised species represent evolutionarily distinct lineages. The species status of *E. canadensis* and *E. ortleppi* could not be confirmed. Coalescent-based analyses represent a promising approach to species delimitation in these closely related taxa. It seems likely, from a comparison of sister species groups, that speciation in the genus has been driven by geographic isolation, but biogeographic scenarios are largely speculative and require further testing.

**Key words:** species concepts, phylogeny; population genetics; speciation; biogeography
1. Introduction

Taeniid tapeworms (Eucestoda: Cyclophyllidae: Taeniidae) are important parasites of people throughout the world. Although as many as 13 genera have been described in the family, the most recent taxonomic revision recognised only four; *Hydatigera*, *Taenia*, *Versteria* and *Echinococcus* (Nakao et al., 2013a). The genus *Echinococcus* is a monophyletic group of species characterised by small adult worms and larvae (metacestodes) with extensive asexual reproduction. Definitive hosts are carnivores, usually canids or felids, and infection is acquired by eating herbivorous or omnivorous intermediate hosts. Humans are accidental intermediate hosts, with the infection being known as echinococcosis or hydatid disease. There are three different types of echinococcosis, which result from infection with different species of *Echinococcus* and are named for the structure of the metacestode; cystic, alveolar or polycystic. Cystic and alveolar echinococcosis are major public health issues in many countries throughout the world and are recognised as neglected parasitic zoonoses (Moro and Schantz, 2009; Torgerson, 2013).

Classification and nomenclature within the genus *Echinococcus* have long been controversial topics, but in recent years molecular phylogenetic analyses have promised a resolution to this controversy. In this paper, I will briefly review the taxonomic history and currently accepted taxonomic designations within the genus, attempt to define an appropriate species concept, examine both the phylogenetic and population genetic data that are required to correctly delimit species according to that concept, apply criteria for delimitation to currently described species and, finally, explore the phenotypic consequences of genetic variation among species.

2. Species of *Echinococcus*
Prior to the widespread application of molecular genetic techniques, a total of 16 species and 13 subspecies had been described in the genus based on morphology, but most of these taxa were subsequently invalidated by Rausch (1953), Vogel (1957), Rausch and Nelson (1963) and Schantz et al. (1976), leaving only four valid species: *E. granulosus* (with the subspecies *E. g. granulosus* and *E. g. canadensis*); *E. multilocularis* (with the subspecies *E. m. multilocularis* and *E. m. sibiricensis*); *E. oligarthra* (= *oligarthus*); and *E. vogeli*. The term “strain” was used to refer to intraspecific variants of uncertain taxonomic status, including many of the invalidated taxa (Thompson, 1986; Thompson and Lymbery, 1988). Strains were initially described from differences in host occurrence, geographic distribution, morphology and developmental biology, with most strains ascribed to the species *E. granulosus*. Molecular genetic studies, based principally on partial sequencing of mtDNA, clarified the extent of strain variation, leading to a “genotype” nomenclature of intraspecific variants within *E. granulosus*: G1 (sheep strain); G2 (Tasmanian sheep strain); G3 (buffalo strain); G4 (horse strain); G5 (cattle strain); G6 (camel strain); G7 (pig strain); G8 (American cervid strain); and G10 (European or Fennoscandian cervid strain) (Bowles et al., 1992, 1994; Bowles and McManus, 1993).

By the mid-1990’s it was suggested from phylogenetic analyses of both morphological (Lymbery, 1992) and mtDNA sequence data (Bowles et al., 1995; Lymbery, 1995) that *E. granulosus* was a paraphyletic group, with some strains being more closely related to *E. multilocularis* than to other strains within the complex. Initial taxonomic revisions split *E. granulosus* (sensu lato) into three species; *E. equinus* for G4, *E. ortleppi* for G5 and *E. granulosus* sensu stricto for G1, G2 and G3 (Thompson et al., 1995; Thompson and McManus, 2002). Nakao et al. (2007) subsequently proposed that the remaining strains (G6, G7, G8 and
be afforded species status as *E. canadensis*. In addition to these taxonomic changes, Xiao et al. (2005) described a new species (*E. shiquicus*), closely related to *E. multilocularis*, from Tibet, and Hüttner et al. (2008) resurrected the species *E. felidis*, originally described by Ortlepp (1934, 1937) from South Africa. As a consequence of these revisions and new descriptions, nine valid species are currently recognised in the genus *Echinococcus* (Table 1).

While substantial progress has been made in elucidating the species-level taxonomy of *Echinococcus*, the current species designations are unlikely to be the final word. Genetically differentiated populations within described species may at some stage warrant further taxonomic revision. For example, Lymbery et al. (2015) suggested that *E. canadensis* may consist of three separate species, based on differences in mtDNA sequences, morphology and life history of (apparently) sympatric populations. Substantial genetic diversity has been found among geographically separated populations of *E. multilocularis* (Nakao et al., 2009), *E. vogeli* (Santos et al., 2012) and *E. oligarthra* (Soares et al., 2013); at present there is no basis for taxonomic recognition of these differences, but that may change with further studies. There may also be undiscovered species, particularly in parts of the world where the parasite fauna has been less well studied than in Europe and North America. *Echinococcus shiquicus*, for example, was recently described from the Tibetan Plateau (Xiao et al., 2005). Prior to this study, the adult stage of *E. shiquicus* in Tibetan foxes (*Vulpes ferrilata*) had been regarded as a different morphological form of *E. multilocularis*, while the larval stage in the plateau pika (*Ochotona curzoniae*) was misidentified as *E. granulosus* (Xiao et al., 2006). In Africa, early taxonomists described a number of species of *Echinococcus* which were subsequently invalidated; recent studies, however, have highlighted that the genetic diversity of *Echinococcus* spp. in African wildlife is
far from clear and it is possible that some of these invalidated taxa may need to be revisited (Romig et al., 2011; Wassermann et al., 2015)

Resolving controversies about species status and inferring the taxonomic rank of newly discovered variant populations requires a clear conception of what constitutes a species and what criteria can be validly used to delimit species. This has not always been the case with respect to the species-level taxonomy of *Echinococcus*. For example, Thompson *et al.* (1995) in initially proposing a taxonomic revision in the genus, favoured an evolutionary species concept; a species is a single lineage of ancestral descendant populations which maintains its identity from other such lineages and which has its own evolutionary tendencies and historical fate (Wiley 1978). Most subsequent taxonomic studies, however, have not identified a species concept or proposed clear criteria for delimiting different species. In many cases, where there have been no obvious differences in morphology between putative species, a genetic yardstick has been used; species status has been recommended if taxa are as genetically different from each other as the difference between well-established species in the same or related genera (e.g. Thompson and McManus, 2002; Nakao *et al.*, 2007; Hüttner *et al.*, 2008). While a genetic yardstick approach may sometimes provide a useful guide to species status, it is essentially an appeal to authority, or at least to past practice. Like all general rules for historical inference it may provide a misleading picture of lineage separation, because of large stochastic variation in evolutionary processes over time and space (Knowles and Carstens, 2007). Nakao *et al.* (2010a, 2013b), in the most recent reviews of the genus, proposed the application of a phylogenetic species concept, defining species as the smallest set whose members are descended from a common ancestor. In practice, however, this definition does not lead to workable criteria for delimiting species (i.e. for determining what is the “smallest set”), and Nakao *et al.* (2010a, 2013b) variously used a genetic
yardstick, the presence (or absence) of fixed morphological and ecological characters, and evidence of lack of gene flow to determine which set of intraspecific variants should be accorded species status.

While deciding on an appropriate species concept may seem to be an arcane academic exercise, in reality it has great practical value if we wish to have an accurate and stable alpha taxonomy. Accurate delimitation of species provides the foundation of our knowledge of parasite life history, distribution and disease processes (Hoberg, 2006). In addition, controversy and instability in species names can have important public health implications. For example, *E. canadensis*, but not *E. granulosus* s.s., is found in Canada (Moro and Schantz, 2009). Government health regulations, however, do not make the distinction between *E. granulosus* s.s. and *E. granulosus* s.l. Officially, Canada is considered endemic for *E. granulosus* and import requirements for dogs into Canada do not require treatment with an anthelmintic. Such treatment could prevent the introduction of *E. granulosus* s.s., a species of far more economic and public health significance than *E. canadensis* (Lymbery et al., 2015).

3. Species concepts and species delimitation

An accurate and stable alpha taxonomy requires agreement about what the term species actually means; without this there can be objective way of deciding whether one particular proposal for species-level nomenclature is any more valid than another proposal. For such a fundamental unit of biological organisation, there has been a surprising amount of debate as to what constitutes a species, with at least 24 different species concepts having been proposed (Mayden, 1997). Most of these concepts, however, differ in their criteria for delimiting species rather than in their
theoretical understanding of how species exist. The basic theoretical view of a species in almost all concepts is that species represent independently evolving lineages (De Queiroz, 1999, 2007; Hey, 2006). This general lineage concept of species is essentially an updating of the evolutionary species concept and I will use the operational definition suggested by Lymbery and Thompson (2012): a species is a single lineage of organisms with a common evolutionary trajectory, distinguishable from other such lineages.

The evolutionary species concept is applicable to all organisms, regardless of their mode of reproduction or breeding system. Because there are two parts to the concept (single lineage and common evolutionary trajectory), delimiting evolutionary species requires two steps. The first step is phylogenetic reconstruction to determine the pattern of evolutionary relationships among lineages. Putative species identified through phylogenetic reconstruction should be both monophyletic (sharing a common ancestor) and exclusive (more closely related to each other than to any organisms outside the group), as these two properties are not necessarily congruent in a reticulate (as opposed to a purely diverging) genealogy (de Queiroz and Donoghue, 1990; Velasco 2009). While monophyly and exclusivity are necessary conditions for a group of organisms to have species status, they are not of themselves sufficient, because they do not address the processes responsible for maintaining a cohesive evolutionary trajectory within the group. The second step in the delimitation of evolutionary species is therefore to assess whether members of the group possess genetic exchangeability (where cohesion is maintained by gene flow among populations) or ecological exchangeability (where cohesion is maintained because populations occupy the same ecological niche and selective regime) (Templeton, 1989; Crandall et al., 2000).
In practice, the lack of exchangeability among groups is much more easily inferred when they are sympatric than when they are allopatric. Groups which maintain fixed genetic differences in sympatry can be confidently asserted to lack genetic exchangeability and therefore to be different species. In allopatry, however, lack of exchangeability is usually inferred from the extent of genetic, phenotypic or life history differences among groups, which is a much more subjective exercise. The distinction between sympatric and allopatric populations is not as clear in parasites as in free-living organisms, as parasite populations may be physically separated by different host associations even within the same geographic area (McCoy, 2003; Huyse et al., 2005). When considering the capacity for genetic exchangeability among parasite populations, therefore, a distinction should be made between occurrence in the same geographic region (which I will call broad sympatry) and occurrence in the same host associations in a region (strict sympatry).

Do the groups which have been conferred species status in the genus *Echinococcus* fit the criteria for evolutionary species? Answering this question requires us to consider both the phylogenetic pattern among groups in the genus and the evolutionary processes responsible for maintaining a cohesive evolutionary trajectory within each group.

### 4. Phylogenetic pattern

The phylogeny of genetic variants within the genus *Echinococcus* has been reconstructed using both mtDNA sequences (e.g. Le et al., 2002; McManus et al., 2002; Obwaller et al., 2004; Lavikainen et al., 2003, 2006; Thompson et al., 2006; Nakao et al., 2007, 2013b,c; Hüttner et al., 2008; Moks et al., 2008) and nuclear DNA sequences (e.g. Lavikainen et al., 2003; Bart et al., 2006; Saarma et al., 2009; Knapp et al., 2011). Figures 1 and 2 show the best current estimates of
these phylogenies, from studies using a number of different sequences and a complete (or almost complete) range of described genetic variants. There are a number of common features to these phylogenies; *E. felidis* and *E. granulosus* sensu stricto are clearly sister species, and *E. ortleppi* is closely related to the different genotypes of *E. canadensis*. However, there are also some major discrepancies. While the mtDNA phylogenies appear quite robust, there are differences between the mtDNA and nuclear DNA phylogenies and among the nuclear DNA phylogenies themselves. These differences principally concern the position of the South American species *E. vogeli* and *E. oligarthra* (basal or non-basal), relationships of clades in the *E. granulosus* s. l. complex (monophyletic or paraphyletic), relationships between the different genotypes of *E. canadensis* (monophyletic or paraphyletic) and whether *E. multilocularis* and *E. shiquicus* are sister species. With respect to the basal position of *E. vogeli* and *E. oligarthra*, the paraphyletic nature of the *E. granulosus* s.l. complex and the sister species relationship between *E. multilocularis* and *E. shiquicus*, the nuclear DNA phylogenies of Knapp et al. (2011) are in agreement with the mtDNA phylogenies of Nakao et al. (2007, 2013b,c) and Hüttner et al. (2008), while the phylogeny of Saarma et al. (2009) suggests a different topology. With respect to *E. canadensis*, all phylogenies indicate a monophyletic origin, except for that based on exon sequences of nuclear DNA genes by Knapp et al. (2011), which suggests that the G6 genotype is more closely related to *E. ortleppi* than to the other genotypes of *E. canadensis*.

Clearly, further analyses are required, using more nuclear DNA sequences, to completely resolve the pattern of relationships among putative species within the genus. Nakao et al. (2013b) argued that broad agreement between the mtDNA phylogenies and the complete nuclear phylogeny of Knapp et al. (2011) provided a solid foundation for delimiting species. From these phylogenies (Figure 1 and Figure 2b), it appears that each of the nine currently described species forms a
monophyletic and exclusive group. They therefore fulfill the first criterion to be recognised as separate evolutionary species. This is not evidence, however, that they fulfill the second criterion of being independent lineages, i.e. on separate evolutionary trajectories. To determine this, we need to consider the evolutionary processes operating within the groups.

5. Evolutionary processes

In contrast to the large number of studies which have aimed at reconstructing the phylogeny of species of Echinococcus, the study of population genetic structure has been relatively neglected. This is unfortunate, because analysing the distribution of genetic variation within and among populations of a species can provide information on evolutionary processes such as gene flow, genetic drift and selection, and on the biological factors, such as mode of reproduction, breeding system, effective population size and dispersal ability, which influence these evolutionary processes.

It has often been suggested that the reproductive biology of species of Echinococcus, with a combination of self-fertilisation and extensive asexual reproduction, has a profound effect on evolutionary processes, leading to the genetic uniformity of local populations and rapid genetic differentiation among populations subject to different selection pressures (Smyth and Smyth, 1964; McManus and Smyth, 1986; Bryant and Flockhart, 1986; Haag et al., 2008; Nakao et al., 2009, 2010a, 2013b). This is thought to occur because the population genetic consequences of obligate self-fertilisation and asexual reproduction are almost complete homozygosity, extensive linkage disequilibrium (non-random association of alleles at different loci) and a distribution of genetic diversity between, rather than within family groups, which will lead to spatial structuring
of genetic variation if dispersal is limited. The empirical population genetic data, although not extensive, suggest a much more complex picture than this.

5.1 Adult worms reproduce by both self-fertilisation and cross-fertilisation

Substantial deficiencies of heterozygous genotypes have been reported in populations of *E. granulosus* in Australia (Lymbery and Thompson, 1988; Lymbery et al., 1990, 1997), and *E. granulosus* and *E. ortleppi* in Brazil (Badaraco et al., 2008). These heterozygote deficiencies cannot be accounted for by unobserved population structuring (Wahlund effect), at least for the Australian samples (see Lymbery et al., 1997), and the most likely explanation is predominant self-fertilisation. Cross-fertilisation must also occur, however, because heterozygous genotypes have been found, not only in *E. granulosus* (Lymbery and Thompson, 1988; Lymbery et al., 1990, 1997; Badaraco et al., 2008), but also in *E. multilocularis* (Nakao et al., 2003; Knapp et al., 2007) and *E. vogeli* (Santos et al., 2012). Furthermore, Haag et al. (1999; 2011) found no evidence of heterozygote deficiencies in populations of *E. granulosus* in Brazil in two separate studies.

It appears, then, that both cross-fertilisation and self-fertilisation occur within species of *Echinococcus*. In organisms with alternating sexual and asexual stages in their life cycle, self-fertilisation can be achieved in two ways; ova may be fertilized by sperm of the same individual (autogamy) or by the sperm of clonally identical individuals (geitonogamy). Both self-insemination and cross-insemination have been observed microscopically in *Echinococcus* spp. (Smyth and Smyth, 1969; Kumaratilake et al., 1986; Wang, 1998). The interesting question is: what is the relative frequency of cross-fertilisation (mating between genetically different
individuals) versus self-fertilisation through either autogamy or geitonogamy? At present, the genetic data are too few for any definitive answer; three studies in Australia and one in Brazil gave found high rates of self-fertilisation, while two studies in Brazil suggest cross-fertilisation is the norm. Lymbery et al. (1997) argued that, while high rates of self-fertilisation were found in Australian populations of *E. granulosus*, an absence of linkage disequilibrium indicated that this selfing rate is an accident of colonisation (i.e. due to low genetic diversity in the founding population, leading to geitonogamy) rather than the natural breeding system of the species.

Cross-fertilisation requires definitive hosts to be multiply infected with genetically different metacestodes and for adult worms derived from these metacestodes to be in contact in the small intestine. Multiple infections are likely to be common, since the distribution of infected definitive hosts is typically overdispersed (e.g. Lahmar et al., 2001; Ziadinov et al., 2010), and worms in infrapopulations are aggregated within the small intestine, which will increase contact rates (Lymbery et al., 1989). The important issue is whether definitive hosts are often concurrently infected with genetically different cysts. A number of studies have used mtDNA markers to demonstrate mixed infections of different species and strains of *Echinococcus* in individual dogs and wolves (Stefanic et al., 2004; Xiao et al., 2006; Zhang et al., 2006; Schurer et al., 2014). More pertinently, nuclear DNA markers have been used to find different genotypes of *E. multilocularis* in individual foxes and *E. granulosus* in individual dogs. Genetically different *E. multilocularis* worms were found in in 38% of 13 foxes in Hokkaido, Japan (Nakao et al., 2003), 52% of 25 foxes in France (Knapp et al., 2008) and 35% of 125 foxes throughout Europe (Knapp et al., 2009), while genetically different *E. granulosus* worms were found in 50% of dogs in Argentina (Haag et al., 2011). These data suggest that it is not unusual for definitive hosts to be infected with genetically different worms of the same species.
In summary, the available evidence indicates that species of *Echinococcus* normally have a mixed mating system, with sexual reproduction by both self-fertilisation and cross-fertilisation. Mixed mating systems are common in both hermaphroditic plants (Goodwillie et al., 2005) and animals (Jarne and Auld, 2005) and present an evolutionary conundrum because theoretical studies predict the evolution of mating systems towards either pure selfing or pure outcrossing (Lande and Schmenske, 1985). The most common explanation for the maintenance of mixed mating systems is the reproductive assurance hypothesis; outcrossing is usually favoured by selection, but selfing can be favoured if mate availability varies, because selfing is better than not mating at all (Holsinger, 1996). For species of *Echinococcus*, the most parsimonious explanation of the available data is that adult worms normally reproduce by mating with other individuals, with high rates of self-fertilisation due to geitonogamy rather than autogamy. Geitonogamy will be common where infection rates of definitive hosts are low, so that multiple infections are rare, or where genetic diversity has been reduced, for example through recent colonisation of an area by a small number of founders. Cross-fertilisation will be common where infection rates and genetic diversity are high.

5.2 Genetic diversity within populations reflects time since colonisation

Initial studies on *E. granulosus* in Australia found much lower levels of genetic diversity at isozyme loci than reported for other species of parasitic helminths or for free-living organisms, consistent with a founder effect from recent colonisation (Lymbery and Thompson, 1988; Lymbery et al., 1990; Lymbery, 1995). Subsequently, much more comprehensive analyses of mtDNA variation in *E. granulosus* from the Middle East, Europe, China, South Asia, Africa,
South America and Australia have found relatively low levels of nucleotide diversity, although greater levels of haplotype diversity, again suggesting rapid population expansion from small founder populations (Moro et al., 2009; Nakao et al., 2010b; Casulli et al., 2012; Yanagida et al., 2012; Sharma et al., 2013). A comparison of haplotype diversities suggests a Middle Eastern origin for this species (or at least for the switch from a wildlife to a domestic host cycle), with subsequent spread to other parts of the world through anthropogenic transport of hosts (Nakao et al., 2013b).

There have been far fewer studies on other species, but Knapp et al. (2009) found microsatellite allelic diversity in *E. multilocularis* in Europe to be greatest in historically endemic areas centred on Switzerland and lower in northern and eastern regions where the parasite has only been recently recorded. Nakao et al. (2010b) found low levels of mtDNA diversity in *E. multilocularis* in the Tibetan Plateau and suggested that the species was introduced to this region by a recent range extension of the red fox (*Vulpes vulpes*), although this hypothesis requires confirmation because there does not seem to be independent evidence of such a range extension (Kutschera et al., 2013). However, relatively high levels of mtDNA diversity have been found for *E. shiquicus* in the Tibetan Plateau (Nakao et al. 2010b) and *E. vogeli* in Brazil (Santos et al. 2012). Both *E. shiquicus* and *E. vogeli* are principally maintained in wildlife cycles and have presumably been present in their respective areas for the last 1.5- 3.0 M years (based on the chronogram of Knapp et al. 2011).

In summary, the data suggest that genetic diversity is greatest in species that persist in (presumably ancestral) wildlife cycles, while in species that cycle (at least partially) in
domesticated hosts, genetic diversity can vary widely and is positively related to the time since colonisation of a region.

5.3 Gene flow prevents differentiation among populations in the absence of geographic or ecological barriers

The data on this point are rather sparse because not many studies have measured genetic variation at the appropriate scale, i.e. within and among local populations of a species. Those data that are available, however, indicate that most genetic variation occurs within, rather than among, geographically defined populations.

In *E. granulosus*, Lymbery et al. (1997) found only 2.4% of total genetic variance between populations separated by approximately 3,500 km in mainland Australia; Haag *et al.* (1999) found effectively no variance among localities in Southern Brazil, Australia and Europe (Germany, Poland, Spain, Switzerland and Ireland); and Casulli *et al.* (2012) found 4% of variance among populations from different countries in eastern Europe (Bulgaria, Hungary and Romania). More substantial genetic structuring has been found in some cases. Lymbery et al. (1997) found 5.8% of genetic variance between populations of *E. granulosus* in mainland Australia and the island state of Tasmania, Casulli *et al.* (2012) found 18.7% of genetic variance between populations in Hungary and Italy, and Wang *et al.* (2014) reported 9.3% of variance among populations in different regions of south-western China. In a few cases, the majority of genetic variation has been found to be distributed among, rather than within populations. Haag *et al.* (2004) found 40.8% of genetic variance in *E. granulosus* from different geographical regions (north, central, south) of Argentina, although nested clade analysis indicated that the most likely
cause of this differentiation was historical differences in the time and origin of livestock introductions (carrying different parasite genotypes), rather than restricted gene flow among regions. On a more local scale, Haag et al. (2011) compared genetic diversity within and among dogs from different farms in the same rural area of southern Brazil and found the majority of variance (61.9%) among farms.

There have been very few studies on the genetic structure of other species of *Echinococcus*. Phylogenetic analysis of mDNA sequences identified geographically distinct clades in Europe, Asia and North America (Nakao et al., 2009). Within Europe, Knapp et al. (2009) measured genetic variation at a microsatellite locus in *E. multilocularis* from foxes in Switzerland, France, Germany, Austria, Czech Republic, Slovakia and Poland. Although they did not statistically partition genetic variation at different hierarchical levels, there was no evidence of geographic clustering of genotypes and no relationship between genetic and geographic distance among regions. In contrast, Santos et al. (2012) found substantial structuring (39% of total genetic variance) in nuclear and mitochondrial markers among populations of *E. vogeli* from eastern and western regions of the Brazilian Amazon, separated by approximately 2,500 km.

In summary, despite the capacity for self-fertilisation and asexual reproduction in species of *Echinococcus*, the available population genetic data suggest that the majority of genetic variation is usually found within, rather than among local populations, with the extent of spatial genetic structuring determined by host vagility.

6. Are currently described species of *Echinococcus* evolutionarily independent?
6.1. *Echinococcus oligarthra* and *E. vogeli*

These taxa are basally placed in most phylogenetic trees and sister species in the nuclear gene phylogeny of Knapp et al. (2011). Both taxa occur throughout South and Central America, often in the same geographic locality (e.g. in Columbia; D’Alessandro and Rausch, 2008), where they maintain consistent differences in nuclear and mtDNA sequences, as well as in adult morphology and host occurrence; *E. oligarthra* using mainly wild felids as definitive hosts, with a wide intermediate host range, while *E. vogeli* is found principally in bush dogs (*Speothos venaticus*) and pacas (*Cuniculus paca*), although there is some overlap in intermediate host occurrence (D’Alessandro and Rausch, 2008; Nakao et al., 2013b). This is strong evidence for a lack of ecological exchangeability between the taxa, confirming their status as different evolutionary species. If *E. oligarthra* and *E. vogeli* are sister species and/or basally placed within the phylogeny of *Echinococcus*, then logically each must also be a different evolutionary species to all other taxa in the genus.

6.2. *Echinococcus granulosus* and *E. felidis*

*Echinococcus granulosus* and *E. felidis* are sister taxa on all mitochondrial and nuclear phylogenies. They occur in the same localities in southern and eastern Africa, with fixed differences in both mtDNA and nuclear DNA sequences, as well as consistent differences in adult morphology and definitive host occurrence; usually canids for *E. granulosus* and lion (*Panthera leo*) or spotted hyena (*Crocuta crocuta*) for *E. felidis* (Hüttner et al., 2008; Kagendo et al., 2014). Separation of host cycles is not complete, however, and eggs of both taxa have been found (using
molecular markers) from a single lion faecal sample (Kagendo et al., 2014). This indicates a lack of both genetic and ecological exchangeability between the taxa, and therefore confirms their status as separate evolutionary species. As they are sister taxa, then they must also be evolutionarily independent of all other taxa in the genus.

6.3. *Echinococcus equinus*

In the phylogenetic tree of Saarma et al. (2009), *E. equinus* clusters with *E. ortleppi* and *E. canadensis*, but in all other phylogenies it forms a separate branch, with no closely related sister taxa. It is distinct genetically and morphologically from all other taxa, and has a strong intermediate host preference for equids (Thompson et al., 1995; Thompson and McManus, 2006). It is found in broad sympatry with *E. granulosus*, and probably also with *E. felidis*, *E. ortleppi*, *E. canadensis* and *E. multilocularis* (Thompson et al., 1995; Romig et al., 2015; Wassermann et al., 2015), although strict sympatry has not been demonstrated in all cases. Nevertheless, the consistent morphological and ecological distinctiveness of *E. equinus* over a wide geographic area is evidence of a lack of ecological exchangeability with all other taxa, and therefore of separate species status.

6.4. *Echinococcus multilocularis* and *E. shiquicus*

In most (but not all) phylogenies, *E multilocularis* and *E. shiquicus* are sister taxa and positioned within the *E granulosus* s.l. complex. The two taxa occur sympatrically in the Tibetan plateau, with a partial, but not complete, separation of life cycles. Mixed infections of *E. multilocularis* and *E. shiquicus* have been found in the Tibetan fox (*Vulpes ferrilata*) (Jiang et al., 2012) and the
plateau pika (Xiao et al., 2006). Importantly, differences in nuclear DNA sequences, as well as mtDNA sequences, have been found in strict sympatry, providing strong evidence for a lack of genetic exchangeability (Xiao et al., 2005). This is supported by small, but consistent differences in adult morphology and by different metacestode structures (alveolar cysts for \emph{E. multilocularis} and unilocular cysts for \emph{E. shiquicus}, which are maintained even in coinfections of the same intermediate host (Xiao et al., 2006). If \emph{E. multilocularis} and \emph{E. shiquicus} are different evolutionary species, and if they are sister taxa, as indicated in most phylogenetic trees, then they must also be separate evolutionary species from all other taxa in the genus.

\textbf{6.5. \emph{Echinococcus canadenis}}

There is still some doubt about the monophyletic origin of \emph{E. canadensis}. While phylogenies from mtDNA and some nuclear gene sequences show \emph{E. ortleppi} as a sister group to the genotypes of \emph{E. canadensis} (Nakao et al., 2007, 2013b; Saarma et al., 2009), other nuclear phylogenies are unable to resolve the relationship (Knapp et al., 2011). If the genotypes of \emph{E. canadensis} are not monophyletic, they clearly cannot be regarded as a single evolutionary species, but if we accept the mtDNA phylogeny as a true reconstruction of evolutionary history, then they may be.

Nakao et al. (2007; 2013b,c) proposed the unification of the G6, G7, G10 and G8 genotypes into the species \emph{E. canadensis}, while Lymbery et al. (2015) suggested that there were three different evolutionary lineages within this group; G6/G7 as one species, G10 as a second and G8 as a third. Both proposals are consistent with the phylogenetic pattern shown by analyses of mtDNA sequence data, as indeed would be a third proposal; to consider each genotype as a separate
species. The challenge is to find evidence that the taxa either have or lack genetic or ecological exchangeability.

The G6 and G7 genotypes are largely allopatric, making it difficult to establish the presence or absence of genetic exchangeability (Lymbery et al., 2015). Although the metacestode of G6 is often found in camels, while that of G7 is often found in pigs, there is overlap in life cycles, as well as similarity in morphological traits and close genetic similarity in mtDNA and nuclear DNA sequences (Nakao et al., 2007, 2013b,c; Lymbery et al., 2015; Romig et al., 2015). These data are all consistent with ecological exchangeability between the G6 and G7 genotypes (and hence conspecific status), although, as pointed out by Romig et al. (2015), there remains the possibility of biologically important variation between these groups.

The G6 genotype has been found in the same area in Far East Russia as the G10 (and G8) genotypes (Konyaev et al. 2013), although strict sympatry has not been demonstrated. In addition, the genotypes in this region have only been characterised at non-recombining mtDNA loci, so a lack of genetic exchangeability cannot be assumed. Lack of ecological exchangeability is suggested by differences in intermediate hosts (principally domestic livestock for G6/G7 and cervids for G10; Thompson, 2008), although this separation is not complete, as G6 has been found in reindeer (Konyaev et al., 2013). Morphological differences between protoscoleces and adult worms of pig or camel origin and those of cervid origin (both G8 and G10 genotypes) are also suggestive of a lack of ecological exchangeability (Lymbery et al., 2015), although some morphological traits can be substantially influenced by intermediate host origin (Lymbery, 1998).
The G8 and G10 genotypes have been found in strict sympatry; not only in the same species of intermediate and definitive hosts in the same localities in northern Eurasia and North America (Thompson et al., 2006; Moks et al., 2008; Konyaev et al., 2013; Schurer et al., 2014), but adult worms of both genotypes have been found in the same individual wolf (Schurer et al., 2014). This is not necessarily an indication of a lack of genetic exchangeability, because these worms were only genotyped at mitochondrial (non-recombining) loci, so it is not clear if gene flow among the genotypes is restricted in sympatry. It could be argued that morphological and ecological (life cycle) differences between these mitochondrial lineages indicate a lack of genetic exchangeability (Lymbery et al., 2015), but confirmation of this would require some knowledge of genetic and environmental effects on these traits.

At present, it is difficult to say whether *E. canadensis*, as currently described, represents a single evolutionary species or several species. The question of genetic exchangeability between the G8 and G10 genotypes could possibly be settled by a population genetic study (using nuclear markers) in sympatry, but strict sympatric occurrence is rare and, even if genetic differences were found at nuclear loci, it may be that the two mitochondrial lineages have only recently come into contact through natural or artificial host movements, leaving insufficient time to detect genetic signatures of introgression (Nakao et al., 2015). The G6, G7 and G10 genotypes are largely allopatric, so genetic exchangeability cannot be tested and arguments for species status rest either on a genetic yardstick or on similarities and differences in life cycles and morphology. These are necessarily subjective judgments. Lymbery et al. (2015) suggested that an integrative or iterative approach (see Padial et al., 2010; Yeates et al., 2011), using molecular, ecological and morphological data together, might provide more objectivity, but unfortunately we know very little about the heritability of morphological and ecological traits in *Echinococcus* spp. The best
way to resolve this issue may be to use multilocus genotypic data in a coalescent-based approach to species delimitation (see section 7).

6.6. *Echinococcus ortleppi*

*Echinococcus ortleppi* is clearly closely related to the genotypes of *E. canadensis*. If we assume that the genotypes of *E. canadensis* are monophyletic and can be regarded as a single species, then we need to consider whether *E. ortleppi* represents a lineage on a different evolutionary trajectory (of course, if the genotypes of *E. canadensis* are regarded as separate species, then they do not need to form a monophyletic group to the exclusion of *E. ortleppi*). *Echinococcus ortleppi* co-occurs with the G6 and G7 genotypes in Africa, South America and Europe (Thompson and McManus, 2002; Tigre et al., 2016) but it is not clear whether they are strictly sympatric and therefore whether genetic exchangeability can be tested. There are differences in morphology and host occurrence, but, as described for *E. canadensis*, the genetic basis of these differences has not been established. Nakao et al. (2007; 2013b,c) used a genetic yardstick to suggest that *E. ortleppi* and *E. canadensis* are separate species; the genetic distances at mtDNA loci between *E. ortleppi* and the genotypes of *E. canadensis* are similar to the distance between *Taenia saginata* and *T. asiatica*, the most closely related known sister species within the family. As previously discussed, while the genetic yardstick may be useful in some situations, it is an arbitrary and unreliable criterion on its own, and more evidence of a lack of ecological exchangeability is required to confidently assign *E. ortleppi* as a separate evolutionary species.

7. A coalescent-based approach to species delimitation
The delimitation of evolutionary species is reasonably straightforward when taxa are sympatric or have well-defined genetic, morphological and ecological differences in allopatry. It is much more problematic, however, for cases such as *E. ortleppi* and the genotypes of *E. canadensis*, where lineage separation appears to be recent or incomplete, so that we cannot recognise fixed diagnostic states or reciprocal monophyly (concordance of all gene trees). In situations such as this, a coalescent-based approach might provide a more objective means of delimitation. Coalescent theory is a mathematical formulation of the genealogical history of neutral alleles in a population and, when considering a number of different populations, the multispecies coalescent can be used to describe a probability distribution for gene trees evolving along the branches of a species tree (Degnan and Rosenberg, 2009). Because coalescent approaches are probabilistic, they allow for gene tree discordance when testing alternative hypotheses of lineage separation (Fujuta et al., 2012).

Coalescent theory tells us that the time at which two alleles share a most recent common ancestor (coalescence time) is a function of the demographic history of the population (Emerson et al., 2001). Individuals that belong to the same species will therefore share demographic parameters such as population size (\( \theta \), the product of effective population size and mutation rate) and lineage divergence time (\( \tau \)). This provides the basis for using a collection of gene trees to evaluate the probabilities of different species delimitation models, i.e. of determining which of several alternative species trees best explains the gene tree data (Rannala and Yang, 2003; Knowles and Carstens, 2007). For example, if we consider the genotypes of *E. canadensis*, there are four different species trees that are compatible with the mtDNA phylogeny and these species trees contain different numbers of population demographic parameters (Figure 3). Bayesian or
maximum likelihood approaches can be used to evaluate the fit of parameter estimates from gene
trees to those expected from different species trees (Fujita et al., 2012).

Coalescent-based species delimitation requires a number of methodological considerations. First,
multilocus data are needed on a number of different individuals of each putative species.
Simulation studies suggest that 5-10 individuals, sequenced at each of 5-10 independent loci
provide sufficient power for species delimitation (Zhang et al., 2011), although a larger number
of individuals may need to be sampled for genetically diverse taxa. Sampling should also
encompass the geographical range of the putative species and incorporate any existing
information on diversity. Second, multispecies coalescent models assume neutrality, so the loci
chosen for study should not be subject to strong natural selection (Degnan and Rosenberg, 2009).
Empirical studies of species delimitation with coalescent-based methods have used both
mitochondrial and nuclear loci, and coding and non-coding sequences (e.g. Camargo et al., 2012;
Harrington and Near, 2012; Niemiller et al., 2012), but loci which are likely to be under active
selection pressure may produce misleading signals of lineage divergence. Third, separate gene
trees need to be constructed for each locus. Although it has been quite common to concatenate
DNA sequences from different loci prior to phylogenetic analysis of species of Echinococcus
(e.g. Nakao et al., 2007, 2013b,c; Saarma et al., 2009; Knapp et al., 2011; but see Hüttner et al.,
2008), this results in a loss of information and will not necessarily yield a more reliable species
tree. When gene trees from different loci are discordant, as is expected when species have
recently diverged, then concatenation across loci can actually provide misleading inferences
about lineage separation (Kubatko and Degnan, 2007). Finally, different analytical methods
should be compared for species delimitation. A range of different software packages are
available, using different approaches to evaluate the probability of alternative species trees (e.g. Pons et al., 2006; Ence and Carstens, 2010; O’Meara, 2010; Yang and Rannala, 2010).

8. Biogeography and speciation

Acceptance of a particular species concept constrains our view of how speciation occurs. If species are regarded as lineages with separate evolutionary trajectories, then speciation must involve the evolution of traits which limit genetic or ecological exchangeability. There is abundant theoretical and empirical evidence that for the majority of free-living organisms, speciation usually occurs as the result of genetic drift or adaptive divergence between allopatric (geographically separated) populations (Turelli et al. 2001; Coyne and Orr 2004). In parasites, however, there has been a long-held view that sympatric speciation, mediated by host switching leading to ecological isolation within the same geographic region, is relatively more common (Price 1980; de Meeûs et al. 1998; Kunz 2002; Huyse et al. 2005).

It has been suggested that the mode of reproduction in Echinococcus spp. (high levels of selfing in definitive hosts and asexual reproduction in intermediate hosts) might predispose to speciation through the rapid (essentially instantaneous) production of genetically different forms (Smyth and Smyth, 1964; Kumaratilake and Thompson, 1982; McManus and Smyth, 1986). This seems rather unlikely, given current knowledge of the breeding system and population genetic structure of species of Echinococcus (see Section 5.1), but the issue of speciation in the genus has rarely been addressed empirically. Speciation can almost never be directly observed, but must be inferred from the current pattern of evolutionary relationships among different species and from the evolutionary processes operating within and among contemporary populations. The sympatric
occurrence of sister species provides an initial indication that speciation may have occurred by host switching. Although there is some disagreement among studies, *E. granulosus* and *E. felidis*, *E. multilocularis* and *E. shiquicus*, *E. canadensis* and *E. ortleppi*, and *E. vogeli* and *E. oligarthra* have all been identified as sister species in at least one phylogeny of the genus, and all of these pairs except *E. canadensis* and *E. ortleppi* have been shown to be broadly sympatric. Current distribution, however, does not necessarily reflect distribution during the time that speciation occurred, and a historical biogeographic perspective is necessary to infer the processes by which speciation has occurred.

For parasitic organisms, biogeographical patterns have to be interpreted within a coevolutionary context and a number of recent studies have attempted such an interpretation for species of *Echinococcus*. If the basal position of the neotropical species *E. oligarthra* and *E. vogeli* in most (but not all) phylogenies is taken as a starting point, then it has been argued that the ancestor of these species invaded South America in carnivorous (canid or felid) definitive hosts with the opening of the Panamanian land bridge around 3 Mya (Knapp et al., 2011). Using this time for calibration of their nuclear exon phylogeny, Knapp et al. (2011) estimated that the *Echinococcus* clade began to diversify from other taeniids around 5.8 M years ago (Figure 4). Nakao et al. (2013b) extrapolated from this to suggest that the ancestor of the *Echinococcus* clade may have originated in either North America, where modern canids evolved approximately 10 M years ago or in Asia where modern felids arose approximately 11 M years ago. Genetic diversity among species of *Echinococcus*, compared to that among other taeniids (Knapp et al. 2011), suggests that speciation and global radiation in the genus has occurred recently and rapidly, although that of course assumes constancy in the molecular clock across the different genera.
Geographic comparisons of genetic diversity have been used to suggest that *E. granulosus* evolved in the Middle East, coincident with the domestication of sheep about 10,000 – 12,000 years ago, and subsequently spread worldwide with livestock movements (Yanagida et al., 2012; Nakao et al., 2013b; Sharma et al., 2013). Huttner et al. (2008) suggested an Asian origin for *E. felidis*, as modern Felidae apparently arose in Asia, prior to spreading into Africa approximately 3 M years ago (Johnson et al., 2006; although the demographic history of lions appears to have involved multiple movements between Africa and Asia, see Barnett et al., 2014). As *E. felidis* and *E. granulosus* are sister species, then presumably they shared a common ancestor in Asia prior to this time, and divergence of the lineages occurred in geographic isolation.

Nakao et al. (2009) suggested an origin for *E. multilocularis* in the region of the land bridge between Asia and North America (Beringia) during the Pleistocene (2.6 M – 11,700 years ago), with subsequent spread via foxes across North America, Asia and Europe. The North American, Asian and European clades are hypothesized to have then undergone cycles of isolation in glacial refugia and dispersal from these refugia during inter-glacial periods (Nakao et al. 2013b). In Europe, for example, Nakao et al. (2009) proposed an introduction of *E. multilocularis* in the late Pleistocene, with contraction and subsequent spread from glacial refugia, such as the Iberian, Italian and Balkan peninsulas, presumably initially to the genetically diverse Alpine arch region of Switzerland, southern Germany, eastern France and Austria (Knapp et al., 2009, 2015).

*Echinococcus shiquicus* is thought to have differentiated from *E. multilocularis* in isolation, presumably during a period of range contraction, with the current sympatric occurrence a result of recent introduction of *E. multilocularis* into the Tibetan Plateau with red foxes (Nakao et al., 2010b).
These biogeographic scenarios are all plausible with respect to the putative timing of species splitting events (Figure 4) and they can be related to the taxon pulse theory, whereby cyclical episodes of isolation and range expansion create opportunities for a complex interplay of cospeciation, host switching and coadaptation (Hoberg and Brooks, 2008, 2010). However, they are all essentially historical narratives, which have not yet been tested empirically. What is required to move these narratives into a more objective realm is to derive from them explicit predictions which can be tested against phylogeographic data. To avoid circularity, it is important that the data used to test the predictions are independent of those used to derive them (Crisp et al., 2011). Until we can get to this stage, hypotheses about the timing and mode of speciation in the genus *Echinococcus* will remain speculative.

**9. The phenotypic consequences of speciation**

Under an evolutionary (or general lineage) species concept, lineages are recognised as different species when they are on different evolutionary pathways. We therefore expect species to diverge phenotypically, as a result of genetic drift or selective responses to the environment. These phenotypic differences may include traits of clinical or epidemiological importance, so the differentiation of species is of importance to the treatment and control of echinococcosis.

Traits of clinical or epidemiological importance, which may differ among species, include host range, propensity to infect people, cyst structure, cyst growth rate, adult size (related to fecundity) and prepatent period (Lymbery, 1995). The genetic architecture of these traits is not known, but it seems likely that they are multifactorial, with both a genetic component, influenced by a number of polygenes or quantitative trait loci, and an environmental component. In some
cases, the genetic contribution to phenotypic differences among species appears quite obvious. The two major contributors to the burden of echinococcosis in humans are *E. granulosus* and *E. multilocularis*. These two species differ markedly in the structure of the metacestode. *Echinococcus granulosus* typically forms single chambered, unilocular cysts, surrounded by a host-produced fibrous adventitial layer (Thompson, 1995). Growth occurs by concentric enlargement and proliferation of the germinal layer is entirely endogenous (within the laminated layer). The metacestode of *E. multilocularis*, on the other hand, consists of numerous small vesicles, with no limiting adventitial layer; exogenous proliferation of the germinal layer leads to infiltrating growth through host tissue (Thompson, 1995). Although some variation exists in each of these metacestode structures, the extent and consistency of the differences between the two species clearly indicates a substantial genetic component. The infiltrative growth form of *E. multilocularis* is different to that of any other described species of *Echinococcus*, including its sister species *E. shiquicus*, so presumably arose in this lineage.

For many traits of clinical or epidemiological importance, however, it is much more difficult to determine the extent to which differences among species are due to genetic, as opposed to environmental factors. For example, approximately 88% of all documented cases of cystic echinococcosis in people are caused by *E. granulosus*, with about 11% due to infection with the G6 and G7 genotypes of *E. canadensis* and less than 1% to *E. ortleppi* and the G8 and G10 genotypes of *E. canadensis*; there have been no recorded cases of human infection with *E. equinus* or *E. felidis* (Alvarez Rojas et al., 2014). This may be indicative of fundamental differences in the infectivity of these lineages to humans, but is also very likely to be influenced by differences in exposure, which will depend on a complex interplay of geographic range and host occurrence of the parasite, and social behavior of people who may interact with these hosts.
Determining the relative importance of genetic and environmental factors to phenotypic differences among species or genotypes is challenging, because of the obvious difficulties with experimental approaches which would allow the same genotype to be expressed in different environments. Lymbery (1998) used a quantitative genetic analysis to estimate the contribution of intermediate host origin to variation of hook size and shape in protoscoleces, but this method is really only applicable to traits expressed by protoscoleces within the metacestode. Of much greater value would be the actual identification of the genes involved in traits of clinical or epidemiological importance, so they can be directly compared among species.

Some progress has been made to this end. Siles-Lucas et al. (2001), for example, found differences in the expression profile of the 14-3-3 gene, involved in abnormal cell proliferation, between *E. granulosus* and *E. multilocularis*, and suggested that this may partially explain differences in metacestode growth behavior. The recent publication of whole genome sequences for *E. granulosus* and *E. multilocularis* (Tsai et al., 2013; Zheng et al., 2013) enhances the potential of this approach. As yet, only limited interpretations have been made from the genomic data, but even these open up new avenues for research. For example, differences in regulation profiles among developmental stages of *E. granulosus* suggest candidate genes for regulating larval development and differentiation, and for strobilisation and reproduction in adults (Zheng et al., 2013); and sequence differences in the apomucin gene family between *E. granulosus* and *E. multilocularis* may indicate the genetic basis for differences in the thickness of the laminated layer in the metacestodes of these species (Tsai et al., 2013). The resources provided by these studies will undoubtedly be used much more in the future to identify candidate genes for traits of
clinical and epidemiological importance, allowing us to evaluate the extent to which differences in these traits between species are genetically determined.

10. Conclusions

When I first reviewed this topic in 1995, there were only four recognised species in the genus, with a number of strains of uncertain taxonomic status. My conclusions at that time were that the molecular data which were starting to be collected were not consistent with the prevailing view of phylogeny within the genus, but it was not yet clear how many species existed and how they were related to each other. Thanks to a large number of careful molecular phylogenetic studies in the intervening 20 years, particularly those by Minoru Nakao and his colleagues, we now have a much more robust species-level taxonomy, with nine described species.

There are still many unanswered questions, however. The transmission patterns and zoonotic potential of a number of described species have not yet been fully elucidated. Phylogenies using different nuclear DNA sequences are not fully congruent, or congruent with phylogenies from mtDNA sequences. There is debate about the species status of some intraspecific variants, for example the G6/7 genotypes in *E. canadensis*, and the recent description of *E. shiquicus* in Tibet suggests that more species may remain to be discovered. Many of these uncertainties, of course, can be resolved with more data, but there is also a need to have a clear idea of what constitutes a species and what criteria are appropriate for delimiting species in the genus. I argue that an evolutionary species concept, which regards a species as a single lineage of organisms with a common evolutionary trajectory, provides an appropriate conceptual framework.
Delimiting species under an evolutionary species concept requires consideration of both the pattern of evolutionary relationships among lineages and the processes responsible for maintaining a cohesive evolutionary trajectory. It is in an understanding of evolutionary processes, such as gene flow and ecological constraints, where we have most to learn. Population genetic studies in species of *Echinococcus* are few and far between, but those that are available indicate that most genetic variation occurs within, rather than between, local populations, with genetic structuring only apparent over larger geographic scales. These studies suggest that, despite asexual reproduction in intermediate hosts, which would tend to favour clumped transmission of clones and effective self-fertilisation (through geitonogamy) in definitive hosts, gene flow among clonal groups may be extensive. This interpretation is supported by ecological data, which, while also rather sparse, indicate multiple infections of definitive hosts and an aggregated distribution of adult worms within the small intestine of definitive hosts. These factors would tend to promote cross-fertilisation among clones and reduce the potential for geographic structuring of genotypes. Under this scenario, the absence of genetic or ecological exchangeability among clades provides good evidence that they are separate evolutionary species.

Accepting that there is still some uncertainty in the species-level phylogeny, seven of the nine currently described species of *Echinococcus* would appear to lack genetic or ecological exchangeability and therefore be on different evolutionary trajectories. I would be hesitant at this point to regard either *E. canadensis* or *E. ortleppi* as good evolutionary species. Further molecular phylogenetic analyses, comparing gene trees in a hypothesis-testing, coalescent framework, and supported by careful morphological and ecological studies, should resolve this issue.
From a practical perspective, the importance of an evolutionarily accurate alpha taxonomy is that it should provide a deeper understanding of differences among species in traits of medical or epidemiological importance and a better predictive capability for the treatment and control of echinococcosis. In 1995, I concluded that the difficulty of applying the statistical techniques of quantitative genetics to partition genetic and environmental influences on traits of medical or epidemiological importance, meant that molecular techniques were needed to locate the genes involved. That seemed a long way off at the time. However, important progress has already been made in this field and the recent publication of whole genome maps for *E. granulosus* and *E. multilocularis* heralds exciting times ahead for research on this fascinating and important genus.

**Acknowledgements**

Thanks to Andy Thompson for encouraging me to write the review and for comments on the manuscript.

**References**


heterogeneity in mitochondrial genes of cytochrome c oxidase-1 and NADH dehydrogenase-1. Parasitology 128, 569-575.


Rausch, R.L., 1953. The taxonomic value and variability of certain structures in the cestode genus *Echinococcus* (Rud., 1801) and a review of recognised species. Thapar Commemoration Volume, 233-246.


Schantz, P.M., Colli, C., Cruz-Reyes, A., Prezioso, U., 1976. Sylvatic echinococcosis in Argentina. II. Susceptibility of wild carnivores to Echinococcus granulosus (Batsch, 1786) and host-induced morphological variation. Tropenmed. Parasitol. 27, 70-78.


Velasco, J.D., 2009. When monophyly is not enough: exclusivity as the key to defining a phylogenetic species concept. Biol. Phil. 24, 473-486.


Table 1. Currently recognised species within the genus *Echinococcus* (following Nakao *et al.* 2013b). Hosts and geographic distribution are indicative. Only those hosts thought to be major contributors to transmission cycles are included. Continental range is shown, but distribution within those continents may be patchy.

<table>
<thead>
<tr>
<th>Species</th>
<th>Definitive hosts</th>
<th>Intermediate hosts</th>
<th>Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. oligartha</em>(^1)</td>
<td>Wild felids</td>
<td>Agouti</td>
<td>Central and South America</td>
</tr>
<tr>
<td><em>E. vogeli</em></td>
<td>Bush dog</td>
<td>Paca</td>
<td>Central and South America</td>
</tr>
<tr>
<td><em>E. granulosus</em>(^2)</td>
<td>Domestic dog</td>
<td>Sheep, many other ungulates</td>
<td>Cosmopolitan</td>
</tr>
<tr>
<td><em>E. felidis</em></td>
<td>Lion, hyena</td>
<td>Unknown</td>
<td>Africa</td>
</tr>
<tr>
<td><em>E. equinus</em></td>
<td>Domestic dog</td>
<td>Horse, other equids</td>
<td>Eurasia, Africa</td>
</tr>
<tr>
<td><em>E. multilocularis</em></td>
<td>Foxes</td>
<td>Arvicolid rodents</td>
<td>Central and northern Eurasia, northern North America</td>
</tr>
<tr>
<td><em>E. shiquicus</em></td>
<td>Tibetan fox</td>
<td>Pika</td>
<td>Tibetan plateau</td>
</tr>
<tr>
<td><em>E. ortleppi</em></td>
<td>Domestic dog</td>
<td>Cattle</td>
<td>Eurasia, Africa</td>
</tr>
<tr>
<td><em>E. canadensis</em></td>
<td>Domestic dog, wolves</td>
<td>Pig, camel, cervids</td>
<td>Eurasia, Africa, South America</td>
</tr>
</tbody>
</table>

\(^1\)Usually referred to as *E. oligarthrus*. Originally described as *Taenia oligartha*. The spelling of the specific epithet appears to have been changed, incorrectly, from *oligarthra* to *oligarthrus* by Cameron (1926); see Hüttner and Romig (2007) for details. By the Principle of Priority, *E. oligarthra* is the correct name.

\(^2\)This is *Echinococcus granulosus* sensu stricto (s.s.); the name *E. granulosus* sensu lato (s.l.) is often used to refer to those taxa formerly included in the species, i.e. *E. granulosus* s.s., *E. felidis*, *E. equinus*, *E. ortleppi* and *E. canadensis*. Throughout this review, the term *E. granulosus* always refers to *E. granulosus* s.s.
Figure captions

Figure 1. Cladograms of species of *Echinococcus* based on phylogenetic analyses of mtDNA sequences. (a) Analysis of concatenated sequences from 12 protein coding, two rRNA and 22 tRNA genes (Nakao et al. 2007, 2013c). (b) Analysis of concatenated sequences from three protein-coding and one rRNA gene (Hüttner et al. 2008). Shaded taxa are those formerly placed in the species *Echinococcus granulosus* s.l.

Figure 2. Cladograms of species of *Echinococcus* based on phylogenetic analyses of nuclear DNA sequences. (a) Analysis of concatenated sequences from five protein coding genes (Saarma et al. 2009). (b) Analysis of concatenated sequences (including both introns and exons) from three protein coding genes (Knapp et al. 2011). (c) Analysis of concatenated sequences (exons only) from three protein coding genes (Knapp et al. 2011). Shaded taxa are those formerly placed in the species *Echinococcus granulosus* s.l.

Figure 3. Example of alternative species trees for the genotypes of *Echinococcus canadensis*. The fully resolved phylogeny for all four genotypes, based on analyses of mtDNA by Nakao et al. 2013c) is shown in grey. The black lines indicate different species delimitation models that are compatible with the phylogeny, with associated population size (θ) and coalescence time (τ) parameters. The number of parameters varies from two, when all genotypes are regarded as a single species, to 10, when each genotype is regarded as a separate species.

Figure 4. Chronogram of species of *Echinococcus*, reconstructed from an exon dataset of nuclear protein-coding genes (Knapp et al. 2011). Time estimates are based on the assumption that the most recent common ancestor of *E. vogeli* and *E. oligarthra* dates from the opening of the Panamanian land bridge between South and North America.
A

E. canadensis G6
E. canadensis G7
E. canadensis G8
E. ortleppi
E. multilocularis
E. shiquicus
E. equinus
E. granulosus
E. felidis
E. oligarthra
E. vogeli

B

E. canadensis G6
E. canadensis G7
E. canadensis G8
E. canadensis G10
E. ortleppi
E. multilocularis
E. shiquicus
E. equinus
E. granulosus
E. felidis
E. vogeli
E. oligarthra