Overview and significance of molecular methods: what role for molecular epidemiology?

R. C. A. THOMPSON*, C. C. CONSTANTINE *and* U. M. MORGAN

WHO Collaborating Centre for the Molecular Epidemiology of Parasitic Infections, and Division of Veterinary and Biomedical Sciences, Murdoch University, Murdoch, Western Australia, Australia 6150

SUMMARY

In this chapter, the contribution of molecular tools in understanding the aetiology and ecology of infectious diseases is examined in the context of molecular epidemiology (ME). ME is seen as providing the 'tools', both laboratory and analytical, which have predictive significance in epidemiological investigations of the causation of disease. A diversity of questions can be addressed with these tools which can conveniently be viewed as particular regions of DNA and grouped according to the different hierarchical levels of specificity by which infectious agents can be characterized. These groupings and the applications of the different molecular tools are described, and consideration given to the most appropriate methods of analysing data from ME investigations.

Key words: Molecular epidemiology, infectious diseases, molecular tools, parasites.

INTRODUCTION

A variety of molecular approaches have been applied to the problem of diagnosing infectious agents, and this is well illustrated in the preceding articles in this issue. In the present chapter, we should like to examine the overall contribution of such molecular tools, and their contribution to what we know about the aetiology of infectious diseases. Such a collective assessment can be viewed in the context of molecular epidemiology (ME), which allows us to go beyond the use of molecular methods for parasite identification, and illustrates the enormous contribution molecular biology is making towards our understanding of the causation of infectious disease and its control.

ME adds another dimension to investigations of disease aetiology and provides information that classical epidemiology cannot. Thus ME is having a great impact on traditional epidemiological concepts. However, although cancer epidemiologists and microbiologists have embraced ME for at least the last decade, parasitologists have been slower to develop the field of ME. It was not until the early 1990s that ME started to figure more prominently in the parasitological literature. A key publication was that by Hide & Tait (1991) which reviewed the techniques available and their potential to answer epidemiological questions using trypanosomiasis and leishmaniasis as examples. During the subsequent 8 years, ME has impacted in exciting ways on a number of areas of parasitology and we will look at these as well as some other infectious diseases. However, before doing so we will try and define what is meant by ME.

* Corresponding Author.

WHAT IS MOLECULAR EPIDEMIOLOGY?

Interestingly, trying to define ME is seen by some as a point of controversy, particularly in medical circles, with concern about sub-specialization and the boundaries of ME investigations (e.g. McMichael, 1995; Poirier, 1995; Weston, 1995; Hall, 1996). However, these mostly semantic and discipline-orientated discussions serve to obscure the value of ME in understanding the causation of disease. The most logical way to view ME is that it provides the 'tools' (both laboratory and analytical) with predictive significance which epidemiologists can use to better define the aetiology of specific diseases; both infectious and non-infectious.

Recent technological advances have made copious genetic sequence data and precise quantifiable measurements of microorganisms and host cell populations available (Lipsitch, 1997). Using PCR, (polymerase chain reaction) defined parasite gene sequences can now be detected from small quantities of parasite material from which data can be used for diagnosis, to assess the effect of interventions on parasite population structure, assessment of intra-species diversity, and transmission studies (Bjorkman *et al*. 1998).

In this chapter, we propose to examine ME in its broadest sense, to encompass (1) the causation of parasitic disease, which includes identification of the causative agent(s) and those host and environmental factors that result in successful transmission, colonization and disease; (2) elucidating the interactions and relationships between parasites at the population level, both outside and within the host; and (3) how this information can be applied to solving problems related to control.

TOOLS EMPLOYED IN MOLECULAR EPIDEMIOLOGY

Our ability, using molecular techniques, to detect and characterize the genetic variability of infectious agents, particularly at the intraspecific level, can be seen as the foundation for the majority of molecular epidemiological studies. The application of appropriate molecular tools will aid in the surveillance of infectious agents and in determining sources of infection. Indeed, outbreak investigations and surveillance may soon be considered incomplete without the application of molecular tools, especially PCR which obviates the need to culture (Hall, 1996). As such, ME investigations principally comprise the analysis of nucleic acids and proteins in studying health and disease determinants in animal, including human, as well as plant populations. In this context, Hall (1996) and Thompson & Lymbery (1996) have emphasized the importance of appropriate analysis and the value of characterizing the genetic diversity of infectious agents at different levels of specificity. The latter requires choosing molecular tools which are capable of discriminating genetic variants at different hierarchical levels and the region of DNA examined must be appropriate to the level of questions being asked (Table 1). It should be noted that these hierarchical levels represent a continuum and different techniques or regions of DNA may be useful at different levels across taxa. This Table shows typical examples of the level of variability for regions of DNA and will not apply to every situation, and some techniques may be useful across several levels.

The level of variability detected depends on the technique used as well as the region of DNA examined. Sequencing detects the highest level of variation right down to individual base pair (bp) changes although time and/or financial constraints may limit its application to short stretches of DNA. RFLP-based techniques can detect only those changes which occur in the restriction site but can be used to sample across the whole genome. Allozymes detect only approximately 60 % of bp changes and can only examine enzyme-coding loci for which an appropriate detection system has been devised.

Highly conserved regions can reveal information about taxonomic relationships between species, moderately conserved regions can differentiate strains or closely related species; moderately variable regions can reveal population genetic structure; and highly variable regions can allow the tracing of particular fingerprinted isolates within a population (see Table 1). Ribosomal RNA (rRNA) is widely used in molecular characterization and phylogenetic studies. rRNA genes are organized into 'rRNA units' which consist of a series of tandemly repeated sequences that contain the 18s, 5·8s and 28s genes and the spacer regions that surround them. As reviewed by Appels & Honeycutt (1986), the ribosomal DNA repeats contain three types of sequences with different rates of evolution: the coding regions, which are highly conserved and have been used to make hypotheses about the basal branches in the phylogeny of life (Hasegawa *et al.* 1985); the transcribed spacer regions which are moderately conserved and have been used to reconstruct events over the past 50×10^6 years (Hillis & Davis, 1986); and the non-transcribed spacer regions which vary both within and between species (Ranzani, Bernini & Crippa, 1984; Hillis & Davis, 1986; Sites & Davis, 1989).

Discrimination above species

By far the most extensively used region of the genome has been the small sub-unit of ribosomal DNA. It has been used to create universal phylogenetic trees linking all life and causing the revision of the traditional kingdom-based classification of life (Olsen & Woese, 1993). It is also useful right down to species level and beyond, producing new revisions of major groups, such as for the phylum Nematoda (Blaxter *et al.* 1998) and the tissue cyst-forming coccidia (Tenter & Johnson, 1997). Features of mitochondrial DNA (mtDNA) such as gene order, genetic code and secondary structure of transfer RNA (tRNA) and rRNA may provide characters for phylogenetic analysis of distantly related organisms (Moritz, Dowling & Brown, 1987).

Discrimination between species

A huge variety of techniques for species diagnosis has been used as seen in the preceding articles of this volume. Moderately variable transcribed rRNA spacer regions such as the internal transcribed spacers (ITS1 and ITS2) are useful in discriminating between species. Extrachromosomal regions such as mtDNA are also useful tools for discriminating both between and within species. mtDNA contains both coding and non-coding regions and rates of sequence evolution vary along the mtDNA molecule (Harrison, 1989). Because of this intra-molecular variability, sequence comparisons of different regions of the molecule can provide useful information at a number of different levels.

Discrimination between intraspecific variants/strains

Isoenzymes/allozymes. Isoenzyme electrophoresis was one of the very early techniques used to detect the different molecular variants of enzymes, and although it can be useful for species level investigations, the majority of studies involve looking for intraspecific variation. Allozymes, a subset of isoenzymes looking only at allelic forms of the same locus, have been used extensively and played a major

Table 1. Molecular tools for characterizing genetic diversity at different hierarchical levels

Function*	Purpose	Regions of DNA
Discrimination above level of species	Taxonomy/evolution	Highly conserved coding regions e.g. SSU rDNA, features of mt DNA
Discrimination between species	Taxonomy/diagnosis/epidemiology	Moderately conserved regions e.g. coding mt DNA genes, ITS rDNA
Discrimination between intraspecific variants/strains	Population genetics/breeding systems/host adapted strains/conservation	Variable regions e.g. allozyme, RAPD, AFLP
Discrimination between individual isolates/clonal lineages/ecological interactions within host	'Fingerprinting' – tracking transmission of genotypes identifying sources of infection and risk factors	Fingerprinting techniques e.g. microsatellites
Genetic markers/linking phenotype and genotype	Identifying phenotypic traits of clinical significance	Genotype linked to phenotype via (i) Genetic map; (ii) RDA; (iii) sequencing of genes thought to be linked to phenotypic traits

* In some cases, there may be overlap between the tools (regions of DNA) used and function. This will depend on the group or organisms being studied and the level of variation detected by a particular tool (see text).

role in the establishment of population genetic concepts. Allozymes are generally only useful below the level of closely related genera, and they are a good method for comparing strains/races and species (Buth, 1984).

RAPDs. Random amplified polymorphic DNA (RAPD) or arbitrarily primed PCR (AP-PCR) was independently developed by Williams *et al.* and Welsh & McClelland in 1990. This procedure detects nucleotide sequence polymorphisms in a PCR-assay without the need for previously determined nucleic acid sequence information. RAPD analysis has been applied to many parasite populations where it has been used to differentiate isolates of *Trichinella* (Bandi *et al.* 1993; Tighe *et al.* 1994), trypanosomes (Majiwa *et al.* 1993; Steindel *et al.* 1993), *Eimeria* (MacPherson & Gajadhar 1993; Procunier, Fernando & Barta, 1993), Tritrichomonads (Fellenstein, 1998), *Toxoplasma* and *Neospora* (Guo & Johnson, 1995 a, b), *Giardia* (Morgan *et al.* 1993), and *Cryptosporidium* (Morgan *et al.* 1995). Being PCR-based, the principal limitations of RAPD fingerprinting arise from its sensitivity to reaction conditions and also to the non-specificity of the reaction as any contaminating DNA will also be amplified. Therefore, stringent procedures to avoid PCR contamination (Kwok & Higuchi, 1989) must be adhered to when utilizing this technique.

AFLPs. Amplified Fragment Length Polymorphism (AFLP) is a new DNA fingerprinting technique which allows the detection of DNA polymorphisms without prior sequence information. In AFLP analysis, genomic DNA is digested with restriction enzymes, ligated to adaptors, and a subset of DNA fragments are amplified. The selective amplification is achieved by the use of primers that extend into the restriction fragments, amplifying only those fragments in which the primer extensions match the nucleotides flanking the restriction sites. Using this method, sets of restriction fragments may be visualized by PCR without knowledge of nucleotide sequence. Typically 50–100 restriction fragments are amplified and detected on denaturing polyacrylamide gels (Vos *et al.* 1995).

AFLPs are likely to be more reliable and reproducible than RAPDs as specific primers are used and therefore high stringency amplification reactions can be performed, thereby avoiding the problems of non-specific amplification which can be a problem with RAPDs due to the low stringency amplification conditions. Research has shown that AFLPs correlate well with RFLP and SSR (microsatellite) markers (Powell *et al.* 1996). The technique has mainly been applied to plant genomes and more recently to insect (Reineke, Karlovsky & Zelitz, 1998) and bacterial genomes (Lin, Kuo & Ma, 1996), where it has successfully been used to track a nosocomial outbreak of *Acinetobacter* (Koelman *et al.* 1997). The main limitation to applying this technique to parasite genomes, particularly protozoan parasites, is the requirement for relatively large amounts of DNA (0·2–1 mg). However, for organisms where DNA can only be obtained in small amounts and for which little sequence information is available, the technique could be used as an initial screening tool on a limited number of isolates to determine useful polymorphisms which could then be cloned and specific primers designed.

Discrimination between individual isolates/clonal lineages/genotypes

For transmission and genetic mapping studies, highly variable genetic markers that are not under selection by the host need to be utilized. A variety of 'fingerprinting' techniques have been developed in

recent years, a number of which have particular application in studies on the epidemiology of parasitic infections.

Mini/microsatellites. Hypervariable loci due to their relatively high degree of variation between individuals are also useful as fingerprinting tools and are known to exist in most if not all species of plants and animals, including prokaryotic species (Nakamura *et al.* 1987) There are two main types of hypervariable loci: minisatellites (Jeffreys, Wilson & Thein, 1985) and microsatellites or simple sequence repeats (SSR) (Weber & May, 1989). Minisatellites consist of repeating motifs of between 10–100 nucleotide base pairs in length and are usually analysed with DNA probes which are specific for a single locus. Microsatellites consist of repeating motifs 2–5 bp in length and usually exhibit smaller alleles than do minisatellites. Microsatellites are more widely distributed throughout genomes, are more easily identified by cloning methods and because of their smaller allele size are more amenable to PCR.

IGS (rDNA intergenomic spacer). The rDNA intergenomic spacer region (IGS) is useful for transmission as it is the most variable region within the rRNA repeat unit. PCR amplification of the IGS region is facilitated because it is flanked by conserved coding regions. Restriction fragment length polymorphism (RFLP) analysis of the IGS region has already been developed for *Giardia* (Hopkins *et al.* unpublished) and has been shown to be a useful epidemiological tool in tracking *Giardia* in localized endemic foci of infection (Hopkins *et al.* 1999).

Genetic markers

This level does not represent an increase in detectable variability but involves linking the highly variable techniques used in the previous level to phenotypic variation. Loci affecting phenotype can be identified by several means some of which are discussed in detail below. Once the loci are identified and studied the means are available for highly applied applications such as to predict phenotype from genotype (e.g. find other organisms likely to respond to a particular drug); redesign existing drugs to work better or to act on different organisms; and design vaccines against specific parasite proteins.

Genetic mapping. The process of genetic mapping involves detecting linkage between genetic markers (such as specific RAPD bands or microsatellite alleles) with loci which influence phenotypic traits. These loci can then be examined more closely to determine the cause of observed phenotypic differences.

RDA. Representational Difference Analysis (RDA) offers an alternative to conventional genetic mapping procedures, and is a revolutionary new PCR-based technique which combines subtractive and kinetic enrichment as a means of targeting specific genetic differences between individuals (Lisitsyn, Lisitsyn & Wigler, 1993; Lisitsyn *et al.* 1994). RDA provides a way to generate genetic markers linked to a particular trait of interest without prior knowledge of the chromosomal location of the gene controlling the trait or the availability of a pre-existing genetic map. The application of such a technique to parasites is particularly exciting as it opens up the prospect of rapidly isolating markers linked to traits of clinical significance such as infectivity or drug sensitivity without time-consuming and expensive genetic mapping. A single method for defining the differences between two DNA populations, for example 2 closely-related virulent and non-virulent strains of a particular parasite, could lead to the isolation of genetic markers for virulence, and on a broader level could lead to the discovery of the genetic basis of many types of biological phenomena.

High-throughput genotyping

The evolution of both PCR and robotic technologies during the last decade has reached a point where genotyping can now be performed routinely, on a large scale, in a dedicated facility (Barlow & Ewen, 1998). The incorporation of fluorescent labels into different primers means that RAPDs, microsatellites and AFLPs can now be detected by light emitted during electrophoresis. Amplified products are differentiated by colour and also by molecular weight. By using different fluorescent labels for different primer pairs several different microsatellites/AFLPs etc. can be applied to one electrophoresis lane. The incorporation of fluorescently-labelled internal standards in each lane of every run, eliminates the problems of run to run and well to well variation. These recent advances mean that genotyping and also diagnostic tests will now be simpler and faster to perform and more reliable (Barlow & Ewen, 1998).

APPLICATIONS

Above the species level

Molecular techniques revealing conserved variation are not only useful for delimiting species, but they can also provide information above the species level (Table 1). The most common above-species investigation is determining a phylogeny; that is, finding out which species are most closely related and estimating sequence of divergence. Phylogenetic trees also provide valuable information about population dynamics and evolutionary history (Harvey & Nee, 1994). This includes both general processes and many which are specific to infectious agents. Parasites offer several interesting characteristics which make them useful to the study of evolutionary processes. First, host/parasite co-evolution, includ-

ing the possibility of co-speciation with their hosts, which can provide 2 independent measures of divergence allowing comparisons of rate of change between taxa. Secondly, each life cycle stage often lives in a unique environment and is exposed to entirely different environmental pressures. Finally, there are nested levels of organization (within the individual host, in host populations, in all host species) with unambiguous boundaries (Bush *et al.* 1995).

Recent studies are addressing basic questions such as the processes by which organisms become parasitic; for example, a new phylogeny of nematodes, using SSU rDNA, suggests that parasitism arose independently at least 4 times (Blaxter *et al.* 1998). Studies are revealing a variety of different strategies of invading the host and of evading the host's immune responses; for example *Trypanosoma brucei* and *T. cruzi* cause different human diseases, and as strategies for immune evasion, *T. brucei* undergoes antigenic variation whereas *T. cruzi* becomes an intracellular organism. This fundamental difference is reflected by major differences in their genetic organizations and recent comparisons of their gene sequences indicate that these 2 trypanosome species are highly divergent evolutionarily (Donelson, 1996). Host switching is another evolutionarily interesting process to test general theories; for example, one would expect that parasites infecting the ancestral host will be more genetically diverse than those which have recently colonized a new host species (Anderson & Jaenicke, 1997). Host switching is also of practical significance as it is one way in which new diseases may emerge.

With respect to evolutionary processes, the study of virulence and its evolution is a particular area of interest with observations that, for many diseases, occurrence in non-adapted hosts can result in 'deadly' infections unlike the situation in natural hosts (e.g. ebola, hantavirus, arenavirus, viral encephalitis, tick fevers; see Bull, 1994). Virulence and/or pathogenicity may be defined as the potential for a pathogen to produce morbidity or mortality in infected hosts (Lipsitch & Moxon, 1997). In theoretical models of microparasite superinfection and transmission, virulence is often considered in terms of reproductive rate; strains of the parasite with the greater reproductive rate being more virulent and always superior in within-host competitive interactions (Bremermann & Pickering, 1983; Nowak & May, 1994). However, equating virulence with reproductive rate misses many features of the parasite/host relationship. From the few studies which have simultaneously measured parasite reproductive rate and disease severity there is little evidence of a positive relationship between the two (Ebert, 1994). There is, therefore, much interest in the ability of infectious agents to undergo rapid genetic variability, the role this plays in their survival

and transmission, and how this affects virulence. However, obtaining the information for constructing the necessary predictive models requires the use of molecular tools which can discriminate between different strains (Lipsitch & Moxon, 1997).

On a more practical level, conserved regions of DNA may have a predictive value for controlling parasites. Although the emergence of drug resistant strains of parasites is likely to be a continuing impediment to control, an alternative to the detection of drug resistant strains is the prediction of drug sensitivity using molecular techniques. Conserved drug targets like rRNA and tubulin structures predict sensitivities to benzimidazoles and aminoglycosides (Meloni *et al.* 1990; Armson *et al.* unpublished; Mclaughlin, 1997). Similarly, although on a more functional level, comparative molecular characterization of protease genes in *Leishmania* and other trypanosomatids, is seen as a prerequisite to the design of drugs which target these enzymes (Sakanari *et al.* 1997).

Species discrimination

One of the fundamental issues for researchers is to determine the taxonomic status of the group they are working with. In order for the information they learn about their organisms to be useful they need to delineate exactly what species are involved so that other researchers can be sure they are studying the same thing. At the most basic level therefore, molecular techniques can be developed to detect the presence of a species. The development of a diagnostic test(s) is the first step in any epidemiological study, and as we have seen in earlier chapters in this issue, the advantages of molecular assays, particularly those based on PCR, offer many advantages over conventional techniques including high sensitivity and the ability to detect infectious agents directly in clinical or environmental samples without the need for laboratory amplification. For example, with vector-borne infections such as those caused by *Onchocerca volvulus*, the success of intervention programmes requires a precise estimation of the annual transmission potentials of vector populations in endemic areas such as Uganda (Fischer *et al.* 1997). However, accurate surveillance may not be possible if the vectors are also capable of transmitting non-human species of *Onchocerca* with stages in the vector morphologically similar to those of *O. volvulus*. The recent development of sensitive PCR-based assays which can differentiate between human and animal filariae species (Fischer *et al.* 1997) will thus enable accurate surveillance of vector populations and an accurate assessment to be made of control interventions in endemic areas.

One major challenge, particularly in some groups, is exactly what constitutes a species. The most commonly used species concept, the 'biological

species concept' (BSC) has several drawbacks particularly for parasitic groups – interbreeding is often difficult to test directly as group may be geographically separated and laboratory crosses may not reflect natural occurrence; most importantly it is applicable only to sexually reproducing, outcrossing organisms (Lymbery, 1992). Recent studies, especially with bacteria, have shown that low levels of recombination occur in many species allowing the BSC to be extended to groups previously thought to be strictly clonal. There is in fact a continuum from panmixic (strictly sexual, outcrossing) to strictly clonal breeding systems and in some species this may vary geographically, between ecological niches or even over time. For example, malaria (caused by *P. falciparum*) in highly endemic areas is close to panmixia, whereas in lower prevalence areas there is less crossing (Walliker, 1994). An example of variation in breeding system between host species is found with *Trypanosoma*, where genetic exchange occurs between the cattle stocks, while an 'epidemic' population structure involving limited genetic exchange is a characteristic of the human infective stocks (Hide *et al.* 1994). The amount of recombination may also vary over time and it is important when measuring gene flow to realize that it may not be current migration but residual evidence of historical interactions, as suggested from studies of mtDNA for *Ascaris* (Anderson, 1995).

Although there is a large number of species concepts and guidelines for designating species, they share common attributes that, to be a species, a group must be cohesive (monophyletic) and distinct from other groups. A comprehensive taxonomic study of a group thus involves studying numerous individuals from the entire geographic range, and individuals from similar species looking for distinct clusters, without making assumptions about what constitutes a population as there may be invisible ecological subdivisions. Molecular data should never be examined in isolation from other types of information. The relationships determined for a single locus (or even entire extra-nuclear DNA) may not reflect the relationships between whole organisms. Furthermore, the amount of genetic variation required to designate a species should be considered in conjunction with biologically significant differences in traits especially as the actual level of genetic difference between species is highly variable across taxa.

With the cestode *Echinococcus granulosus*, morphological and developmental differences between the adult parasite of British horse and sheep origin have been shown to reflect considerable genetic variability. Initially, these forms of the parasite were described as distinct subspecies but were subsequently invalidated taxonomically because of their sympatric occurrence and other taxonomic arguments (Thompson & Lymbery, 1988). Genetic characterization has now revealed that the genetic differences between these forms are not only conserved and occur consistently in isolates from horses and sheep throughout the world, but are also greater than differences between currently recognized species (Thompson *et al.* 1994; Lymbery & Thompson, 1996). In view of these findings and data from a number of other host adapted variants of *Echinococcus*, a taxonomic revision of the genus *Echinococcus* has been proposed (Lymbery & Thompson, 1996; Thompson & Lymbery, 1996; Thompson, 1997).

In a number of parasite groups, molecular studies have revealed subdivisions not apparent using morphology which represent host-adapted 'strains' between which there is very little gene flow and which probably represent distinct species. This has important epidemiological consequences that can help to target control efforts most effectively as it means no effective transmission is occurring between host species. For example, the molecular characterization of *Ascaris* from humans and pigs in Guatemalan villages where both hosts occur in sympatry, has revealed that parasite populations occur in 2 host associated clusters (Anderson, Romero-Abal & Jaenike, 1993; Anderson & Jaenike, 1997). This suggests that in Guatemala, cross-infection is extremely rare and/or of very short duration which contrasts with the situation in North America where cross-infection has been shown to occur (Anderson, 1995). From a practical consideration, the results suggest that programmes aiming to control *Ascaris* infection in the human population can safely ignore zoonotic infection from pigs. The authors suggest that human and pig-associated *Ascaris* populations should continue to be regarded as sibling species both for convenience and because there are strong reproductive isolating mechanisms limiting gene flow between them (Anderson & Jaenike, 1997).

Similarly, with *Cryptosporidium*, the recent proposals to limit species in the genus to 6 (see O'Donoghue, 1995) or 8 (Fayer, Speer & Dubey, 1997) would seem to be an oversimplification in light of earlier taxonomic studies and recent data obtained from a comprehensive genetic characterization of isolates from different species of hosts (Morgan *et al.* 1997, 1998 *a, c*). A range of host-adapted, geographically conserved, genotypes have been identified which appears likely to reflect taxonomic discrimination as well as providing a valuable basis for the predictive epidemiology of cryptosporidial infections.

Intraspecific variants and host-adapted strains

Perhaps the most exciting applications of ME are those studying within species genetic variation. The amount of variation within and between populations

reveals the population genetic structure, the breeding system and amount of geneflow. During the 1980s, the challenge was to demonstrate that the phenotypic variability described for many parasites had a genetic basis. The application of isoenzyme electrophoresis, and subsequently RFLP analysis was used widely for this purpose. For example, research on the heterogeneity of the causative agent of Chagas' disease, *Trypanosoma cruzi*, demonstrates the important role molecular characterization techniques have played in elucidating the epidemiological and clinical significance of intraspecific variation. One of the basic problems in understanding the aetiology of Chagas' disease is that it can present in a variety of clinical forms. Because of this diversity and the significant geographical variation in the severity and clinical course of the different clinical sequelae, it had long been presumed that *T. cruzi* was composed of genetically distinct populations. Research over the last 15 years, initially using isoenzyme analysis and more recently DNA characterization techniques, has unequivocally demonstrated the central role the genetic constitution of *T. cruzi* plays in determining the clinical characteristics of Chagas' disease and the association between 'strain/genotype/clone' with transmission cycle and drug sensitivity (reviewed in Miles & Cibulskis, 1986; Thompson, 1988; Macedo & Pena, 1998).

Because of the extensive genetic and phenotypic variation that has been demonstrated in *Echinococcus*, it is very important to characterize the aetiological agents in different endemic areas in order to determine transmission patterns, particularly where there is the possibility of interaction between cycles (Thompson, 1995; Thompson *et al.* 1994; Thompson, Lymbery & Constantine, 1990). Unlike the situation in Britain, where ME has shown that 2 cycles of transmission occur involving 2 distinct strains, probably species, of *Echinococcus* with different intermediate host preferences and thus minimal opportunities for interaction (Thompson & Smyth, 1975; Thompson, 1991), there is no evidence of genetic distinctness between *E. granulosus* maintained in different cycles of transmission on the Australian mainland (Lymbery, Thompson & Hobbs, 1988; Thompson & Lymbery, 1990; Hope, Bowles & McManus, 1991). This is significant to the control of hydatid disease in Australia since wild and domestic cycles of transmission will interact in areas where they overlap (Thompson, 1992; Constantine *et al.* 1993).

ME investigations on *Echinococcus* in Australia have also found that gene flow is restricted between populations of the parasite on the mainland of Australia and in the island state of Tasmania, and that these populations represent distinct strains (Lymbery & Thompson, 1988; Thompson & Lymbery, 1988). Despite the genetic differences between mainland and Tasmanian populations, however,

Lymbery (1992) calculated that migration between the populations was of sufficient magnitude to be responsible for occasional breakdowns in the largely successful Tasmanian hydatid control campaign.

Protozoan parasites may exhibit considerable genetic variability making it difficult to predict transmission characteristics from one endemic area to another. This is well illustrated in *Giardia*, where the question of zoonotic transmission has been clouded by numerous laboratory-based investigations on potentially, culture-selected isolates. It was not until the development of appropriate PCR-based techniques which allowed direct genotypic characterization of *Giardia* in faecal samples collected from localized endemic areas, that it has been possible to study the frequency of zoonotic transmission. An interesting picture has emerged. Although humans and dogs may harbour identical genotypes, it appears that in areas where the frequency of transmission is high, competitive interactions between apparently 'host-adapted' genotypes may preclude establishment of other isolates of *Giardia* (Hopkins *et al.* 1997). In areas where the frequency of transmission is less intense, evidence suggests that infections resulting from zoonotic transmission are more likely to occur.

Outbreaks of cryptosporidiosis are frequently reported and it is important, particularly in the case of waterborne outbreaks, to determine the source of infection (Morgan & Thompson, 1998). Molecular characterization of isolates of *Cryptosporidium parvum* has identified 2 major genotypes; 1 common to domestic livestock, particularly cattle, and the other restricted to humans (Morgan *et al.* 1997, 1998*a*, *b*; Peng *et al.* 1997; Spano *et al.* 1998). Interestingly, the 'cattle' genotype is found in less than 20% of human cases (Morgan *et al.* 1998*b*). The application of PCR-based genotyping tools can now be applied to outbreak situations in order to determine the source of infection. In the notorious Milwaukee outbreak where over 400000 people became infected, a human source has recently been identified as responsible for infection (Peng *et al.* 1997). Other potential sources of environmental contamination include wild animals and recent molecular characterization has revealed distinct genotypes in some species of native fauna (Morgan *et al.* 1998*a*).

Conservation

Although parasitology is recognized as having far-reaching implications for the ecology and evolution of species, its importance in the area of conservation is also increasingly been emphasized (e.g. see Loye & Carroll, 1995). ME has much to offer in this field, particularly with the analysis of field samples in studies on species survival and the impact of pathogens affecting native fauna. Endangered species often decline to very low numbers with subsequent

loss of genetic diversity which can greatly increase the populations at risk of debilitating disease from pathogens that overcome the immune defences of a single individual (O'Brien, 1994). Exotic diseases often cause greater mortality in new host species than ones with which they have coevolved, thus the introduction of new diseases has the potential to cause extinction in already endangered species.

For example, feral cats adversely affect Australian wildlife in a number of ways (Cross, 1990) of which cat-related diseases such as toxoplasmosis are suspected of being a serious cause of mortality and may also be contributing to the decline of some endangered species (Cross, 1990; Obendorf, Statham & Driessen, 1996). Australian native animals are highly susceptible to infection with *Toxoplasma*, and usually suffer severely from the disease (Lindsay *et al.* 1995; Obendorf *et al.* 1996). However, the epidemiology of toxoplasmosis in Australian native fauna is poorly understood. The cat is essential for the maintenance of *Toxoplasma* in endemic areas and is believed to be the main source of infection for native animals in Australia (Hartley & Munday, 1974; Obendorf *et al.* 1996). Although *Toxoplasma* has been isolated from numerous species of wild animals no studies have tried to correlate the occurrence of *Toxoplasma* in native mammals and birds with that in feral cats occurring in the same geographical area. The major limiting factor in obtaining useful epidemiological data has been the inadequacy of available diagnostic techniques which are reliant on serology, which is often unreliable (James *et al.* 1996) and unable to detect latent infections. The recent advent of ME techniques based on PCR (James *et al.* 1996; Muller *et al.* 1996) will overcome these problems and offer greater sensitivity, specificity and rapidity for the direct detection of *Toxoplasma* in faeces, blood, as well as a variety of other tissues in which both the active or latent stages of *Toxoplasma* may be present.

Tracking transmission

As emphasized above, using appropriate DNA techniques allows different levels of genetic discrimination to be revealed. At the finest level, 'fingerprinting' approaches utilizing micro/mini-satellites or RFLP/ITS, may allow isolates/clonal lineages to be followed in order to better understand inter-and intra-familial transmission at a localized level in endemic communities, as well as in longitudinal studies. Such approaches can also be used to study the transmission dynamics of different parasite genotypes and to relate this information to symptomatology and response to treatment. For example, with *Giardia* it has been possible to 'type' isolates of the parasites recovered directly from the stools of chronically infected children over a two year period (see below), providing evidence of the occurrence of mixed infections, the response of particular genotypes to drugs and the frequency of inter-familial transmission (Hopkins *et al.* 1997, 1999).

Similarly, with studies on *Pneumocystis*, a fingerprinting approach utilizing the ITS region of nuclear rRNA suggested that both re-activation of a previously acquired infection and reinfection from an exogenous source may be occurring in patients with recurrent episodes of *P. carinii* pneumonia (Tsolaki *et al.* 1996). These authors also found evidence of case clustering suggesting the possibility of person-to-person spread.

Probably the most elegant ME transmission studies have been undertaken on the spread of *Mycobacterium tuberculosis* in endemic communities. For example, research by Genewein *et al.* (1993) involving the genetic typing of individual strains of *M. tuberculosis* demonstrated that networks of marginalized drug users are susceptible to each others' infections (including TB and HIV), and that these infections will spread to specific individuals in adjacent groups, such as restaurant workers.

ME will be of most value in endemic areas where rapid intervention strategies are likely to enhance control efforts. For example, in diarrhoeal outbreaks of cryptosporidiosis, the ability to 'type' the genotypes from affected individuals will determine whether the source of the outbreak is of livestock origin (i.e. run-off from pasture) or as a result of a failure in the human sanitation system (Morgan & Thompson, 1998). At a more localized level, the availability of sensitive molecular tools that can detect asymptomatic carriers of *Cryptosporidium* will be of immense value in limiting the risk of infection at the institutional level, such as in day-care centres.

Thus in general, ME interventions will provide information on which members of a community become infected with a particular 'strain' of infectious agent and by whom. For example, we will know what proportion of giardial or cryptosporidial infections in infants come from the parents and what proportion from siblings; and which social groups are mainly responsible for transmitting a particular strain of TB, in the population, and what kind of contact is high risk? All these answers will allow more rational planning of interventions.

Ecological interactions within the host

At a more fundamental level, ME investigations are concerned with our lack of understanding about within-host infection dynamics. Molecular techniques are being developed which will enable the kinetics of parasite infection to be studied in the host. For example, quantitative competitive PCR has been used to study parasite abundance in susceptible and resistant strains of mice infected with *T. gondii* (Luo *et a.* 1997). Increasing attention

has also been given to the role of genetic variation in the dynamics of host-parasite interactions, and the significance of such interactions with respect to evolutionary considerations and the clinical outcome of parasite infections. Only a few of these studies have considered the complications which arise from the co-infection of a single host with genetically different strains/genotypes of a parasite, and these have usually been restricted to the case of super-infection, where one strain usually out competes another within the host (Bonhoeffer & Nowak, 1994; May & Nowak, 1994; Nowak & May, 1994). Although such theoretical studies can help to answer basic questions about the evolution of virulence (Bonhoeffer & Nowak, 1994) and applied problems such as vaccination strategies (Gupta *et al.* 1994), there is at present only a very narrow empirical base for either testing their predictions or determining the validity of their assumptions (Thompson & Lymbery, 1996). We require better ecological and genetic data to address questions such as: How frequently do multiple infections of a single host occur? Is there a competitive dominance hierarchy among different parasite strains within a host? Are more virulent strains competitively superior? Can virulence be equated with reproductive rate? (Thompson & Lymbery, 1996). The application of ME to these problems is providing the tools for collecting such data (Hide & Tait, 1991; Bull, 1994; Thompson *et al.* 1994).

Mixed-genotype infections of microparasites are considered to be common, but almost nothing is known about how competitive interactions within hosts affect the subsequent transmission success of individual genotypes (Taylor, Walliker & Read, 1997). We also do not understand the mechanisms that maintain strain structure in pathogen populations (Gupta *et al.* 1996; Lipsitch, 1997). Macedo & Pena (1998) have recently put forward a complex picture of clonal interactions in infections with *Trypanosoma cruzi* that can only be resolved by molecular techniques. They propose that 'swarms' of *T. cruzi* clones may form symbiotic relationships in the same host but that they must also compete fiercely for available resources. One would expect that some clones would be eliminated by their inability to compete, although different clones could present tropism for different tissues thus obviating the need to compete for the same resources. *In vivo* isolation techniques are highly selective for one or more clones from the original population present in the patient's blood. Consequently, techniques dependent on laboratory amplification are not suitable. These authors have recently shown that it is possible to achieve genetic typing directly from clinical tissues using PCR which will provide a valuable tool to study the ME and pathogenesis of Chagas' disease (Macedo & Pena, 1998).

Elegant research reported by Ranford-Cartwright *et al.* (1997) has demonstrated the ability of PCR techniques to study mixed infections with different genotypes of *Plasmodium falciparum*. Of particular value was the ability to identify different parasite genotypes in pre- and post-treatment samples thus allowing recrudescent and/or drug resistant genotypes to be identified, as well as genotypes resulting from a new infection or as a result of sequestration. This approach offers a powerful tool in our understanding of the within-host interactions of *P. falciparum* genotypes, and the epidemiology of malaria. In future, it should be possible to type parasites at the loci responsible for the resistance although for the majority of antimalarial drugs in use, these have not been identified (Ranford-Cartwright *et al.* 1997).

Studies on the ME of *Giardia* infections in localized endemic areas where the frequency of transmission is high and extensive genetic heterogeneity exists within the parasite population, have shown that mixed infections with more than 1 genetic variant may occur in children (Hopkins *et al.* 1999). Longitudinal studies have also revealed that mixed genotype populations do not coexist over time and eventually one genotype predominates. Whether this is due to variable sensitivity to drug interventions or competition is not clear. However, experimental studies have demonstrated that competitive interactions do occur *in vitro* between genetically distinct isolates of *G. duodenalis* with faster-growing isolates out-competing those with slower growth rates (Thompson *et al.* 1996). The addition of sub-lethal concentrations of metronidazole to clonal mixtures *in vitro* prevented the competitive exclusion which was seen in normal culture. This apparently occurred because the drug reduced the growth rate, and presumably also the maximum stationary phase concentration, of the faster-growing but not the slower-growing clone (Thompson *et al.* 1996). In asexually reproducing organisms, competition between clones for a restricted resource is expected to lead to monopolisation by the competitively superior clone, and therefore a reduction in clonal diversity in the absence of disturbing forces (Sebens & Thorne, 1985; Wilson & Hebert, 1992; Weider, 1992). We have argued previously that interference with the normal homogenizing influence of clonal competition may account for the greater than expected levels of genetic diversity found in populations of *Giardia* from remote Aboriginal communities in Australia (Thompson & Meloni, 1993). In these endemic communities, the disturbing force may be frequent exposure of the parasite to regular, suboptimal chemotherapeutic interventions which may interfere with the normal homogenizing influence of clonal competition in *Giardia* and thus serve to maintain genetic diversity through pre-emption of competitive exclusion (Thompson, 1991; Thompson & Meloni, 1993; Thompson *et al.* 1996).

Genetic markers

Although phenotypic variants have been shown to differ genetically, the big challenge which remains is to correlate phenotypic and genetic differences in the context of transmission and markers of epidemiological/clinical significance (e.g. for drug resistance and virulence). The relationship between phenotype characters such as virulence or drug-resistance and genotype is important in many ME studies. The ability to find genetic markers for phenotypic traits depends greatly on the population genetic structure which is determined by factors such as amount of recombination and geographic or ecological isolation.

At a more practical level, there is a role for ME in discriminating between well defined virulence phenotypes such as those associated with congenital abnormalities caused by toxoplasmosis. Genetic markers based on microsatellites in the beta-tubulin gene of *Toxoplasma gondii* have potential for direct screening of biological samples. Since they are PCR-based they can be used for typing *T. gondii* isolates without any bias due to laboratory cultivation of the parasite (Costa *et al*. 1997). This area of investigation is likely to be a huge growth area in ME, although few studies have been carried out to date. The combination of information from genome projects, such as the *P. falciparum* genome project (Dame *et al*. 1996), the large increase in available markers (due to advances in speed and reduction in costs), and new techniques such as RDA will assist in the detection of loci responsible for phenotypic differences right down to specific nucleotide changes. This will lead directly to developing practical control measures, particularly new prophylactics and therapeutics.

However, as infectious agents are subject to evolutionary adaptation, any intervention, drug or vaccine, imposes a selective advantage on variants capable of circumventing that intervention (Ottesen *et al*. 1998). The inherent genetic heterogeneity of infectious agents, and their ability to produce antigenic and drug resistant variants may thus limit our changes of effecting sustained control. Genetic variability may also decrease the effectiveness of diagnostic tools. There is, therefore, a need to apply appropriate ME techniques for the surveillance of parasite populations for the rapid detection of new variants. If not, as new vaccines and therapeutics are developed and used against variant populations of infectious agents, such as malaria, we may once again face the prospect of pathogens evolving to outrun our control measures (Lipsitch, 1997).

Analysis

The resulting explosion of molecular data has been followed, after a significant lag period, with new methods of analysis. It is also causing a constant review of theories concerned with processes of evolution and speciation. All analyses make certain assumptions, usually to simplify problems enough so that we can express them mathematically, such as: that populations are in a state of equilibrium, that is, all forces acting on them currently have been acting at similar levels for a very long time. Such idealized situations are probably never completely met. It is important to be aware that assumptions are made and if these assumptions are seriously violated the result may be invalid. New methods such as coalescence analysis allow us to recover information about the shared history of genes and directly estimate effective population size, migration rates, recombination rates, population growth and selection (Beerli, 1997).

Simply 'plugging' data into any available programme and accepting the result without question is not acceptable. The production of dendograms is often the extent of analysis for many published studies, despite the fact that any random data will produce a dendogram. Recent research is revealing that different clustering techniques and new approaches to handling the data can have extreme effects on the resultant dendogram (Nadler, 1990; Schoniger & Vonhaeseler, 1995). Techniques are currently being developed to attempt to quantify the 'power' of a resultant dendogram (Crowley, 1992; Penny, Hendy & Steel, 1992; Crandall, 1994). General guidelines include using more than 1 locus, comparing multiple data sets, contrasting different analysis methods, and bootstrapping (or other re-sampling techniques) to determine the consistency of the data. Low sample sizes and low levels of variability greatly increase the chance of Type II errors (accepting the null hypothesis incorrectly). Isolates identical by 1 method may be further subdivided by a more discriminating method; for example the ITS 2 sequences of *Metastrongylus salmi* and *M. confusus* were found to be identical, whereas RAPD analysis, morphological and epidemiological data indicated unambiguously that the 2 species are distinct (Leignel, Humbert & Elard, 1997).

There is a lack of communication between the various fields, they use different terminology, utilize different techniques and widely different frameworks. Although it is true that generalizing across all organisms is very difficult, the development of a common framework is important. In this respect, evolutionary biology and population biology have much to offer to infectious disease research in terms of techniques and cross-discipline collaborative studies (e.g. phylogenies and population-dynamic modelling) and testable, sometimes counterintuitive, hypotheses about the mechanisms and transmission of disease (Lipsitch, 1997). The provision of quantifiable molecular data will greatly enhance the predictive power of such multidisciplinary investigations.

As we have tried to demonstrate, ME may involve the application of techniques to distinguish closely related infectious agents or to document patterns of transmission of 'strains' and species within populations. ME provides a predictive and quantifiable power to epidemiological investigations not previously available. This is allowing precise determinations to be made about the aetiological agent, its characteristics and the source of infection. Phenotypic markers will hopefully soon be available so that more accurate predictions can be made about the clinical course of any resultant disease and the most appropriate therapy. It would appear that the future of infectious disease epidemiology may be largely dependent on DNA approaches. However, as we have tried to emphasize, developments in DNA characterization must be combined with traditional methods and matched by appropriate methods of analysis and interpretation.

Ideally, ME techniques should also allow determinations of why infections occur and if they are likely to occur in the future. Techniques that utilize virulence markers for parasites and other microorganisms and susceptibility markers of human populations are also needed to provide such determinations; for example, ligands and receptors that mediate the attachment of infectious agents to the surface of host cells are being described at the molecular level (McDade & Anderson, 1996). Similarly, processes associated with the invasion of tissues by infectious agents and dissemination within the host are also being investigated. Characterization of the molecules and genes involved may provide the basis for development of newer techniques. For example, recent studies have shown a significant association of recurrent candidal (*Candida albicans*) vaginitis with a particular blood group antigen, suggesting that certain variants of the parasite preferentially or exclusively infect these women (Hilton *et al*. 1995). It has therefore been suggested that the composition of carbohydrates comprising the blood group antigen molecules determine which strains will attach and invade a particular person (McDade & Anderson, 1996).

REFERENCES

ANDERSON, T. J. C. (1995). *Ascaris* infections in humans from North America: molecular evidence for cross-infection. *Parasitology* **110**, 215–219.

ANDERSON, T. J. C. & JAENIKE, J. (1997). Host specificity, evolutionary relationships and macrogeographic differentiation among *Ascaris* populations from humans and pigs. *Parasitology* **115**, 325–342.

ANDERSON, T. J. C., ROMERO-ABAL, M. E. & JAENIKE, J. (1993). Genetic structure and epidemiology of *Ascaris* populations: patterns of host affiliation in Guatemala. *Parasitology* **107**, 319–334.

APPELS, R. & HONEYCUTT, R. L. (1986). rDNA: Evolution over a billion years. In *DNA Systematics* (ed. Dutta, S. K.). p. 81–135. Boca Raton, Florida: CRC Press.

BANDI, C., LA ROSA, G., BARDIN , M. G., DAMIANI, G., DE CARNERI, I. & POZIO, E. (1993). Arbitrarily primed polymerase chain reaction of individual *Trichinella* specimens. *Journal of Parasitology* **79**, 437–440.

BARLOW, J. W. & EWEN, K. R. (1998). High-throughput genotyping. *Today's Life Science* **10**, 16–21.

BEERLI, P. (1997). Statistical analyses of population genetic data. *TREE* **12**, 488.

BJÖRKMAN, A., DO RASÁRIO, V. E., SNOUNOU, G. & WALLIKER, D. (1998). Standardizing PCR for molecular epidemiology studies of malaria. *Parasitology Today* **14**, 85.

BLAXTER, M. L., DE LEY, P., GAREY, J. R., LIU, L. X., SCHELDEMAN, P., VIERSTRAETE, A., VANFLETEREN, J. R., MACKEY, L. Y., DORRIS, M., FRISSE, L. M., VIDA, J. T. & THOMAS, K. W. (1998). A molecular evolutionary framework for the phylum Nematoda. *Nature* **392**, 71–75.

BONHOEFFER, S. & NOWAK, M. A. (1994). Mutation and the evolution of virulence. *Proceedings of the Royal Society of London* **258**, 133–140.

BREMERMANN, H. J. & PICKERING, J. (1983). A game-theoretical model of parasite virulence. *Journal of Theoretical Biology* **100**, 411–426.

BULL, J. J. (1994). Evolution of virulence. *Evolution* **48**, 1423–1437.

BUSH, A. O., CAIRA, J. N., MINCHELLA, D. J., NADLER, S. A. & SEED, J. R. (1995). Parasitology year 2000. *Journal of Parasitology* **81**, 835–842.

BUTH, D. G. (1984). The application of electrophoretic data in systematic studies. *Annual Reviews of Ecology and Systematics* **15**, 501–522.

CONSTANTINE, C. C., THOMPSON, R. C. A., JENKINS, D. J., HOBBS, R. P. & LYMBERY, A. J. (1993). Morphological characterization of adult *Echinococcus granulosus* as a means of determining transmission patterns. *Journal of Parasitology* **79**, 57–61.

COSTA, J. M., DARDE, M. L., ASSOULINE, B., VIDAUD, M. & BRETAGNE, S. (1997). Microsatellite in the beta-tubulin gene of *Toxoplasma gondii* as a new genetic marker for use in direct screening of amniotic fluids. *Journal of Clinical Microbiology* **35**, 2542–2545.

CRANDALL, K. A. (1994). Intraspecific cladogram estimation: Accuracy at higher levels of divergence. *Systematic Biology* **43**, 222–235.

CROSS, J. (1990). The feral cat – justifications for its control. Thesis for Associate Diploma in Applied Science. Charles Sturt University, Riverina, Australia.

CROWLEY, P. H. (1992). Resampling methods for computation-intensive data analysis in ecology and evolution. *Annual Reviews of Ecology and Systematics* **23**, 405–447.

DAME, J. B., ARNOT, D. E., BOURKE, P. F., CHAKRABARTI, D., CHRISTODOULOU, Z., COPPEL, R. L., COWMAN, A. F., CRAIG, A. G., FISCHER, K., FOSTER, J., GOODMAN, N., HINTERBERG, K., HOLDER, A. A., HOLT, D. C., KEMP, D. J., LANZER, M., LIM, A., NEWBOLD, C. I., RAVETCH, J. V., REDDY, G. R., RUBIO, J., SCHUSTER, S. M., SU, X. Z., THOMPSON, J. K., VITAL, F. *et al*. (1996). Current status of the *Plasmodium falciparum* genome project. *Molecular and Biochemical Parasitology* **79**, 1–12.

DONELSON, J. E. (1996). Genome research and evolution in Trypanosomes. *Evolution* **6**, 699–703.

EBERT, D. (1994). Genetic differences in the interactions of a microsporidian parasite and four clones of its cyclically parthenogenetic host. *Parasitology* **108**, 11–16.

FAYER, R., SPEER, C. A. & DUBEY, J. P. (1997). The general biology of *Cryptosporidium*. In *Cryptosporidium and Cryptosporidiosis* (ed. Fayer, R.), pp. 1–42. CRC Press, Boca Raton.

FELLEISEN, R. S. J. (1998). Comparative genetic analysis of Tritrichomonadid protozoa by the Random Amplified Polymorphic DNA technique. *Parasitology Research* **84**, 153–156.

FISHER, P., YOCHA, J., RUBAALE, T. & GARMS, R. (1997). PCR and DNA hybridization indicate the absence of animal filariae from vectors of *Onchocerca volvulus* in Uganda. *Journal of Parasitology* **83**, 1030–1034.

GENEWEIN, A., TELENTI, A., BERNASCONI, C., MORDASINI, C., WEISS, S., MAURER, A., RIEDER, H. L., SCHOPFER, K. & BODMER, T. (1993). Molecular approach to identifying route of transmission of tuberculosis in the community. *The Lancet* **342**, 841–844.

GUO, Z. G. & JOHNSON, A. M. (1995*a*). Genetic characterization of *Toxoplasma gondii* strains by random amplified polymorphic DNA polymerase chain reaction. *Parasitology* **111**, 127–132.

GUO, Z. G. & JOHNSON, A. M. (1995*b*). Genetic comparison of *Neospora caninum* with *Toxoplasma* and *Sarcocystis* by Random Amplified Polymorphic – DNA-polymerase chain reaction. *Parasitology Research* **81**, 365–370.

GUPTA, S., MAIDEN, M. C. J., FEAVERS, I. M., NEED, S., MAY, R. M. & ANDERSON, R. M. (1996). The maintenance of strain structure in populations of recombining infectious agents. *Nature Medicine* **2**, 437–442.

GUPTA, S., TRENHOLME, K., ANDERSON, R., M. & DAY, K. P. (1994). Antigenic diversity and the transmission dynamics of *Plasmodium falciparum*. *Science* **263**, 961–963.

HALL, A. (1996). Editorial: What is molecular epidemiology? *Tropical Medicine and International Health* **1**, 407–408.

HARRISON, R. G. (1989). ANIMAL MITOCHONDRIAL DNA AS A GENETIC MARKER IN POPULATION AND EVOLUTIONARY BIOLOGY. *TREE* **4**, 6–11.

HARTLEY, W. J. & MUNDAY, B. L. (1974). Felidae in the dissemination of toxoplasmosis to man and other animals. *Australian Veterinary Journal* **50**, 224–228.

HARVEY, P. H. & NEE, S. (1994). Phylogenetic epidemiology lives. *TREE* **9**, 361–363.

HASEGAWA, M., IIDA, Y., YANO, T., TAKAIWA, F. & IWABUCHI, M. (1985). Phylogenetic relationships among eukaryotic kingdoms inferred from ribosomal RNA sequences. *Journal of Molecular Evolution* **22**, 32–38.

HIDE, G. & TAIT, A. (1991). The molecular epidemiology of parasites. *Experientia* **47**, 128–142.

HIDE, G., WELBURN, S. C., TAIT, A. & MAUDLIN, I. (1994). Epidemiological relationships of *Trypanosoma brucei* stocks from South East Uganda: evidence for different population structures in human infective and non-human infective isolates. *Parasitology* **109**, 95–111.

HILLIS, D. M. & DAVIS, S. K. (1986). Evolution of ribosomal DNA: Fifty million years of recorded history in the frog genus *Rana*. *Evolution* **40**, 1275–1288.

HILTON, E., CHANDRASEKARAN, V., RINDOS, P. ISENBERG, H. D. (1995). Association of recurrent candidal vaginitis with inheritance of Lewis blood group antigens. *Journal of Infectious Diseases* **172**, 1616–1619.

HOPE, M., BOWLES, J. & McMANUS, D. P. (1991). A reconsideration of the *Echinococcus granulosus* strain situation in Australia following RFLP analysis of cystic material. *International Journal for Parasitology* **21**, 471–475.

HOPKINS, R. M., MELONI, B. P., GROTH, D. M., WETHERALL, J. D., REYNOLDSON, J. A. & THOMPSON, R. C. A. (1997). Ribosomal RNA sequencing reveals differences between the genotypes of *Giardia* isolates recovered from humans and dogs living in the same locality. *Journal of Parasitology* **83**, 44–51.

HOPKINS, R. M., CONSTANTINE, C. C., GROTH, D. A., WETHERALL, J. D., REYNOLDSON, J. A. & THOMPSON, R. C. A. (1999). PCR-based DNA fingerprinting of *Giardia duodenalis* isolates using the intergenic rDNA spacer. *Parasitology* (in press).

JAMES, G. S., SINTCHENKO, V. G., DICESON, D. J. & GILBERT, G. L. (1996). Comparison of cell culture, mouse inoculation, and PCR for the detection of *Toxoplasma gondii*. *Journal of Clinical Microbiology* **34**, 1572–1575.

JEFFREYS, A. J., WILSON, V. & THEIN, S. L. (1985). Hypervariable 'minisatellite' regions in human DNA. *Nature* **314**, 67–73.

KOELEMAN, J. G. M., PARLEVLIET, G. A., DIKJSHOORN, L., SAVELKOUL, P. H. M. & VANDENBROUCKEGRAULS, C. M. J. E. (1997). Nosocomial outbreak of multi-resistant *Acinetobacer baumannii* on a surgical ward – epidemiological and risk factors for acquisition. *Journal of Hospital Infection* **37**, 113–123.

KWOK, S. & HIGUCHI, R. (1989). Avoiding false positives with PCR. *Nature* **339**, 237–238.

LEIGNEL, V., HUMBERT, J. G. & ELARD, L. (1997). Study by ribosomal DNA ITS 2 sequencing and RAPD analysis on the systematics of four *Metastrongylus* species (Nematoda: Metastronyloidea). *Journal of Parasitology* **83**, 606–611.

LIN, J. J., KUO, J. & MA, J. (1996). A PCR-based DNA fingerprinting technique – AFLP for molecular typing of bacteria. *Nucleic Acids Research* **24**, 3649–3650.

LINDSAY, D. S., GASSER, R. B., HARRIGAN, K. E., MADILL, D. N. & BLAGBURN, B. L. (1995). Central nervous system toxoplasmosis in roller canaries. *Avian Diseases* **39**, 204–207.

LIPSITCH, M. (1997). Evolution in health and disease. *Trends in Microbiology* **5**, 303–305.

LIPSITCH, M. & MOXON, R. E. (1997). Virulence and transmissibility of pathogens: what is the relationship? *Trends in Microbiology* **5**, 31–37.

LISITSYN, N., LISITSYN, N. & WIGLER, M. (1993). Cloning the differences between two complex genomes. *Science* **259**, 946–951.

LISITSYN, N. A., SEGRE, J. A., KUSUMI, K., LISITSYN, N. M., NADEAU, J. H., FRANKEL, W. N., WIGLER, M. H. & LANDER, E. S. (1994). Direct isolation of polymorphic markers linked to a trait by genetically directed

representational difference analysis. *Nature Genetics* **6**, 57–63.

LOYE, J. & CARROLL, S. (1995). Birds, bugs and blood: avian parasitism and conservation. *TREE* **10**, 232–235.

LUO, W., AOSAI, F., UEDA, M., YAMASHITA, K., SHIMIZU, K., SEKIYA, S. & YANO, A. (1997). Kinetics in parasite abundance in susceptible and resistant mice infected with an avirulent strain of *Toxoplasma gondii* by using quantitative competitive PCR. *Journal of Parasitology* **83**, 1070–1074.

LYMBERY, A. J. (1992). Interbreeding, monophyly and the genetic yardstick: species concepts in parasites. *Parasitology Today* **8**, 208–211.

LYMBERY, A. J. & THOMPSON, R. C. A. (1988). Electrophoretic analysis of genetic variation in *Echinococcus granulosus* from domestic hosts in Australia. *International Journal for Parasitology* **18**, 803–811.

LYMBERY, A. J. & THOMPSON, R. C. A. (1996). Species in the genus *Echinococcus*: pattern and process. *Parasitology Today* **12**, 486–491.

LYMBERY, A. J., THOMPSON, R. C. A. & HOBBS, R. P. (1988). Genetic diversity and genetic differentiation in *Echinococcus granulosus* (Batch, 1786) from domestic and sylvatic hosts on the mainland of Australia. *Parasitology* **101**, 283–289.

MACEDO, A. M. & PENA, S. D. J. (1998). Genetic variability of *Trypanosoma cruzi*; implications for the pathogenesis of Chagas disease. *Parasitology Today* **14**, 119–124.

MACPHERSON, J. M. & GAJADHAR, A. A. (1993). Differentiation of seven *Eimeria* species by random amplified polymorphic DNA. *Veterinary Parasitology* **45**, 257–266.

MAJIWA, P. A. O., MAINA, M., WAITUMBI, J. N., MIHOK, S. & ZWEYGARTH, E. (1993). *Trypanosoma* (*Nannomonas*) *congolense*: molecular characterization of a new genotype from Tsavo, Kenya. *Parasitology* **106**, 151–162.

MAY, R. M. & NOWAK, M. A. (1994). Superinfection, metapopulation dynamics and the evolution of diversity. *Journal of Theoretical Biology* **170**, 95–114.

MCDADE, J. E. & ANDERSON, B. E. (1996). Molecular epidemiology: Applications of nucleic acid amplification and sequence analysis. *Epidemiologic Reviews* **18**, 90–97.

MCLAUGHLIN, G. L. (1997). Taxonomy's role in biomedical research on pathogenic protozoa. *Proceedings of the American Society of Parasitology Annual Meeting*, p.46#28, Nashville, Tennessee.

MCMICHAEL, A. J. (1995). The author replies. *American Journal of Epidemiology* **142**, 225.

MELONI, B. P., THOMPSON, R. C. A., REYNOLDSON, J. A. & SEVILLE, P. (1990). Albendazole: a more effective antigiardial agent *in vitro* than metronidazole or tinidazole. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **84**, 375–379.

MILES, M. A. & CIBULSKIS, R. E. (1986). Zymodeme characterization of *Trypanosoma cruzi*. *Parasitology Today* **2**, 94–97.

MORGAN, U. M., CONSTANTINE, C. C., GREENE, W. K. & THOMPSON, R. C. A. (1993). RAPD (Random Amplified Polymorphic DNA) analysis of *Giardia* DNA and correlation with isoenzyme data. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **87**, 702–705.

MORGAN, U. M., CONSTANTINE, C. C., O'DONOGHUE, P., MELONI, B. P., O'BRIEN, P. A. & THOMPSON, R. C. A. (1995). Molecular characterization of *Cryptosporidium* isolates from humans and other animals using RAPD (Random Amplified Polymorphic DNA) analysis. *American Journal of Tropical Medicine and Hygiene* **52**, 559–564.

MORGAN, U. M., CONSTANTINE, C. C., FORBES, D. A. & THOMPSON, R. C. A. (1997). Differentiation between human and animal isolates of *Cryptosporidium parvum* using rDNA sequencing and direct PCR analysis. *Journal of Parasitology* **83**, 825–830.

MORGAN, U. M., FORBES, D. A. & THOMPSON, R. C. A. (1998*c*). Molecular epidemiology of *Cryptosporidium parvum* infection. *European Journal of Protistology* **34**, 262–266.

MORGAN, U. M., PALLANT, L., DWYER, B., FORBES, D. A. & THOMPSON, R. C. A. (1998*b*). Comparison of PCR versus microscopy for the detection of *Cryptosporidium* – A clinical trial. *Journal of Clinical Microbiology* (in press).

MORGAN, U. M., SARGENT, K. D., DEPLAZES, P., FORBES, D. A., SPANO, F., HERTZBERG, H., ELLIOT, A. & THOMPSON, R. C. A. (1998*a*). Molecular characterisation of *Cryptosporidium* from various hosts. *Parasitology* **117**, 31–37.

MORGAN, U. M. & THOMPSON, R. C. A. (1998). The detection and characterisation of *Cryptosporidium* – PCR the way forward? *Parasitology Today* (in press).

MORITZ, C., DOWLING, T. E. & BROWN, W. M. (1987). Evolution of animal mitochondrial DNA: relevance for population biology and systematics. *Annual Reviews of Ecology and Systematics* **18**, 269–292.

MULLER, N., ZIMMERMANN, V., HENTRICH, B. & GOTTSTEIN, B. (1996). Diagnosis of *Neospora caninum* and *Toxoplasma gondii* by PCR and DNA hybridization immunoassay. *Journal of Clinical Microbiology* **34**, 2850–2852.

NADLER, S. (1990). Molecular approaches to studying helminth population genetics and phylogeny. *International Journal for Parasitology* **20**, 11–29.

NAKAMURA, Y., LEPPERT, M., O'CONNELL, P., WOLFF, R., HOLM, T., CULVER, M., MARTIN, C., FUMIMOTO, E., HOFF, M. & KUMLIN, E. (1987). Variable number of tandem repeat (VNTR) markers for human gene mapping. *Science* **235**, 98–102.

NOWAK, M. A. & MAY, R. M. (1994). Superinfection and the evolution of parasite virulence. *Proceedings of the Royal Society of London* **255**, 81–89.

O'BRIEN, S. J. (1994). A role for molecular genetics in biological conservation. *Proceedings of the National Academy of Sciences, USA* **94**, 5748–5755.

O'DONOGHUE, P. J. (1995). *Cryptosporidium* and cryptosporidiosis in man and animals. *International Journal for Parasitology* **25**, 139–195.

OBENDORF, D. L., STATHAM, P. & DRIESSEN, M. (1996). Detection of agglutinating antibodies to *Toxoplasma gondii* in sera from free-ranging eastern barred bandicoots (*Perameles gunnii*). *Journal of Wildlife Diseases* **32**, 623–626.

OLSEN, G. J. & WOESE, C. R. (1993). Ribosomal RNA: a key to phylogeny. *FASEB Journal* **7**, 113–123.

OTTESEN, E. A., DOWDLE, W. R., FENNER, F., HABERMEHL, K. O., JOHN, T. J., KOCH, M. A., MEDLEY, G. F., MULLER, A. S., OSTROFF, S. M. & ZEICHHARDT, H. (1998). Group report: how is eradication to be defined and what are the biological criteria? In *The Eradication of Infectious Diseases* (ed. Dowdle, W. R. & Hopkins, D. R.), pp. 47–59. New York, Wiley.

PENG, M. M., XIAO, L., FREEMAN, A. R., ARROWOOD, M. J., ESCALANTE, A. A., WELTMAN, A. C., ONG, C. S. L., MACKENZIE, W. R., LAL, A. A. & BEARD, C. B. (1997). Genetic polymorphism among *Cryptosporidium parvum* isolates: evidence of two distinct human transmission cycles. *Emerging Infectious Diseases* **3**, 567–573.

PENNY, D., HENDY, M. D. & STEEL, M. A. (1992). Progress with methods for constructing evolutionary trees. *Trends in Evolution and Ecology* **7**, 73–79.

POIRIER, M. C. (1995). Re: "Invited commentary – 'molecular epidemiology': new pathway or new travelling companion?" *American Journal of Epidemiology* **142**, 223–224.

POWELL, W., MORGANTE, M., ANDRE, C., HANAFEY, M., VOGEL, J., TINGEY, S. & RAFALSKI, A. (1996). The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Molecular Breeding* **2**, 225–238.

PROCUNIER, J. D., FERNANDO, M. A. & BARTA, J. R. (1993). Species and strain differentiation of *Eimeria* spp. of the domestic fowl using DNA polymorphisms amplified by arbitrary primers. *Parasitology Research* **79**, 98–102.

RANFORD-CARTWRIGHT, L. C., TAYLOR, J., UMASUNTHAR, T., TAYLOR, L. H., BABIKER, H. A., LELL, B., SCHMIDT-OTT, J. R., LEHMAN, L. G., WALLIKER, D. & KREMSNER, P. G. (1997). Molecular analysis of recrudescent parasites in a *Plasmodium falciparum* drug efficacy trial in Gabon. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **91**, 719–724.

RANZANI, G. N., BERNINI, L. F. & CRIPPA, M. (1984). Inheritance of rDNA spacer length variants in man. *Molecular and General Genetics* **196**, 141–145.

REINEKE, A., KARLOVSKY, P. & ZEBITZ, C. P. W. (1998). Preparation and purification of DNA from insects for AFLP analysis. *Insect Molecular Biology* **7**, 95–99.

SAKANARI, J. A., NADLER, S. A., CHAN, V. J., ENGEL, J. C., LEPTAK, C. & BOUVIER, J. (1997). *Leishmania major* – comparison of the cathepsin L- and B-like cysteine protease genes with those of other trypanosomatids. *Experimental Parasitology* **85**, 63–76.

SCHONIGER, M. & VONHAESELER, A. (1995). Performance of the maximum likelihood, Neighbor Joining, and Maximum Parsimony methods when sequence sites are not independent. *Systematic Biology* **44**, 533–547.

SEBENS, K. P. & THORNE, B. L. (1985). Coexistence of clones, clonal diversity, and the effects of disturbance. In *Population Biology and Evolution of Clonal Organisms* (ed. Jackson, J. B. C., Buss, L. W. & Cook, R. E.), pp. 357–98. Yale University Press, New Haven.

SITES, J. W. & DAVIS, S. K. (1989). Phylogenetic relationships and molecular variability within and among six chromosome races of *Sceloporus grammicus* (Sauria, Iguandiae), based on nuclear and mitochondrial markers. *Evolution* **43**, 296–317.

SPANO, F., PUTIGNANI, L., CRISANTI, A., SALLICANDRO, P., MORGAN, U. M., LE BLANCQ, S. M., TCHACK, L., TZIPORI, S. & WIDMER, G. (1998). Multilocus genotypic analysis of *Cryptosporidium parvum* isolates from different hosts and geographical origin. *Journal of Clinical Microbiology* **36**, 3255–3259.

STEINDEL, M., NETO, E. D., DE MENEZES, C. L. P., ROMANHA, A. & SIMPSON, A. J. G. (1993). Random amplified polymorphic DNA analysis of *Trypanosoma cruzi* strains. *Molecular and Biochemical Parasitology* **60**, 71–80.

TAYLOR, L. H., WALLIKER, D. & READ, A. F. (1997). Mixed-genotype infections of malaria parasites: within-host dynamics and transmission success of competing clones. *Proceedings of the Royal Society London B Biological Sciences* **264**, 927–935.

TENTER, A. M. & JOHNSON, A. M. (1997). Phylogeny of the tissue cyst-forming coccidia. *Advances in Parasitology* **39**, 69–139.

THOMPSON, R. C. A. (1988). Intraspecific variation and Epidemiology. In *Parasitology in Focus* (ed. Mehlhorn, H.), pp. 391–411. Springer, Berlin.

THOMPSON, R. C. A. (1991). *Echinococcus* and *Giardia*: variation on a theme. *International Journal for Parasitology* **21**, 291–297.

THOMPSON, R. C. A. (1992). Parasitic zoonoses: problems created by people, not animals. *International Journal for Parasitology* **22**, 555–562.

THOMPSON, R. C. A. (1995). Biology and systematics of *Echinococcus*. In *Echinococcus and Hydatid Disease* (ed. by Thompson, R. C. A. & Lymbery, A. J.) pp. 1–50, Wallingford, CAB International.

THOMPSON, R. C. A. (1997). Taxonomy of *Echinococcus* spp. *Archivos Internacionales de la Hidatidosis* **32**, 121–125.

THOMPSON, R. C. A. & LYMBERY, A. J. (1988). The nature, extend and significance of variation within the genus *Echinococcus*. *Advances in Parasitology* **27**, 210–258.

THOMPSON, R. C. A. & LYMBERY, A. J. (1990). Intraspecific variation in parasites – what is a strain? *Parasitology Today* **6**, 345–348.

THOMPSON, R. C. A. & LYMBERY, A. J. (1996). Genetic variability in parasites and host-parasite interactions. *Parasitology* **112**, 57–522.

THOMPSON, R. C. A., LYMBERY, A. J. & CONSTANTINE, C. C. (1990). Variation in *Echinococcus*: towards a taxonomic revision of the genus. *Advances in Parasitology* **35**, 145–176.

THOMPSON, R. C. A., LYMBERY, A. J., MELONI, B. P., MORGAN, U. M., BINZ, N., CONSTANTINE, C. C. & HOPKINS, R. M. (1994). Molecular epidemiology of parasite infections. In *Biology of Parasitism* (ed. Ehrlich, R. & Nieto, A.) pp. 167–185. Montevideo, Uruguay, Edicio Trilce.

THOMPSON, R. C. A., LYMBERY, A. J., PEARCE, D. A., FINN, K. C., REYNOLDSON, J. A. & MELONI, B. P. (1996). *Giardia duodenalis*: exposure to metronidazole inhibits competitive interactions between isolates of the parasite *in vitro*. *Journal of Parasitology* **82**, 679–683.

THOMPSON, R. C. A. & MELONI, B. P. (1993). Molecular variation in *Giardia* and its implications. *Acta Tropica* **53**, 167–184.

THOMPSON, R. C. A. & SMYTH, J. D. (1975). Equine hydatidosis: a review of the current status in Great Britain and the results of an epidemiological survey. *Veterinary Parasitology* **1**, 107–127.

TIGHE, P. J., GOYAL, P. K., WILSON, Z. A., WAKELIN, D. & PRICHARD, D. I. (1994). Analysis of genetic variation in isolates of *Trichinella* using random amplified polymorphic DNA. *Molecular and Biochemical Parasitology* **63**, 175–178.

TSOLAKI, A. G., MILLER, R. F., UNDERWOOD, A. P., BANERJI, S. & WAKEFIELD, A. E. (1996). Genetic diversity at the internal transcribed spacer regions of the rRNA operon among isolates of *Pneumocystis carinii* from AIDS patients with recurrent pneumonia. *Journal of Infectious Diseases* **174**, 141–156.

VOS, P., HOGERS, R., BLEEKER, M., REIJANS, M., VANDELEE, T., HORNES, M., FRIJTERS, A., POT, J., PELEMAN, J., KUIPER, M. & ZABEAU, M. (1995). AFLP – A new technique for DNA fingerprinting. *Nucleic Acids Research* **23**, 4407–4414.

WALLIKER, D. (1994). The role of molecular genetics in field studies on malaria parasites. *International Journal for Parasitology* **24**, 799–808.

WEBER, J. L. & MAY, P. E. (1989). Abundant class of human DNA polymorphisms which can be typed using PCR. *American Journal of Human Genetics* **44**, 388–396.

WEIDER, L. J. (1992). Disturbance, competition and the maintenance of clonal diversity in *Daphnia pulex*. *Journal of Evolutionary Biology* **5**, 505–22.

WELSH, J. & MACCLELLAND, M. (1990). Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Research* **18**, 7213–7218.

WESTON, A. (1995). Re: ''Invited commentary – 'molecular epidemiology': new pathway or new travelling companion?'' *American Journal of Epidemiology* **142**, 224–225.

WILLIAMS, J. G. K., KUBELIK, A. R., LIVAK, K. J., RAFALSKI, J. A. & TINGEY, S. V. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research* **18**, 6531–6535.

WILSON, C. C. & HERBERT, P. D. N. (1992). The maintenance of taxon diversity in an asexual assemblage: an experimental analysis. *Ecology* **73**, 1462–1472.