

# What has molecular epidemiology ever done for wildlife disease research? Past contributions and future directions

Clare H. Benton · Richard J. Delahay · Hannah Trewby · David J. Hodgson

Received: 25 September 2014 / Revised: 25 November 2014 / Accepted: 30 November 2014 / Published online: 20 December 2014  
© Springer-Verlag Berlin Heidelberg 2014

**Abstract** The increasing availability of novel molecular techniques has transformed the study of human health and disease epidemiology. However, uptake of such approaches has been more conservative in the field of wildlife disease epidemiology. We consider the reasons for this and discuss current and potential applications of molecular techniques in a variety of relevant areas within the field of wildlife disease research. These include conducting wildlife disease surveillance, identifying sources of pathogen emergence, uncovering host-pathogen dynamics and managing current outbreaks, including the development and monitoring of wildlife vaccines. We highlight key examples of applications of molecular epidemiological approaches to wildlife disease scenarios and draw parallels from human disease research to suggest potential future directions. The potential value of next generation sequencing technologies to the field of wildlife disease research is discussed, and initial applications are highlighted, balanced against consideration of the challenges involved. Using a wide range of examples drawn from research into human, livestock and wildlife diseases, we demonstrate the value of using molecular epidemiological approaches at all

scales of wildlife disease research, from pathogen strains circulating at a global scale to intra-individual host-pathogen dynamics. The potential future contribution of these technologies to the field of wildlife disease epidemiology is substantial. In particular, they are likely to play an increasingly important role in helping us to address a principal challenge in the management of wildlife diseases which is how to tease apart the transmission dynamics of complex multi-host systems in order to develop effective and sustainable interventions.

**Keywords** Disease surveillance · Next generation sequencing · Wildlife vaccines · Wildlife reservoirs · Host-pathogen evolution · Disease transmission

## Introduction

Since its emergence in the 1970s, the term ‘molecular epidemiology’ has appeared in a vast number of publications in a wide range of scientific disciplines. Initially, the term was used primarily in the study of human cancer to describe the process of identifying biomarkers within populations which improved identification of subgroups at greater risk of developing disease (Vineis and Perera 2007). However, the term is now widely used in the field of infectious disease biology, where it has been defined as involving ‘the various techniques derived from immunology, biochemistry and genetics for typing and sub-typing pathogens’ (Tibayrenc 1998). A broader definition that goes some way to capturing the breadth of the subject is ‘a science which utilises molecular biology to define the distribution of disease within a population and relies heavily on integration of traditional epidemiological approaches to identify the etiological determinants of this distribution’ (Snow 2011). The influence of molecular epidemiology in the field of human health research has been extensive. Ongoing surveillance of the spatio-temporal distribution of

---

Communicated by C. Gortázar

C. H. Benton (✉) · R. J. Delahay  
National Wildlife Management Centre, Animal and Plant Health Agency (APHA), Woodchester Park, Tinkley Lane, Nympsfield, Stonehouse, Gloucestershire GL10 3UJ, UK  
e-mail: clare.benton@apha.gsi.gov.uk

H. Trewby  
Boyd Orr Centre for Population and Ecosystem Health, Institute of Biodiversity, Animal Health and Comparative Medicine, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow G12 8QQ, UK

D. J. Hodgson  
University of Exeter, Centre for Ecology and Conservation, Daphne du Maurier, College of Life and Environmental Sciences, Penryn Campus, Penryn TR10 9EZ, UK

disease strains has helped uncover drivers of disease transmission (Liu et al. 2008), infer the geographic origin of pathogens (Hemelaar et al. 2011) and provided a baseline against which changes can be detected (Koopmans et al. 2000). In addition, transmission routes of zoonotic pathogens have been identified (Feng and Xiao 2011; Salyer et al. 2012), and the evolutionary provenance of pathogenic strains (Byrnes et al. 2010) and antibiotic resistance mechanisms (Kumarasamy et al. 2010; Hoffmann et al. 2007) have been pinpointed. Molecular technologies have also been widely used in the development of vaccines against human pathogens (Serruto et al. 2009; Santolaya et al. 2012; Seib et al. 2009) and in responsive investigations of disease outbreaks (Grad et al. 2012; Gardy et al. 2011; Rasko et al. 2011). In this review, we will explore similar applications of molecular technologies in the field of wildlife diseases and suggest directions for future applications.

### Bibliometrics

A comprehensive literature search for journal articles published since 1980 using the term ‘molecular epidemiology’ revealed over 111,000 results. In order to consider the growth of molecular approaches in epidemiology, the proportion of molecular epidemiology articles within ‘epidemiology’ articles for the fields of clinical, livestock, zoonoses and wildlife research was calculated. Time series analysis and forecasting of publication trends illustrate the proportion of epidemiology papers incorporating molecular epidemiological approaches for each of the four categories (Fig. 1). When the overall numbers of molecular epidemiology publications within each field are considered, unsurprisingly, the vast majority fall within the field of human clinical research (Fig. 2). The application of molecular epidemiological techniques in livestock, zoonoses and wildlife disease research has steadily grown from the mid 2000s onwards, but such studies still represent only a small proportion of the total observed. The origin of the term in human clinical research is indicated by its earlier appearance in this field, filtering into the other fields shortly afterwards (Fig. 1). Time series forecasting indicates that the proportion of molecular epidemiology publications is predicted to increase relatively steeply in the fields of livestock and zoonotic research, but the trend is less certain in the field of wildlife disease research as indicated by the wider confidence intervals (Fig. 1).

The bias towards human disease is perhaps not surprising or unexpected given that molecular approaches were first developed in this field and human diseases naturally attract a greater level of attention and funding than diseases of animals. It may also reflect the greater focus on non-infectious diseases related to genetic or environmental factors in human systems, for example the large number of studies which employ

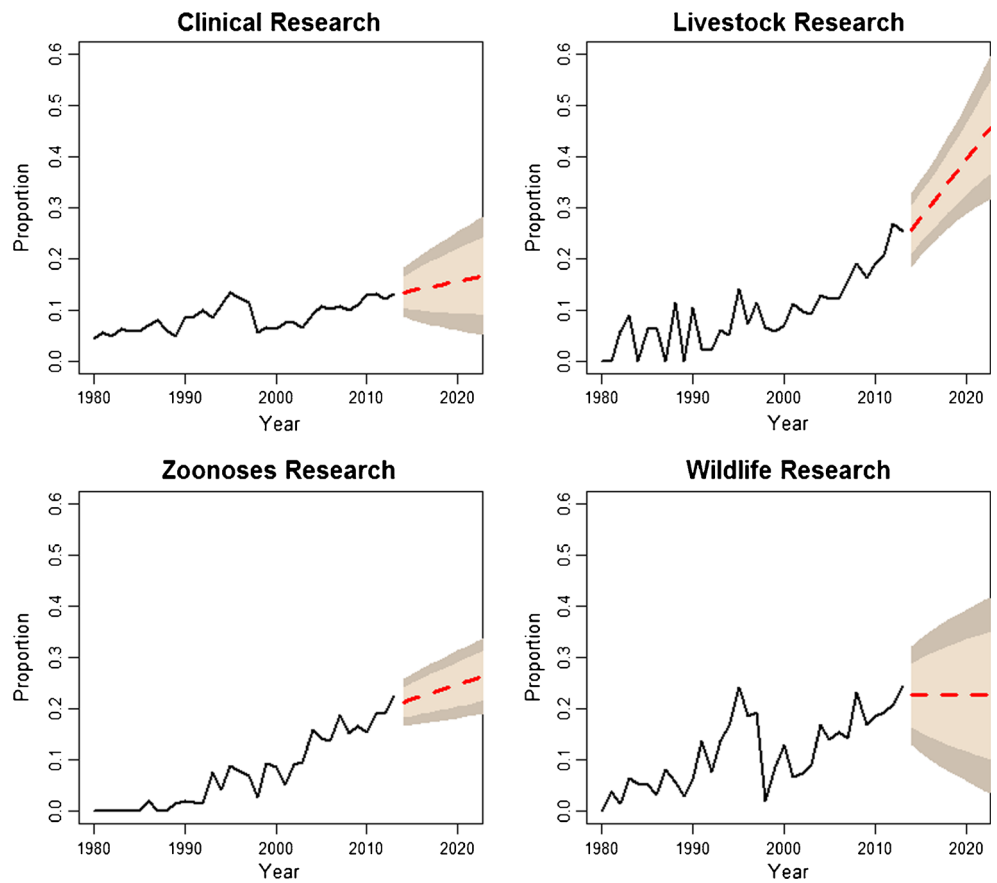
molecular epidemiological approaches to identify biomarkers associated with cancer. In general, diseases of livestock or wildlife have only been considered important when agriculture or human health is potentially threatened (Daszak et al. 2000), which is consistent with the smaller number of publications and conservative uptake of molecular epidemiological techniques in the field of non-zoonotic wildlife disease. There are also a range of practical challenges involved in any study of disease epidemiology in wild populations, including access to individuals, absence of validated diagnostic tests, logistics and costs of sampling, poor baseline surveillance and inherent uncertainty surrounding species ecology and behaviour (Delahay et al. 2008). Nevertheless, as we draw parallels with studies using molecular techniques in other fields, we will see that there are many potentially valuable applications of molecular methods to the epidemiology of disease in wildlife populations. In this review, we consider how molecular epidemiological approaches can help wildlife managers address key questions about disease dynamics, and we suggest directions and opportunities for their wider application in this field.

### Molecular methods

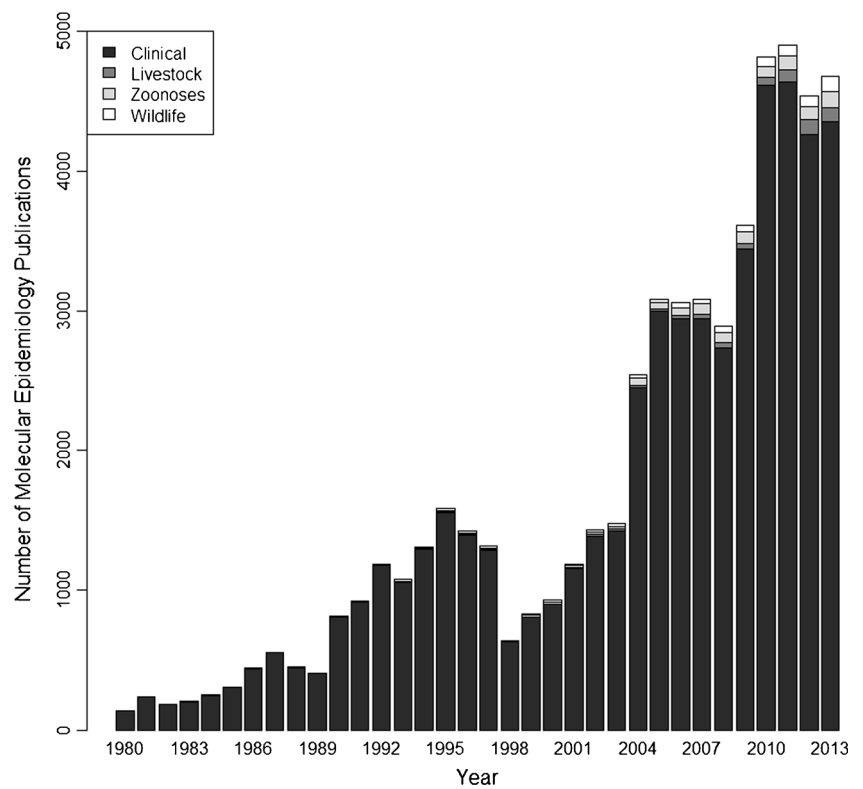
Although DNA-based techniques have been in use for less than 50 years (Medini et al. 2008), during recent decades, a wide range of molecular techniques have emerged for the study of pathogens. One example is typing based on 16S ribosomal RNA (rRNA), in which the percentage of sequence similarity of rRNA molecules between samples is used to classify species (99 % sequence identity is used as the cut-off between separate species) (Medini et al. 2008). In alternative techniques, other genomic elements are sequenced and used as the basis of classification such as housekeeping gene fragments in multi-locus sequence typing (MLST) and enzyme profiles in multi-locus enzyme electrophoresis (MLEE) in bacteria. However, these systems can struggle to distinguish amongst very similar strains (Achtman 2001), such as members of the *Mycobacterium tuberculosis* complex (Frothingham 1995; Köser et al. 2012) or strains of *Bacillus anthracis* (Keim et al. 2001), the causative agent of anthrax.

The choice of an appropriate molecular typing technique requires an understanding of the genomic structure of the pathogen in question. Bacterial genomes consist of a core genome, common to all strains, dispensable genes which are not present in all strains and genomic islands, clusters of contiguous genes with a specialised function (e.g. virulence) (Relman 2011). Classical methods of classifying bacterial pathogens are based on phenotypic characteristics such as cellular structure (Medini et al. 2008), colony morphology (Baron 1996) and antibiotic susceptibility (Harwood et al. 2000). In contrast, viral pathogens contain small genomes, are highly diverse and some can evolve very rapidly (Fierer

**Fig. 1** Time series analysis and forecasting of the proportion of ‘epidemiology’ journal articles that are ‘molecular epidemiology’ in the fields of clinical research, livestock research, zoonoses research and wildlife research. Based on ISI Web of Science search within topic field conducted in April 2014. Time series forecasting carried out using the ‘forecast’ package in R (Hyndman and Khandakar 2007)



**Fig. 2** Number of ‘molecular epidemiology’ journal articles published in the fields of clinical research, livestock research, zoonoses research and wildlife research. Search conducted using ISI Web of Science in April 2014



et al. 2007). For example, the ability of the influenza virus to rapidly change its antigenic profile requires continual development of vaccines (Relman 2011). Genetic diversity in fungal pathogens depends on (a) the mode of reproduction of the species, sexual or asexual, and (b) the presence of ‘transposable elements’, mobile genetic elements which can insert themselves within genes, changing their structure and function (Daboussi 1997). Protozoans such as the human pathogens *Trypanosoma brucei* and *Leishmania major* have genomes with species-specific surface antigens and variable strategies of invading hosts and evading immune responses. Conservation of gene order between species is high, indicating the presence of a strong selection pressure to conserve certain gene clusters and their associated function (El-Sayed et al. 2005).

The underlying genomic structure and diversity within a particular pathogen will influence the choice of typing method. All the typing techniques described above are fundamentally limited as they only examine a small section of the host genome. Depending on the biology of the pathogen, the section under examination will vary in the degree to which it is representative of the whole genome. For example, the typing methods traditionally used to categorize strains of *Mycobacterium bovis* (the causative agent of bovine TB) are spoligotyping (spacer-oligonucleotide typing) and variable number tandem repeat (VNTR) typing, both of which are based on small, hyper-variable genomic regions that are generally evolving at a higher rate than the rest of the genome (Joshi et al. 2012). Such methods are therefore potentially more useful for differentiating between species than detecting finer scale intra-specific variation, although this will depend largely on the genetic diversity within the particular pathogen complex under examination. Rapidly evolving viruses may generate sufficient diversity for intra-specific strains to be differentiated on the basis of more restricted areas of the genome than bacterial pathogens, but only by examining the complete genome of a species can finer genetic structuring be uncovered (Medini et al. 2008). The development of ‘next generation’ sequencing approaches, which base classification on the identification of single nucleotide polymorphisms (SNPs—single base substitutions); insertions or deletions that vary between individual genomes have facilitated the rapid sequencing of whole genomes, opening the door for studies which were previously impossible. In order to construct a phylogenetic tree from a group of sequenced isolates, phylogenetically informative SNPs (i.e. those shared by two or more isolates) are identified through the examination of the maximum sequenceable genome of each isolate. SNPs occur in both coding and non-coding regions, but those in the latter are less likely to exert a phenotypic effect and therefore are less likely to be affected by selective pressures. Hence, whole genome sequencing infers phylogenetic relationships from the maximum amount of genetic information available.

The current next generation sequencing (NGS) technologies are based on breaking the original genome into fragments, which are then extensively sequenced to yield short read sequences. These reads, which can then either be mapped to a reference genome (where present) or assembled against each other (‘*de novo*’ assembly) in order to identify potential SNPs. SNPs which successfully pass the required quality checks can then be used to produce phylogenetic trees and inform transmission models. The cost and required infrastructure for these technologies have so far limited their widespread uptake (Metzker 2010). However, costs are rapidly falling (Köser et al. 2012), uptake is increasing and the incorporation of NGS into routine human disease surveillance (Roetzer et al. 2013) and clinical diagnostics (Boyd 2013) appears to be imminent. Full sequence data has all the potential applications of strain typing but at a far higher resolution and gives the opportunity to determine the extent of the differences amongst strains rather than to simply distinguish them from one another. However, there are challenges in these approaches in relation to the storage, quality control and manipulation of the enormous amounts of data that they generate (Pop and Salzberg 2008). Also, lack of standardisation in bioinformatics protocols limits the extent to which sequences can be compared across laboratories (Köser et al. 2012) and technologies (Metzker 2010). Nevertheless, NGS has enormous potential to uncover fine-scale disease transmission dynamics, which may otherwise remain hidden to epidemiologists.

### Disease surveillance

The importance of surveillance for wildlife diseases is well established (Delahay et al. 2008). Ongoing surveillance can act as an early warning system for outbreaks of new or emerging diseases, allowing pre-emptive management interventions and potentially helping to inform assessment of risk related to conservation interventions such as translocations of endangered populations (Artois et al. 2009).

It is important to assess the extent of genetic diversity within a pathogen population as this has implications for how refined molecular tools need to be in order to investigate disease transmission events. For example, with a low-diversity pathogen such as *Bacillus anthracis*, the most divergent strains are thought to be 99.99 % similar in terms of nucleotide sequencing (Rasko et al. 2011), and therefore, most isolates will appear homogenous, no matter how rigorous the sequencing method. In contrast, within the HIV-1 virus, there is a wealth of genetic diversity which is organised into ‘sub-types’ within which genetic variation can range between 8 and 17 % and can be as much as 35 % amongst sub-types (Hemelaar et al. 2011). Genetically diverse populations of the virus, termed ‘quasi-species’, can be harboured by a single host individual (Pandit and Sinha 2010). Only through

ongoing surveillance of circulating strains can the intrinsic genetic diversity of a pathogen be captured.

Ongoing surveillance of strain diversity can inform the development of diagnostic tests which may be employed for the identification and potentially selective removal of infected individuals in livestock and wildlife populations. The removal of test positive individuals could potentially exert a selection pressure on pathogen populations, with selection favouring strains that produce a weak or negative diagnostic test response. Hence, diagnostic test development is ideally an ongoing process, which seeks to keep one step ahead of such selection pressure. An example from human health is that of *Neisseria meningitidis*, the bacterial cause of meningitis A. Molecular approaches indicate that within this bacterial complex, there are a number of clonal groups, some of which cause disease and others which live commensally within human hosts (Achtman 2001). Horizontal gene transfer between group members can generate genetic diversity. The identity of the most frequent genotype in a population can vary as changing forces of selection favour different strains (Achtman 2001) with clear implications for the design of effective vaccines.

Comparison of pathogen genotypes in regions where infection is endemic versus those where infections only occur sporadically may uncover genetic differences associated with the two scenarios, related for example to pathogenicity factors. Molecular typing may also have important applications in detecting the emergence of new pathogenic strains in populations. In the case of bacteria, the jump from benign to pathogenic could potentially occur relatively rapidly, through the acquisition of a genomic island which codes for a pathogenicity factor (Hacker and Carniel 2001) and molecular typing may aid the detection of such events.

### Phylogeography

Molecular techniques are widely used to describe the spatio-temporal distribution of variant pathogen strains. For example, the characteristic home ranges of *M. bovis* genotypes in cattle have been mapped across the affected areas of the UK (Smith et al. 2003). Routine mapping of this kind may identify the appearance of atypical strains in an area; this may indicate that a 'novel' transmission event has occurred (e.g. the translocation of an infected host animal from another region). Geographic differences in virulence between pathogen strains may also occur, as has been identified for the fungal pathogen *Cryptococcus gatti* (Byrnes et al. 2010). Spatial mapping exercises can also tell us something about the evolution of pathogen strains as geographically dispersed genotypes may be considered more likely to be ancestral strains than those with a restricted home range (Smith et al. 2003). Examining the prevalence of disease in a region can also be used to infer

risk factors which could inform management strategies. Incorporating molecular information into these investigations can provide greater insight into possible causes than simply comparing populations with and without disease (Cowled and Ward 2012).

By examining the strains that are appearing at the moving edge of an epidemic front, it may be possible to gain insights into the factors that are driving disease spread. For example, molecular epidemiology may be a useful tool in determining the proximate causes of new cases of bovine tuberculosis infection in UK cattle at the fringes of the endemic areas, helping to distinguish whether infection is seeded from livestock movements or the presence of infected wildlife. A very different example, focused on conservation of a highly threatened species, is provided by devil facial tumour disease (DFTD) in Tasmanian devils (*Sarcophilus harrisii*), where identifying the location of the disease front has informed management options. Geographic differences have been noted in the epidemiology and population effects of DFTD on devils. Genotyping techniques could be applied to suggest whether this variation is a result of functional sequence variations between strains or differences in disease resistance alleles between populations (Hamede et al. 2012).

When phylogenetic trees of a particular pathogen are overlaid with epidemiological data (such as geographic location of outbreaks), they can be used to map spatial disease spread. This can help epidemiologists infer where transmission events have occurred and therefore potentially to predict and manage future disease risks. For example, examining the geographic localisation of strains of *Mycobacterium leprae*, the bacterial cause of leprosy in humans indicated that global disease spread was most likely linked with historic human migration patterns and trade routes (Monot et al. 2009). Epidemiological linkages between particular populations or geographic locations can be identified if shared genotypes are recorded more often than would be expected by chance (Archie et al. 2009). Host geography has also been found to play a role in rates of pathogen evolution. In the case of *Lyssavirus* (rabies) in bat populations, rates of viral evolution by nucleotide substitution vary depending on whether the host species is in a temperate or tropical environment, which may be related to differences in the seasonality of bat activity and the influence of climate on rates of virus transmission (Streicker et al. 2012). Examining pathogen phylogenies can provide an understanding of rates of new strain emergence, helping epidemiologists to predict and prepare for new disease outbreaks. Also, where transmission rates vary between strains of the same pathogen, either due to differences in infectivity amongst strains or the availability of susceptible hosts, this could be identified through considering rates of spread. Phylogeographic investigations have been conducted on a wide range of human pathogens, including the zoonotic bacteria *Yersinia pestis* (Vogler et al. 2011), dengue virus (Nunes et al. 2012) and

influenza (Cheng et al. 2012). In the case of vector-borne diseases, the same approach can be used to investigate vector distributions, as carried out in a study of *Triatoma infestans*, the primary insect vector of Chagas disease (Perez de Rosas et al. 2011). Phylogeographic approaches have also been used, albeit to a lesser extent, in wildlife and livestock diseases, for example to consider the ecological drivers behind foot and mouth disease (FMD) in cattle (de Carvalho et al. 2013), rates of viral evolution driving infectious bursal disease virus in farmed poultry (Cortey et al. 2012) and the role of the global expansion of fish farming in the spread of salmonid proliferative kidney disease (Henderson and Okamura 2004).

### Roots of emergence

The construction of pathogen phylogenetic trees has made an enormous contribution to the study of human disease, leading to the emergence of the field of evolutionary medicine (Bull 1994). Virulence is known to differ amongst pathogen strains, and this variation is the result of evolutionary processes. Genetic signatures in pathogen phylogenies allow us to look back at the underlying ecological selection pressures which have previously been exerted on a pathogen and shaped its evolution (Biek and Real 2010). Correct inference of ancestry (i.e. determining which strains of a particular pathogen are ancestral and which are descendant) is key to building a clear picture of pathogen population structures (Medini et al. 2008). For example, the population structure of *Mycobacterium bovis* genotypes in the UK suggests a ‘clonal expansion’ of genotype evolution from a common ancestor, through a combination of selection and ‘ecological opportunity’ as invasion into new geographic areas occurred (Smith et al. 2003). Inferring ancestry is also extremely valuable for dating disease transmission events and tracing cross-species transmission in multi-host disease complexes, such as SIV/HIV and hepatitis B in humans and non-human primates (Neel 2010; Starkman et al. 2003). If a pathogen has been transmitted from one species to another, the phylogeny within the recipient species should be nested within that of the source species (Archie et al. 2009). Disease introduction through migration or translocation events can be suggested where there is a genetic mismatch with resident strains, as was recently been inferred for some species of blood parasites in wild birds in Japan where strains of *Leucocytozoon* from migratory and resident birds were phylogenetically separated (Yoshimura et al. 2014). Hence, phylogenetic investigation can be used to identify risk factors for future disease outbreaks.

A substantial body of work exists where whole genome sequencing has been applied to the study of human viral pathogens, such as influenza (Holmes et al. 2005) and HIV (Henn et al. 2012), and in recent years, this approach has also been applied to investigations of viral pathogens in wildlife,

including the detection of highlands J virus in a critically endangered species of crane (Ip et al. 2014), the development of a genome database of orbiviruses (Maan et al. 2013) and an investigation into encephalitis cases in captive polar bears (Szentiks et al. 2014). Work is currently underway to determine the phylogeny of the pathogenic fungus *Geomyces destructans*, the causative agent of white-nose disease in bats (Blehert 2011), with a draft sequence recently published (Chibucos et al. 2013). A number of pathogens of veterinary importance have had at least one isolate sequenced, including African swine fever virus (Chapman et al. 2011), *Mycoplasma haemofelis* (the causative agent of feline infectious anaemia) (Barker et al. 2011) and *Streptococcus equi* (Paillot et al. 2010), although these studies have focused primarily on describing pathogenicity factors, rather than epidemiological outcomes.

The potential for next generation sequencing to infer the origin and population structure of veterinary and wildlife pathogens is substantial. We may expect uptake to be initially greater in the fields of zoonotic and livestock diseases, where the potential human ‘cost’ is perceived to be higher. As illustrated above, phylogenetic approaches offer so much more than an opportunity to delve backwards into the evolutionary history of a pathogen. They can also help us to understand the drivers of the current distribution of pathogens and help us predict their likely distribution in the future.

### Routes of transmission

When investigating the dynamics of infection in a given host population, we reasonably assume that transmission is more likely to have occurred between individuals infected with the same strain of a pathogen than amongst those infected with different strains (Wylie et al. 2005). Pathogen genotyping can therefore help to rule out or implicate particular transmission pathways, which may be valuable in tracing the initial source of infection and preventing further disease spread. The availability of next generation sequencing technologies has allowed contact networks and transmission pathways to be inferred with greater confidence and accuracy (Gardy et al. 2011). Given the relatively recent availability of these technologies, and their decreasing cost, their full potential in the field of human health has yet to be realised (Walker et al. 2013), and to date, their use in relation to livestock and wildlife diseases has been limited. However, there are notable examples such as studies of TB in cattle and badgers in the UK (Biek et al. 2012), brucellosis in livestock and wildlife (Foster et al. 2009) and MRSA in livestock (Price et al. 2012). In studies of human pathogens such as *M. tuberculosis* (Cook et al. 2007; Gardy et al. 2011; Walker et al. 2013), MRSA (Harris et al. 2013), *Clostridium difficile* (Eyre et al. 2013) and *Chlamydia trachomatis* (Wylie et al. 2005), genotyping of pathogenic isolates has informed contact tracing, suggested

the existence of undetected carriers and helped to both construct and verify the conclusions of social network analysis of disease outbreaks. Clinical disease outbreaks in human populations are often treated on a ‘case by case’ basis, on the understanding that no two events are epidemiologically identical. On the other hand, wildlife disease managers are often called upon to use simple management strategies to tackle disease in multiple socially structured populations, without information on the particular transmission dynamics in each situation. The overlaying of data on pathogen strain diversity onto ecological information could be used in wildlife populations to assess transmission rates in relation to population structure (e.g. social groups and herds). In the case of the European badger, the prevailing social structure in high-density populations has been associated with the clustering of infection within social groups (Delahay et al. 2000). Disruption of this social structure, as observed following culling, leads to a reduction in this clustering, as surviving individuals range more widely (Jenkins et al. 2007). Further information on the role of social behaviour in the spread of infection may be achievable by investigating the genetic diversity of *M. bovis* strains in badger populations. If social structure acts as a barrier to disease spread, then we would expect the degree of relatedness amongst *M. bovis* strains within badger social groups to be greater than that observed amongst social groups. Wherever wildlife is implicated as a reservoir of zoonotic and/or livestock disease, such approaches may be valuable in identifying chains of disease transmission between species and could potentially indicate the direction of disease transmission.

In order to make meaningful inferences about transmission dynamics, a pathogen must be acquiring mutations within an epidemiologically meaningful timeframe, and the genotyping method applied must have the ability to detect this variation (Grad et al. 2012). Epidemiologists studying pathogens with very little variation between strains will require a typing method that is able to detect small differences between isolates. Where discrimination between isolates is not possible using conventional methods, whole genome sequencing (WGS) may be the only tool suitable for looking at fine-scale transmission dynamics. The exceptionally high level of genetic resolution achievable using WGS means that even sequencing a restricted number of isolates can reveal a wealth of epidemiologically valuable information. Where access to long-term wildlife studies is possible, a ‘phylogenetic’ approach (Grenfell et al. 2004) of overlaying pathogen phylogenies onto well-documented epidemiological systems is potentially very powerful. Novel molecular approaches are not a replacement for traditional epidemiological investigations but are complementary, allowing a finer scale approach. For this reason, long-term, well-studied epidemiological systems are the ideal scenarios in which to explore the contribution of cutting edge sequencing to uncovering the drivers of disease

transmission. Examples of such well-studied systems include TB infection in wild badgers (Delahay et al. 2000) and meerkats (Drewe 2010), chronic wasting disease in white-tailed deer (Williams et al. 2002) and DFTD in Tasmanian devils (Hamede et al. 2009). It is important to note, however, that, even with the added resolution provided by WGS, there are considerable challenges to identifying pathogen transmission chains. The point at which mutations are acquired in a given transmission sequence is unknown, and when mutation rates are slow compared to pathogen generation time, closely related isolates may appear genetically identical as they lack informative mutations (Kao et al. 2014).

### Host-pathogen dynamics

Pathogens can have widely differing effects in different host species, as is the case for the squirrel parapox virus which causes severe disease in the European Red Squirrel (*Sciurus vulgaris*) but has no observed effects on the North American Grey Squirrel (*Sciurus carolinensis*) (Sainsbury et al. 2000). Variability in the observed costs of pathogen infection has also been observed amongst individuals of the same species. Heterogeneities in susceptibility to infection among individuals can affect the estimation of the transmission parameter  $R_0$  (the basic reproductive number) (Hudson et al. 2002). In such instances, molecular techniques may allow us to distinguish between differences in pathogenicity which arise from strain variation and those which reflect heterogeneity in host immune responses. Scaling up these effects can impact on host population dynamics as regulation by a pathogen requires its *per capita* impact to outweigh the intrinsic population growth rate (Hudson et al. 2002). If the *per capita* impact on host fitness is widely variable amongst individuals, then inferring population regulation is more complicated. Variation in how a pathogen physiologically affects individuals within a population has implications for onward transmission and persistence of disease (Cross et al. 2005). For example, inter-individual variation in the amount or concentration of pathogenic material excreted and the duration over which this occurs is likely to affect the number of secondary cases observed. Ignoring this individual heterogeneity and assuming that each infected individual contributes to the same number of secondary infections can lead to highly inaccurate estimations of  $R_0$ . Molecular techniques can also allow us to examine individual variation in susceptibility and resistance within a host population by assessing the genetic basis of the immune response: an approach known as immunogenetics. Individuals with greater allelic diversity within immune genes such as the major histocompatibility complex (MHC) are able to mount an appropriate immune response against a greater variety of pathogens (Castro-Prieto et al. 2012). Accounting for heterogeneity in individual susceptibility and for differential strain

pathogenicity is likely to allow  $R_0$  to be estimated with a greater degree of accuracy.

It has been suggested that immunogenetic data should be used more widely to complement wildlife management decisions, particularly where populations are restricted or fragmented with a limited gene pool, as is often the case for highly endangered species (Acevedo-Whitehouse and Cunningham 2006). A key example of this approach is investigation into the spread by biting of a contagious cancer amongst Tasmanian devils which is thought to have caused a 90 % population decline (Siddle et al. 2010). The absence of an immune response in infected devils is thought to be linked to the limited genetic diversity within their MHC complex (Siddle et al. 2007). Examination of MHC genetic diversity within devil populations in other areas of Tasmania, where the disease is absent or at low prevalence, have identified some unique profiles which may confer disease resilience. If this were the case, then selective breeding and translocation of resilient individuals has been suggested as a means of controlling disease spread (Hamede et al. 2012). In contrast, as MHC profiles are likely to be adapted to local pathogenic selection pressures, a poor choice of origin or destination could leave a translocated individual unable to cope with a different pathogen community and so at a selective disadvantage (Castro-Prieto et al. 2012). Hence, the application of sequence-based approaches to assess the immunogenetic characteristics of populations of endangered species may have the potential to increase the likelihood of successful translocation (Boyce et al. 2011).

Molecular epidemiology also allows us to zoom in to an even finer scale than that of inter-individual variation, and to consider intra-individual host-pathogen dynamics. Infectious pathogens persist in the context of a co-evolutionary arms race with the host (Acevedo-Whitehouse and Cunningham 2006) which can be considered as a habitat ‘patch’ occupied by a parasite ‘community’ (Hudson et al. 2002). Where an individual host is infected with multiple strains of the same pathogen, strain competition can occur, with certain strains favoured owing to their faster growth rate or ability to grow in a certain tissue (Bull 1994). It is interesting to consider, however, that the traits which allow a particular strain to dominate within the host environment may not necessarily optimise onward transmission although they may increase pathogen virulence (Bull 1994). However, outcomes of multiple infection on evolution of virulence and subsequent effects on individual host fitness are variable (Rankin et al. 2007). The application of suitable molecular techniques to detect multiple strains of a pathogen within a host and to detect within-host pathogen strain evolution or strain competition may have important implications, as for example the scale of competition between bacterial strains is thought to influence the evolution of virulence (Griffin et al. 2004). Comparative genomics studies, in which the aim is to link genetic sequence differences between strains with

phenotypic differences in the host (e.g. differential pathogenicity), have acquired valuable additional resolution from the development of next generation sequencing technologies (Hu et al. 2011).

Understanding variations in the impact of pathogens both amongst and within individuals may be critical to achieving effective management at the population level. Furthermore, disregarding heterogeneity in host responses and failing to acknowledge within host-pathogen dynamics (such as the role of multiple infection) may result in unexpected, potentially adverse, outcomes of management interventions. Molecular approaches have much to offer at both scales of analysis.

### Vaccine development and monitoring

Vaccination is currently being used or considered as a management option in several high profile wildlife disease scenarios, including the control of bovine TB in badgers in the UK (Chambers et al. 2010; Carter et al. 2012), chlamydia in koalas in Australia (Carey et al. 2010), haemorrhagic disease and myxomatosis in rabbits in Europe (Spibey et al. 2012). Molecular epidemiology has a key role to play in the development of effective vaccines for wildlife and monitoring their impacts on disease epidemiology. Vaccine development against human pathogens has greatly benefited from technological advances in gene sequencing, and now, the sequences of many pathogens are available. This has led to the emergence of the field of ‘reverse vaccinology’ which typically involves mining the pathogen sequence for antigens that may be suitable as vaccine targets (Serruto et al. 2009). In pan-genome reverse vaccinology, multiple isolates of a pathogen species are considered. This is based on the idea of the existence of a ‘pan-genome’, which acknowledges that any single isolate of a pathogen does not exhibit all the genetic diversity within that species, especially if they are capable of generating genetic diversity through recombination or horizontal gene transfer. Consequently, it is necessary to sequence multiple genomes in order to get a better measure of the entire genomic repertoire of a species (Tettelin et al. 2008). In comparative reverse vaccinology, sequences of pathogenic strain isolates are compared with those of non-pathogenic isolates of the same species, in order to identify antigens associated with pathogenicity. The first human pathogen for which a vaccine has been developed and recently licensed using this approach is serogroup B meningitis (*N. meningitidis*), responsible for 80 % of meningitis cases in Europe (Santolaya et al. 2012).

As well as informing the development of vaccines, genome sequencing technologies also have applications for monitoring the effectiveness of vaccine deployment. As an increasing proportion of a population is immunised, the selection pressure favouring strain variants that are unaffected by vaccination will grow. The emergence of these strains, known as



‘escape mutants’, could be monitored by sequencing isolates before and after vaccination to look for new mutations related to immunity in the targeted proteins (Seib et al. 2009). This sort of approach could be extremely valuable in monitoring the impacts of vaccination in wildlife populations. Comparing the genetic diversity of pathogen populations before, during and after vaccine deployment could provide valuable information on the potential emergence of resistant strains or differential vaccine performance against variant strains. Strain typing could also help monitor reversion to virulence of live vaccines, as vaccination has on occasion been observed to result in clinical disease. This was recently reported in a red fox (*Vulpes vulpes*) in which strain typing was able to identify the live rabies vaccine as the aetiological agent (Hostnik et al. 2014). Live vaccines, such as the oral rabies vaccine, may be derived from multiple strains. Genetic characterisation of these strains can uncover the genetic basis for their attenuation (Geue et al. 2008). Population coverage of vaccines which may be horizontally transmissible within a population (Angulo and Juan 2007) could also be monitored using molecular diagnostics.

Understanding the antigenic diversity of a pathogen is key in vaccine design and is only possible through ongoing surveillance as the most frequent antigenic strain of a pathogen in a population may change in response to the selection pressure of immunisation (Achtman 2001), favouring new antigenic types which are able to evade the acquired immunity of the host (Bull 1994). Pathogens such as HIV, malaria and influenza have particularly high antigenic diversity (Buckee et al. 2011). A vaccine must either induce cross-immunity to all antigenic strains of a pathogen circulating within a population, or different vaccines may be required according to which antigenic strain is predominant in a given situation. Different strains of a pathogen may also vary in terms of the magnitude or type of immune response invoked (Wedlock et al. 2007). In the case of human seasonal influenza, it has been suggested that it is the changing immune response within the host population which creates the conditions for the emergence of the next dominant strain (Recker et al. 2007). Molecular typing approaches offer powerful tools for furthering our understanding of antigenic diversity in wildlife populations and the role of vaccination in disease control.

### Identifying reservoirs

Another key challenge for wildlife managers is identifying populations that may act as reservoirs of infection for livestock, humans or other wildlife of conservation or economic importance. Assessment of the risks of onward spread requires a clear understanding of transmission dynamics within and amongst the species concerned (Hudson et al. 2002). Disease reservoirs can potentially be composed of one or more

epidemiologically connected populations or environments where the pathogen can be permanently maintained (Haydon et al. 2002). Molecular techniques may be of value in inferring transmission routes amongst multiple host species, although confusion can arise if pathogens are capable of remaining infectious in the environment. Inter-specific transmission has been inferred through strain comparison of *Giardia* (Feng and Xiao 2011) and *Cryptosporidium* in human and animal hosts (Xiao and Ryan 2004). Transmission between wildlife and commercially important livestock can also be inferred through comparing pathogen genotypes, as demonstrated for bovine TB in cattle and badgers in the UK (Goodchild et al. 2012; Biek et al. 2012; Woodroffe et al. 2005) and for *Babesia bovis* and *B. bigemia*, the bacteria responsible for cattle tick fever in white-tailed deer in the USA (Holman et al. 2011).

Molecular typing techniques provide valuable insights into multi-host systems when considering populations of conservation concern. Examples include hookworm and feline leukaemia virus transmission from domestic cats (*Felis catus*) to the critically endangered Iberian lynx (*Lynx pardinus*) (Millan and Blasco-Costa 2012; Meli et al. 2009) and transmission of canine parvovirus and rabies from domestic dogs (*Canis lupus familiaris*) to endangered African wild dogs (*Lycaon pictus*) (Woodroffe et al. 2012). Strain typing of pathogens can also indicate the presence of an undetected wildlife reservoir, or even multiple reservoirs, where strain diversity appears too high to have been generated by mutation alone. However, in order to make such assessments, a sufficient number of samples should ideally be available from all host species in the system under study, and any host-related variation in pathogen mutation rates should be known (Kao et al. 2014).

### Management strategies

One of the greatest challenges faced by wildlife disease managers is unpredictability in the outcome of interventions. This is in part due to the fundamental challenges of working with free-ranging wildlife, but is exacerbated by a lack of reporting when unintended outcomes occur, which has limited the degree to which we can learn from past interventions (Lloyd-Smith et al. 2005). Coupling genetic information from hosts and pathogens, with ecological factors, can help to predict patterns of disease emergence, spread and control (Biek and Real 2010). Employing molecular approaches can help managers to monitor the epidemiological impacts of interventions with a potentially high degree of resolution and hence allow a more informed approach to refining management actions. Where a novel or re-emerging pathogen appears and wildlife populations are implicated, either as the reservoir or target of disease, managers may be called on to advise on potential interventions. Rapid molecular typing can quickly reveal a wealth of information about a disease outbreak and help to

identify true transmission events, trace individual contacts and identify the true source of a particular pathogen. In the field of public health, molecular strain typing has been used to trace the source for outbreaks of a wide range of pathogens including *E. coli* (Grad et al. 2012), TB (Gardy et al. 2011), pneumonia (Snitkin et al. 2012) and even deliberately introduced pathogens associated with bioterrorism (Rasko et al. 2011). Molecular epidemiological investigations during an outbreak can also suggest the existence of undetected carriers through using pathogen phylogenies in association with social network analysis, as conducted in investigations of human TB outbreaks (Walker et al. 2013) and can help identify super-spreading individuals who make a disproportionately large contribution to secondary infections (Woolhouse et al. 1997). Through the comparison of multiple isolates from the same host individual over time, pathogen micro-evolution can be examined (Gardy et al. 2011). Understanding the rate at which a pathogen can acquire mutations has important implications for choosing appropriate diagnostic tests, predicting the emergence of new strains and informing potential intervention strategies, such as vaccination. Comparison of pathogen strains prior to and during an outbreak can indicate whether the epidemic is due to a genetic change in the pathogen or rather to some social or environmental trigger (Gardy et al. 2011). The genetic diversity amongst isolates associated with a particular disease outbreak can also provide information about the size of the initial infection; limited diversity among isolates may indicate that a population bottleneck has occurred, suggesting that the outbreak could have been caused by few initially infected individuals (Grad et al. 2012). However, this requires pre-existing knowledge regarding what level of diversity is typical for that pathogen, which highlights the importance of ongoing disease surveillance.

Molecular epidemiological investigations have been carried out on a wide range of disease outbreaks in livestock, including Newcastle disease in poultry (Gould et al. 2001), FMD in cattle (Cottam et al. 2006) and bluetongue virus in sheep (Maan et al. 2004; Barros et al. 2007). Of these, only the investigation of FMD employed complete genome sequencing. Examples of molecular epidemiological investigations in wildlife include studies on outbreaks of phocine distemper in seals on the Danish coast (Line Nielsen et al. 2009), salmonella in passerines (Hernandez et al. 2012), viruses of anthropogenic origin in protected ape populations (Köndgen et al. 2008) and the source of DFTD in Tasmanian devils (Murchison et al. 2012).

A major wildlife disease outbreak which represents a real threat to global biodiversity is the recent emergence of amphibian chytridiomycosis, caused by the fungus *Batrachochytrium dendrobatidis*. This pathogen has been isolated from all continents where its hosts are found (Fisher et al. 2009) and is thought to be the principal cause of decline in over 200 species of amphibian. There is substantial variation

in observed host responses to infection with some species appearing to be resistant whilst others succumb quickly to its lethal effects, and virulence has been found to vary amongst strains (Blaustein et al. 2005). The full genomes of two geographically diverse chytrid isolates were sequenced and used to identify areas of variation within the genome. Low sequence diversity was observed between the two isolates, but genomic areas with some variation were targeted by multi-locus sequence typing of a global set of chytrid isolates which were used to create a phylogenetic tree, illustrating the geographic origin of each isolate and its host species (James et al. 2009). From examining the tree, it was apparent that all the isolates could feasibly have originated from a single clonal lineage as, in one host animal, the same chytrid sequence diversity existed as was found in the whole global sample (James et al. 2009). This molecular signature is consistent with the rapid spread of a novel pathogen, and hence, movement of animals for trade purposes has been suggested as a potential explanation for its current global distribution (Fisher et al. 2009).

Combining molecular epidemiological approaches, in particular high resolution sequencing, with traditional epidemiological techniques may be a powerful approach in disease outbreak investigation. This is made even more powerful where data is also available from background pathogen surveillance.

## Discussion

The value of molecular epidemiology in the study of human disease is well-established. We now have at least one complete genomic sequence for nearly all bacteria responsible for human disease. An extraordinary amount of genetic diversity has been uncovered, including variation from within clonal cultures (Medini et al. 2008). Phylogenetic tools can be applied to genetic sequence data within open source packages such as BEAST (Drummond et al. 2012) providing powerful insights into pathogen spread within host populations. Also, pathogen sequence data can improve the performance of disease transmission models by reducing the number of candidate transmission trees (Kao et al. 2014), as pathogen genetic data is integrated with epidemiological data to inform transmission model construction within a Bayesian framework (Jombart et al. 2014). Despite the particular challenges involved in applying molecular technologies in the field of wildlife disease, the emergence of a number of high profile zoonotic diseases and dramatic declines in the abundance of some wildlife populations in recent years have raised awareness of this area of study (Daszak et al. 2004) and the application of cutting edge molecular tools is increasing.

The application of molecular technologies poses significant challenges even when used in tandem with traditional

**Table 1** Summary of applications of molecular epidemiology to wildlife disease research, including key examples

Application	Summary	Examples of use in wildlife
Disease surveillance	Ongoing surveillance of circulating pathogen strains can help to capture the intrinsic genetic diversity of the pathogen, informing the use and development of appropriate diagnostics	Avian influenza viruses in wild birds (Hoye et al. 2010) Bovine tuberculosis in wildlife (Romero et al. 2008)
Phylogeography	Spatio-temporal mapping of variant pathogen strains can highlight the appearance of atypical strains in an area, identify geographic differences in pathogen virulence and help to infer risk factors	Devil facial tumour disease (Hamede et al. 2012) Rabies virus (Streicker et al. 2012) Foot & Mouth Disease (de Carvalho et al. 2013)
Roots of emergence	The construction of phylogenetic trees from genotype data can indicate which pathogen strains are ancestral and which are descendant. This can be useful in dating disease transmission between populations and multiple host species.	<i>Leucocytozoon</i> blood parasites in wild birds (Yoshimura et al. 2014) White Nose Disease in bats (Chibucos et al. 2013) Orbiviruses (Maan et al. 2013)
Routes of transmission	Pathogen genotyping can help to rule out or implicate particular transmission pathways, which may be valuable in tracing the initial source of infection and preventing further disease spread	Bovine TB in badgers and cattle (Biek et al. 2012) Brucellosis (Foster et al. 2009)
Host-pathogen dynamics	Molecular techniques can allow us to examine individual variation in susceptibility and resistance within a host population by assessing the genetic basis of the immune response	Devil facial tumour disease (Siddle et al. 2010) MHC differentiation in Namibian cheetahs (Castro-Prieto et al. 2012) Bovine TB resistance in wild boar (Acevedo-Whitehouse and Cunningham 2006)
Vaccine development and monitoring	Molecular epidemiology has a key role to play in the development of effective vaccines for wildlife and monitoring their impacts on disease epidemiology	Reversion to virulence of rabies vaccine strain (Hostnik et al. 2014) Oral rabies vaccine strains in wildlife (Geue et al. 2008)
Identifying reservoirs	Molecular techniques may be of value in inferring transmission routes amongst multiple host species. Strain typing of pathogens can also indicate the presence of an undetected wildlife reservoir, or even multiple reservoirs, where strain diversity appears too high to have been generated by mutation alone	Rabies in African wild dogs (Woodroffe et al. 2012) Feline Leukaemia Virus in Iberian lynx (Millan and Blasco-Costa 2012), (Meli et al. 2009) Cattle Tick Fever (Holman et al. 2011)
Management strategies	Employing molecular approaches can help managers to identify the sources of disease outbreaks and to monitor the epidemiological impacts of interventions with a potentially high degree of resolution	Phocine Distemper in seals (Line Nielsen et al. 2009) Salmonella in passerines (Hernandez et al. 2012) Viruses in protected ape populations (Köndgen et al. 2008)

epidemiological approaches. Techniques such as whole genome sequencing produce huge amounts of data which can be expensive to store and computationally costly to handle. However, the availability of online ‘cloud’ storage provides a potential solution (Baker 2010), and as uptake of these technologies increases, we can expect further developments in data storage and handling capabilities. DNA amplification required for next generation sequencing can introduce sequencing errors and a lack of standardised quality control procedures between laboratories can add uncertainty to sequence data (Kao et al. 2014). Particular challenges for the application of molecular approaches in wildlife populations include the presence of multiple hosts and the possibility of environmental persistence of the pathogen. Such circumstances mean that even whole genome sequencing cannot pick out individual transmission pathways as there will often be multiple routes by which the same pattern of genotypes could have arisen.

Future developments in molecular technology could have exciting applications in the field of wildlife disease. Rapid, field sequencing of isolates from populations and their environment (for example using hand-held

sequencers) could allow a ‘forensic’ approach to investigating disease outbreaks, in which localised management might be tailored to the particular source of infection. This could be useful for example in the case of bTB in UK cattle, where the source of infection is likely to vary widely between herds and geographic locations. In the field of human health, interest is growing in ‘precision medicine’ whereby the entire human genome of a patient is sequenced and a tailored health plan produced based on the patient’s particular genetic composition. It is conceivable that human genome sequencing will become a routine procedure at birth, allowing the development of ‘personalised programmes of lifelong health promotion’ (Tonellato et al. 2011). As we have seen by considering a variety of examples above, technological advances first developed in the field of human health are subsequently employed in livestock and wildlife. It is plausible therefore that as sequencing costs fall, individual level, genome tailored approaches may become attractive for the management of disease in wildlife species of very high conservation value. The management of DFTD in individual Tasmanian devils might be a case in point (Table 1).

Developments in the field of metagenomics, in which multiple microbe genomes could be sequenced directly from environmental samples (Doolittle and Zhaxybayeva 2010), may provide valuable tools for wildlife disease management and research. Such an approach could be used to screen for pathogens prior to translocation of threatened species or for clarifying transmission routes where a pathogen can persist in the environment. The latter would for example be useful in studies of bTB transmission amongst wildlife and livestock, where there is a potential role for environmental contamination with *M. bovis* (Duffield and Young 1985; Vicente, in press).

In summary, molecular technologies allow us to consider pathogens at a wide range of spatial and temporal scales: from individual host-pathogen dynamics, to global patterns of strain diversity. Following their emergence in the field of human health, they have begun to be adopted for the purposes of investigation and management of disease in wildlife. At the present time, these tools have a range of applications in wildlife disease research from the local investigation of disease outbreaks to unearthing the evolutionary history and global spread of pathogens. The potential future contribution of these technologies to the field of wildlife disease epidemiology is substantial. In particular, they are likely to play an increasingly important role in helping us to address a principal challenge in the management of wildlife diseases which is how to tease apart the transmission dynamics of complex multi-host systems in order to develop effective and sustainable interventions.

## References

- Acevedo-Whitehouse K, Cunningham AA (2006) Is MHC enough for understanding wildlife immunogenetics? *Trends Ecol Evol* 21:433–438
- Achtman M (2001) A phylogenetic perspective on molecular epidemiology. *Molecul Med Microbiol* 1–3:485–509
- Angulo E, Juan B (2007) Towards a unique and transmissible vaccine against myxomatosis and rabbit haemorrhagic disease for rabbit populations. *Wildl Res* 34:567–577
- Archie EA, Luikart G, Ezenwa VO (2009) Infecting epidemiology with genetics: a new frontier in disease ecology. *Trends Ecol Evol* 24:21–30
- Artois M, Bengis R, Delahay RJ, Duchêne M-J, Duff JP, Ferroglio E, Gortazar C, Hutchings MR, Kock RA, Leighton FA (2009) Wildlife disease surveillance and monitoring. In: *Management of disease in wild mammals*. Springer, pp 187–213
- Baker M (2010) Next-generation sequencing: adjusting to data overload. *Nat Meth* 7:495–499
- Barker EN, Darby AC, Helps CR, Peters IR, Heesom KJ, Arthur CJ, Crossett B, Hughes MA, Radford AD, Tasker S (2011) Molecular characterization of the uncultivable hemotropic bacterium *Mycoplasma haemofelis*. *Vet Res* 42:83
- Baron S (1996) *Medical microbiology*, 4th edn. University of Texas Medical Branch, Galveston
- Barros SC, Ramos F, Luis TM, Vaz A, Duarte M, Henriques M, Cruz B, Fevereiro M (2007) Molecular epidemiology of bluetongue virus in Portugal during 2004–2006 outbreak. *Vet Microbiol* 124:25–34
- Biek R, Real LA (2010) The landscape genetics of infectious disease emergence and spread. *Mol Ecol* 19:3515–3531
- Biek R, O'Hare A, Wright D, Mallon T, McCormick C, Orton RJ, McDowell S, Trewby H, Skuce RA, Kao RR (2012) Whole genome sequencing reveals local transmission patterns of *Mycobacterium bovis* in sympatric cattle and badger populations. *PLoS Pathog* 8:e1003008
- Blaustein AR, Romansic JM, Scheessele EA, Han BA, Pessier AP, Longcore JE (2005) Interspecific variation in susceptibility of frog tadpoles to the pathogenic fungus *Batrachochytrium dendrobatidis*. *Conserv Biol* 19:1460–1468
- Blehert DS (2011) Population genetics of *Geomyces destructans* from whole genome sequences US geological survey
- Boyce WM, Weisenberger ME, Penedo MCT, Johnson CK (2011) Wildlife translocation: the conservation implications of pathogen exposure and genetic heterozygosity. *BMC Ecol* 11(1):5
- Boyd SD (2013) Diagnostic applications of high-throughput DNA sequencing. *Annu Rev Pathol* 8:381–410
- Buckee CO, Recker M, Watkins ER, Gupta S (2011) Role of stochastic processes in maintaining discrete strain structure in antigenically diverse pathogen populations. *Proc Natl Acad Sci U S A* 108:15504–15509
- Bull JJ (1994) Perspective: virulence. *Evolution* 48:1423–1437
- Byrnes EJ III, Li W, Lewit Y, Ma H, Voelz K, Ren P, Carter DA, Chaturvedi V, Bildfell RJ, May RC, Heitman J (2010) Emergence and pathogenicity of highly virulent *Cryptococcus gattii* genotypes in the Northwest United States. *PLoS Pathog* 6:e1000850
- Carey AJ, Timms P, Rawlinson G, Brumm J, Nilsson K, Harris JM, Beagley KW (2010) A multi-subunit chlamydial vaccine induces antibody and cell-mediated immunity in immunized koalas (*Phascolarctos cinereus*): comparison of three different adjuvants. *Am J Reprod Immunol* 63:161–172
- Carter SP, Chambers MA, Rushton SP, Shirley MDF, Schuchert P, Pietravalle S, Murray A, Rogers F, Gettinby G, Smith GC, Delahay RJ, Hewinson RG, McDonald RA (2012) BCG vaccination reduces risk of tuberculosis infection in vaccinated badgers and unvaccinated badger cubs. *PLoS one* 7:e49833
- Castro-Prieto A, Wachter B, Melzheimer J, Thalwitzer S, Hofer H, Sommer S (2012) Immunogenetic variation and differential pathogen exposure in free-ranging cheetahs across namibian farmlands. *PLoS One* 7:e49129
- Chambers MA, Rogers F, Delahay RJ, Lesellier S, Ashford R, Dalley D, Gowtage S, Davé D, Palmer S, Brewer J, Crawshaw T, Clifton-Hadley R, Carter S, Cheeseman C, Hanks C, Murray A, Palphramand K, Pietravalle S, Smith GC, Tomlinson A, Walker NJ, Wilson GJ, Corner LA, Rushton SP, Shirley MD, Gettinby G, McDonald RA, Hewinson RG (2010) Bacillus Calmette-Guérin vaccination reduces the severity and progression of tuberculosis in badgers. *Proc Biol Sci* 278(1713):1913–1920
- Chapman DA, Darby AC, Da Silva M, Upton C, Radford AD, Dixon LK (2011) Genomic analysis of highly virulent Georgia 2007/1 isolate of African swine fever virus. *Emerg Infect Dis* 17:599–605
- Cheng X, Tan Y, He M, Lam TT-Y, Lu X, Viboud C, He J, Zhang S, Lu J, Wu C, Fang S, Wang X, Xie X, Ma H, Nelson MI, H-f K, Holmes EC, Cheng J (2012) Epidemiological dynamics and phylogeography of influenza virus in Southern China. *J Infect Dis* 207(1):106–114
- Chibucos MC, Crabtree J, Nagaraj S, Chaturvedi S, Chaturvedi V (2013) Draft genome sequences of human pathogenic fungus *Geomyces pannorum* sensu lato and bat white nose syndrome pathogen *Geomyces (Pseudogymnoascus) destructans*. *Genome Announce* 1(6):e01045–13
- Cook VJ, Sun SJ, Tapia J, Muth SQ, Argüello DF, Lewis BL, Rothenberg RB, McElroy PD (2007) Transmission network analysis in tuberculosis contact investigations. *J Infect Dis* 196:1517–1527

- Cortey M, Bertran K, Toskano J, Majo N, Dolz R (2012) Phylogeographic distribution of very virulent infectious bursal disease virus isolates in the Iberian Peninsula. *Avian Pathol* 41:277–284
- Cottam EM, Haydon DT, Paton DJ, Gloster J, Wilesmith JW, Ferris NP, Hutchings GH, King DP (2006) Molecular epidemiology of the foot-and-mouth disease virus outbreak in the United Kingdom in 2001. *J Virol* 80:11274–11282
- Cowled BD, Ward MP (2012) Integrating survey and molecular approaches to better understand wildlife disease ecology. *PLoS One* 7:1–8
- Cross PC, Lloyd-Smith JO, Johnson PLF, Getz WM (2005) Duelling timescales of host movement and disease recovery determine invasion of disease in structured populations. *Ecol Lett* 8:587–595
- Daboussi MJ (1997) Fungal transposable elements and genome evolution. In: Capy P (ed) *Evolution and impact of transposable elements*. Springer, Netherlands, pp 253–260
- Daszak P, Cunningham AA, Hyatt AD (2000) Emerging infectious diseases of wildlife - threats to biodiversity and human health. *Science* 287:443–449
- Daszak P, Tabor GM, Kilpatrick AM, Epstein J, Plowright R (2004) Conservation medicine and a new agenda for emerging diseases. *Ann N Y Acad Sci* 1026:1–11
- de Carvalho LM, Santos LB, Faria NR, de Castro SW (2013) Phylogeography of foot-and-mouth disease virus serotype O in Ecuador. *Infect Genet Evol* 13:76–88
- Delahay RJ, Langton S, Smith GC, Clifton-Hadley RS, Cheeseman CL (2000) The spatio-temporal distribution of *Mycobacterium bovis* (bovine tuberculosis) infection in a high-density badger population. *J Anim Ecol* 69:428–441
- Delahay RJ, Smith GC, Hutchings MR (2008) The science of wildlife disease management. In: Delahay RJ, Smith GC, Hutchings MR (eds) *Management of disease in wild mammals*. Springer, pp 1–8
- Doolittle WF, Zhaxybayeva O (2010) Metagenomics and the units of biological organization. *BioSci* 60:102–112
- Drewe JA (2010) Who infects whom? Social networks and tuberculosis transmission in wild meerkats. *Proc Biol Sci* 277:633–642
- Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol* 29:1969–1973
- Duffield B, Young D (1985) Survival of *Mycobacterium bovis* in defined environmental conditions. *Vet Microbiol* 10:193–197
- El-Sayed NM, Myler PJ, Blandin G, Berriman M, Crabtree J, Aggarwal G, Caler E, Renaud H, Worthey EA, Hertz-Fowler C, Ghedin E, Peacock C, Bartholomeu DC, Haas BJ, Tran A-N, Wortman JR, Alsmark UCM, Angiuoli S, Anupama A, Badger J, Bringaud F, Cadag E, Carlton JM, Cerqueira GC, Creasy T, Delcher AL, Djikeng A, Embley TM, Hauser C, Ivens AC, Kummerfeld SK, Pereira-Leal JB, Nilsson D, Peterson J, Salzberg SL, Shallow J, Silva JC, Sundaram J, Westenberger S, White O, Melville SE, Donelson JE, Andersson B, Stuart KD, Hall N (2005) Comparative genomics of trypanosomatid parasitic protozoa. *Science* 309:404–409
- Eyre DW, Cule ML, Wilson DJ, Griffiths D, Vaughan A, O'Connor L, Ip CLC, Golubchik T, Batty EM, Finney JM, Wyllie DH, Didelot X, Piazza P, Bowden R, Dingle KE, Harding RM, Crook DW, Wilcox MH, Peto TEA, Walker AS (2013) Diverse sources of *C. difficile* infection identified on whole-genome sequencing. *New Engl J Med* 369:1195–1205
- Feng Y, Xiao L (2011) Zoonotic potential and molecular epidemiology of *Giardia* species and giardiasis. *Clin Microbiol Rev* 24:110–140
- Fierer N, Breitbart M, Nulton J, Salamon P, Lozupone C, Jones R, Robeson M, Edwards RA, Felts B, Rayhawk S, Knight R, Rohwer F, Jackson RB (2007) Metagenomic and small-subunit rRNA analyses reveal the genetic diversity of bacteria, archaea, fungi, and viruses in soil. *Appl Environ Microbiol* 73:7059–7066
- Fisher MC, Garner TW, Walker SF (2009) Global emergence of *Batrachochytrium dendrobatidis* and amphibian chytridiomycosis in space, time, and host. *Annu Rev Microbiol* 63:291–310
- Foster JT, Beckstrom-Sternberg SM, Pearson T, Beckstrom-Sternberg JS, Chain PSG, Roberto FF, Hnath J, Brettin T, Keim P (2009) Whole-genome-based phylogeny and divergence of the genus *Brucella*. *J Bacteriol* 191:2864–2870
- Frothingham R (1995) Differentiation of strains in *Mycobacterium tuberculosis* complex by DNA sequence polymorphisms, including rapid identification of *M. bovis* BCG. *J Clin Microbiol* 33:840–844
- Gardy JL, Johnston JC, Sui SJH, Cook VJ, Shah L, Brodtkin E, Rempel S, Moore R, Zhao Y, Holt R, Varhol R, Birol I, Lem M, Sharma MK, Elwood K, Jones SJM, Brinkman FSL, Brunham RC, Tang P (2011) Whole-genome sequencing and social-network analysis of a tuberculosis outbreak. *New Engl J Med* 364:730–739
- Geue L, Schares S, Schnick C, Kliemt J, Beckert A, Freuling C, Conraths FJ, Hoffmann B, Zanoni R, Marston D, McElhinney L, Johnson N, Fooks AR, Tordo N, Müller T (2008) Genetic characterisation of attenuated SAD rabies virus strains used for oral vaccination of wildlife. *Vaccine* 26:3227–3235
- Goodchild AV, Watkins GH, Sayers AR, Jones JR, Clifton-Hadley RS (2012) Geographical association between the genotype of bovine tuberculosis in found dead badgers and in cattle herds. *Vet Rec* 170:259
- Gould AR, Kattenbelt JA, Selleck P, Hansson E, Della-Porta A, Westbury HA (2001) Virulent Newcastle disease in Australia: molecular epidemiological analysis of viruses isolated prior to and during the outbreaks of 1998–2000. *Virus Res* 77:51–60
- Grad YH, Lipsitch M, Feldgarden M, Arachchi HM, Cerqueira GC, FitzGerald M, Godfrey P, Haas BJ, Murphy CI, Russ C, Sykes S, Walker BJ, Wortman JR, Young S, Zeng Q, Abouelleil A, Bochicchio J, Chauvin S, DeSmet T, Gujja S, McCowan C, Montmayeur A, Steelman S, Frimodt-Møller J, Petersen AM, Struve C, Krogfelt KA, Bingen E, Weill F-X, Lander ES, Nusbaum C, Birren BW, Hung DT, Hanage WP (2012) Genomic epidemiology of the *Escherichia coli* O104:H4 outbreaks in Europe, 2011. *Proc Natl Acad Sci* 109:3065–3070
- Grenfell BT, Pybus OG, Gog JR, Wood JLN, Daly JM, Mumford JA, Holmes EC (2004) Unifying the epidemiological and evolutionary dynamics of pathogens. *Science* 303:327–332
- Griffin AS, West SA, Buckling A (2004) Cooperation and competition in pathogenic bacteria. *Nature* 430:1024–1027
- Hacker J, Carniel E (2001) Ecological fitness, genomic islands and bacterial pathogenicity. *EMBO Rep* 2:376–381
- Hamede RK, Bashford J, McCallum H, Jones M (2009) Contact networks in a wild Tasmanian devil (*Sarcophilus harrisi*) population: using social network analysis to reveal seasonal variability in social behaviour and its implications for transmission of devil facial tumour disease. *Ecol Lett* 12:1147–1157
- Hamede R, Lachish S, Belov K, Woods G, Kreiss A, Pearse AM, Lazenby B, Jones M, McCallum H (2012) Reduced effect of Tasmanian devil facial tumor disease at the disease front. *Conserv Biol* 26:124–134
- Harris SR, Cartwright EJP, Török ME, Holden MTG, Brown NM, Ogilvy-Stuart AL, Ellington MJ, Quail MA, Bentley SD, Parkhill J, Peacock SJ (2013) Whole-genome sequencing for analysis of an outbreak of methicillin-resistant *Staphylococcus aureus*: a descriptive study. *Lancet Infect Dis* 13:130–136
- Harwood VJ, Whitlock J, Withington V (2000) Classification of antibiotic resistance patterns of indicator bacteria by discriminant analysis: use in predicting the source of fecal contamination in subtropical waters. *Appl Environ Microbiol* 66:3698–3704
- Haydon DT, Cleaveland S, Taylor LH, Laurenson MK (2002) Identifying reservoirs of infection: a conceptual and practical challenge. *Emerg Infect Dis* 8:1468–1473
- Hemelaar J, Gouws E, Ghys PD, Osmanov S (2011) Global trends in molecular epidemiology of HIV-1 during 2000–2007. *AIDS* 25: 679–689

- Henderson M, Okamura B (2004) The phylogeography of salmonid proliferative kidney disease in Europe and North America. *Proc Roy Soc Biol Sci* 271:1729–1736
- Henn MR, Boutwell CL, Charlebois P, Lennon NJ, Power KA, Macalalad AR, Berlin AM, Malboeuf CM, Ryan EM, Gnerre S (2012) Whole genome deep sequencing of HIV-1 reveals the impact of early minor variants upon immune recognition during acute infection. *PLoS Pathog* 8:e1002529
- Hernandez SM, Keel K, Sanchez S, Trees E, Gerner-Smidt P, Adams JK, Cheng Y, Ray A III, Martin G, Presotto A, Ruder MG, Brown J, Blehert DS, Cottrell W, Maurer JJ (2012) Epidemiology of a *Salmonella enterica subsp enterica* serovar typhimurium strain associated with a songbird outbreak. *Appl Environ Microbiol* 78:7290–7298
- Hoffmann C, Minkah N, Leipzig J, Wang G, Arens MQ, Tebas P, Bushman FD (2007) DNA bar coding and pyrosequencing to identify rare HIV drug resistance mutations. *Nucleic Acids Res* 35:e91
- Holman PJ, Carroll JE, Pugh R, Davis DS (2011) Molecular detection of *Babesia bovis* and *Babesia bigemina* in white-tailed deer (*Odocoileus virginianus*) from Tom Green County in central Texas. *Vet Parasitol* 177:298–304
- Holmes EC, Ghedin E, Miller N, Taylor J, Bao Y, St George K, Grenfell BT, Salzberg SL, Fraser CM, Lipman DJ (2005) Whole-genome analysis of human influenza A virus reveals multiple persistent lineages and reassortment among recent H3N2 viruses. *PLoS Biol* 3:e300
- Hostnik P, Picard-Meyer E, Rihtaric D, Toplak I, Cliquet F (2014) Vaccine-induced rabies in a red fox (*Vulpes vulpes*): isolation of vaccine virus in brain tissue and salivary glands. *J Wildl Dis* 50(2):397–401
- Hoye BJ, Munster VJ, Nishiura H, Klaassen M, Fouchier RAM (2010) Surveillance of wild birds for avian influenza virus. *Emerg Infect Dis* 16:1827–1834
- Hu B, Xie G, Lo C-C, Starkenburg SR, Chain PSG (2011) Pathogen comparative genomics in the next-generation sequencing era: genome alignments, pangenomics and metagenomics. *Brief Funct Genomics* 10:322–333
- Hudson PJ, Rizzoli A, Grenfell BT, Heesterbeek J, Dobson AP (2002) Ecology of wildlife diseases. Oxford University Press
- Hyndman RJ, Khandakar Y (2007) Automatic time series for forecasting: the forecast package for R. (No. 6/07). Monash University, Department of Econometrics and Business Statistics
- Ip HS, Wiley MR, Long R, Palacios G, Shearn-Bochsler V, Whitehouse CA (2014) Identification and characterization of Highlands J virus from a Mississippi sandhill crane using unbiased next-generation sequencing. *J Virol Methods* 206:42–45
- James TY, Litvintseva AP, Vilgalys R, Morgan JAT, Taylor JW, Fisher MC, Berger L, Weldon C, du Preez L, Longcore JE (2009) Rapid global expansion of the fungal disease chytridiomycosis into declining and healthy amphibian populations. *PLoS Pathog* 5:e1000458
- Jenkins HE, Woodroffe R, Donnelly CA, Cox DR, Johnston WT, Bourne FJ, Cheeseman CL, Clifton-Hadley RS, Gettinby G, Gilks P, Hewinson RG, McNemey JP, Morrison WI (2007) Effects of culling on spatial associations of *Mycobacterium bovis* infections in badgers and cattle. *J Appl Ecol* 44:897–908
- Jombart T, Cori A, Didelot X, Cauchemez S, Fraser C, Ferguson N (2014) Bayesian reconstruction of disease outbreaks by combining epidemiologic and genomic data. *PLoS Comput Biol* 10:e1003457
- Joshi D, Harris NB, Waters R, Thacker T, Mathema B, Krieswirth B, Sreevatsan S (2012) Single nucleotide polymorphisms in the *Mycobacterium bovis* genome resolve phylogenetic relationships. *J Clin Microbiol* 50:3853–3861
- Kao RR, Haydon DT, Lycett SJ, Murcia PR (2014) Supersize me: how whole-genome sequencing and big data are transforming epidemiology. *Trends Microbiol* 22:282–291
- Keim P, Smith KL, Keys C, Takahashi H, Kurata T, Kaufmann A (2001) Molecular investigation of the Aum Shinrikyo anthrax release in Kameido, Japan. *J Clin Microbiol* 39:4566–4567
- Köndgen S, Kühl H, N’Goran PK, Walsh PD, Schenk S, Ernst N, Biek R, Formenty P, Mätz-Rensing K, Schweiger B, Junglen S, Ellerbrok H, Nitsche A, Briese T, Lipkin WI, Pauli G, Boesch C, Leendertz FH (2008) Pandemic human viruses cause decline of endangered great apes. *Curr Biol* 18:260–264
- Koopmans M, Vinjé J, de Wit M, Leenen I, van der Poel W, van Duynhoven Y (2000) Molecular epidemiology of human enteric caliciviruses in The Netherlands. *J Infect Dis* 181:S262–S269
- Köser CU, Ellington MJ, Cartwright EJP, Gillespie SH, Brown NM, Farrington M, Holden MTG, Dougan G, Bentley SD, Parkhill J, Peacock SJ (2012) Routine use of microbial whole genome sequencing in diagnostic and public health microbiology. *PLoS Pathog* 8:e1002824
- Kumarasamy KK, Toleman MA, Walsh TR, Bagaria J, Butt F, Balakrishnan R, Chaudhary U, Doumith M, Giske CG, Irfan S, Krishnan P, Kumar AV, Maharjan S, Mushtaq S, Noorie T, Paterson DL, Pearson A, Perry C, Pike R, Rao B, Ray U, Sarma JB, Sharma M, Sheridan E, Thirunarayan MA, Turton J, Upadhyay S, Warner M, Welfare W, Livermore DM, Woodford N (2010) Emergence of a new antibiotic resistance mechanism in India, Pakistan, and the UK: a molecular, biological, and epidemiological study. *Lancet Infect Dis* 10:597–602
- Liu C, Graber CJ, Karr M, Diep BA, Basuino L, Schwartz BS, Enright MC, O’Hanlon SJ, Thomas JC, Perdreau-Remington F, Gordon S, Gunthorpe H, Jacobs R, Jensen P, Leoung G, Rumack JS, Chambers HF (2008) A population-based study of the incidence and molecular epidemiology of methicillin-resistant *Staphylococcus aureus* disease in San Francisco, 2004–2005. *Clin Infect Dis* 46:1637–1646
- Lloyd-Smith JO, Cross PC, Briggs CJ, Daugherty M, Getz WM, Latta J, Sanchez MS, Smith AB, Swee A (2005) Should we expect population thresholds for wildlife disease? *Trends Ecol Evol* 20:511–519
- Maan S, Samuel AR, Maan NS, Attoui H, Rao S, Mertens PP (2004) Molecular epidemiology of bluetongue viruses from disease outbreaks in the Mediterranean Basin. *Vet Ital* 40:489–496
- Maan S, Belaganahalli M, Maan N, Mertens PC (2013) Whole genome sequencing strategies and development of orbivirus sequence database: implications for novel dsRNA virus detection. In: Salar RK, Gahlawat SK, Siwach P, Duhan JS (eds) *Biotechnology: prospects and applications*. Springer, India, pp 237–255
- Medini D, Serruto D, Parkhill J, Relman DA, Donati C, Moxon R, Falkow S, Rappuoli R (2008) Microbiology in the post-genomic era. *Nat Rev Microbiol* 6:419–430
- Meli ML, Cattori V, Martínez F, López G, Vargas A, Simón MA, Zorrilla I, Muñoz A, Palomares F, López-Bao JV, Pastor J, Tandon R, Willi B, Hofmann-Lehmann R, Lutz H (2009) Feline leukemia virus and other pathogens as important threats to the survival of the critically endangered Iberian Lynx (*Lynx pardinus*). *PLoS One* 4:e4744
- Metzker ML (2010) Sequencing technologies -the next generation. *Nat Rev Genet* 11:31–46
- Millan J, Blasco-Costa I (2012) Molecular evidence of shared hookworm *Ancylostoma tubaeforme* haplotypes between the critically endangered Iberian lynx and sympatric domestic cats. *Vet Parasitol* 186:518–522
- Monot M, Honore N, Garnier T, Zidane N, Sherafi D, Paniz-Mondolfi A, Matsuoka M, Taylor GM, Donoghue HD, Bouwman A, Mays S, Watson C, Lockwood D, Khamispour A, Dowlati Y, Jianping S, Rea TH, Vera-Cabrera L, Stefani MM, Banu S, Macdonald M, Sapkota BR, Spencer JS, Thomas J, Harshman K, Singh P, Busso P, Gattiker A, Rougemont J, Brennan PJ, Cole ST (2009) Comparative genomic and phylogeographic analysis of *Mycobacterium leprae*. *Nat Genet* 41:1282–1289
- Murchison EP, Schulz-Trieglaff OB, Ning Z, Alexandrov LB, Bauer MJ, Fu B, Hims M, Ding Z, Ivakhno S, Stewart C, Ng BL, Wong W,

- Aken B, White S, Alsop A, Becq J, Bignell GR, Cheetham RK, Cheng W, Connor TR, Cox AJ, Feng ZP, Gu Y, Grocock RJ, Harris SR, Khrebtkova I, Kingsbury Z, Kowarsky M, Kreiss A, Luo S, Marshall J, McBride DJ, Murray L, Pearse AM, Raine K, Rasolonjatovo I, Shaw R, Tedder P, Tregidgo C, Vilella AJ, Wedge DC, Woods GM, Gormley N, Humphray S, Schroth G, Smith G, Hall K, Searle SM, Carter NP, Papenfuss AT, Futreal PA, Campbell PJ, Yang F, Bentley DR, Evers DJ, Stratton MR (2012) Genome sequencing and analysis of the Tasmanian devil and its transmissible cancer. *Cell* 148:780–791
- Neel C (2010) Molecular epidemiology of simian immunodeficiency virus infection in wild-living gorillas. *J Virol* 84:1464–1476
- Nielsen L, Arctander P, Jensen TH, Dietz H-H, Hammer AS, Banyard AC, Barrett T, Blixenkrone-Moller M (2009) Genetic diversity and phylogenetic analysis of the attachment glycoprotein of phocine distemper viruses of the 2002 and 1988 epizootics. *Virus Res* 144:323–328
- Nunes MR, Faria NR, Vasconcelos HB, Medeiros DB, Silva de Lima CP, Carvalho VL, Pinto da Silva EV, Cardoso JF, Sousa EC, Nunes KN, Rodrigues SG, Abecasis AB, Suchard MA, Lemey P, Vasconcelos PF (2012) Phylogeography of dengue virus serotype 4, Brazil, 2010–2011. *Emerg Infect Dis* 18:1858–1864
- Paillet R, Darby AC, Robinson C, Wright NL, Steward KF, Anderson E, Webb K, Holden MT, Efstratiou A, Broughton K, Jolley KA, Priestnall SL, Marotti Campi MC, Hughes MA, Radford A, Eyles K, Waller AS (2010) Identification of three novel superantigen-encoding genes in *Streptococcus equi subsp. zooepidemicus*, *szef*, *szen*, and *szep*. *Infect Immun* 78:4817–4827
- Pandit A, Sinha S (2010) Using genomic signatures for HIV-1 sub-typing. *BMC Bioinform* 11:S26
- Perez de Rosas AR, Segura EL, Garcia BA (2011) Molecular phylogeography of the Chagas' disease vector *Triatoma infestans* in Argentina. *Heredity* 107:71–79
- Pop M, Salzberg SL (2008) Bioinformatics challenges of new sequencing technology. *Trends Genet* 24:142–149
- Price LB, Stegger M, Hasman H, Aziz M, Larsen J, Andersen PS, Pearson T, Waters AE, Foster JT, Schupp J, Gillece J, Driebe E, Liu CM, Springer B, Zdovc I, Battisti A, Franco A, Żmudzki J, Schwarz S, Butaye P, Jouy E, Pomba C, Porrero MC, Ruimy R, Smith TC, Robinson DA, Weese JS, Arriola CS, Yu F, Laurent F, Keim P, Skov R, Aarestrup FM (2012) *Staphylococcus aureus* CC398: host adaptation and emergence of methicillin resistance in livestock. *MBio* 3(1):e00305–e00311
- Rankin DJ, Bargum K, Kokko H (2007) The tragedy of the commons in evolutionary biology. *Trends Ecol Evol* 22:643–651
- Rasko DA, Worsham PL, Abshire TG, Stanley ST, Bannan JD, Wilson MR, Langham RJ, Decker RS, Jiang L, Read TD, Phillippy AM, Salzberg SL, Pop M, Van Ert MN, Kenefic LJ, Keim PS, Fraser-Liggett CM, Ravel J (2011) *Bacillus anthracis* comparative genome analysis in support of the Amerithrax investigation. *Proc Natl Acad Sci U S A* 108:5027–5032
- Recker M, Pybus OG, Nee S, Gupta S (2007) The generation of influenza outbreaks by a network of host immune responses against a limited set of antigenic types. *Proc Natl Acad Sci* 104:7711–7716
- Relman DA (2011) Microbial genomics and infectious diseases. *N Engl J Med* 365:347–357
- Roetzer A, Diel R, Kohl TA, Rückert C, Nübel U, Blom J, Wirth T, Jaenicke S, Schuback S, Rüsche-Gerdes S, Supply P, Kalinowski J, Niemann S (2013) Whole genome sequencing versus traditional genotyping for investigation of a *Mycobacterium tuberculosis* outbreak: a longitudinal molecular epidemiological study. *PLoS Med* 10:e1001387
- Romero B, Aranaz A, Sandoval A, Alvarez J, de Juan L, Bezos J, Sanchez C, Galka M, Fernandez P, Mateos A, Dominguez L (2008) Persistence and molecular evolution of *Mycobacterium bovis* population from cattle and wildlife in Donana National Park revealed by genotype variation. *Vet Microbiol* 132:87–95
- Sainsbury AW, Nettleton P, Gilray J, Gurnell J (2000) Grey squirrels have high seroprevalence to a parapoxvirus associated with deaths in red squirrels. *Anim Conserv* 3:229–233
- Salzer SJ, Gillespie TR, Rwegu IB, Chapman CA, Goldberg TL (2012) Epidemiology and molecular relationships of cryptosporidium spp. in people, primates, and livestock from Western Uganda. *PLoS Negl Trop Dis* 6(4):e1597
- Santolaya ME, O'Ryan ML, Valenzuela MT, Prado V, Vergara R, Munoz A, Toneatto D, Grana G, Wang H, Clemens R, Dull PM (2012) Immunogenicity and tolerability of a multicomponent meningococcal serogroup B (4CMenB) vaccine in healthy adolescents in Chile: a phase 2b/3 randomised, observer-blind, placebo-controlled study. *Lancet* 379:617–624
- Seib KL, Dougan G, Rappuoli R (2009) The key role of genomics in modern vaccine and drug design for emerging infectious diseases. *PLoS Genet* 5:e1000612
- Serruto D, Serino L, Maignani V, Pizza M (2009) Genome-based approaches to develop vaccines against bacterial pathogens. *Vaccine* 27:3245–3250
- Siddle HV, Kreiss A, Eldridge MDB, Noonan E, Clarke CJ, Pycroft S, Woods GM, Belov K (2007) Transmission of a fatal clonal tumor by biting occurs due to depleted MHC diversity in a threatened carnivorous marsupial. *Proc Natl Acad Sci* 104:16221–16226
- Siddle HV, Marzec J, Cheng Y, Jones M, Belov K (2010) MHC gene copy number variation in Tasmanian devils: implications for the spread of a contagious cancer. *Proc Biol Sci* 277:2001–2006
- Smith NH, Dale J, Inwald J, Palmer S, Gordon SV, Hewinson RG, Smith JM (2003) The population structure of *Mycobacterium bovis* in Great Britain: clonal expansion. *Proc Natl Acad Sci U S A* 100:15271–15275
- Snitkin ES, Zelazny AM, Thomas PJ, Stock F, Henderson DK, Palmore TN, Segre JA (2012) Tracking a hospital outbreak of carbapenem-resistant *Klebsiella pneumoniae* with whole-genome sequencing. *Sci Transl Med* 4:148ra116
- Snow M (2011) The contribution of molecular epidemiology to the understanding and control of viral diseases of salmonid aquaculture. *Vet Res* 42:56
- Spibey N, McCabe VJ, Greenwood NM, Jack SC, Sutton D, van der Waart L (2012) Novel bivalent vectored vaccine for control of myxomatosis and rabbit haemorrhagic disease. *Vet Rec*. doi:10.1136/vr.100366
- Starkman SE, MacDonald DM, Lewis JCM, Holmes EC, Simmonds P (2003) Geographic and species association of hepatitis B virus genotypes in non-human primates. *Virology* 314:381–393
- Streicker DG, Lemey P, Velasco-Villa A, Rupprecht CE (2012) Rates of viral evolution are linked to host geography in bat rabies. *PLoS Pathog* 8:e1002720
- Szentiks CA, Tsangaras K, Abendroth B, Scheuch M, Stenglein MD, Wohlsein P, Heeger F, Höveler R, Chen W, Sun W, Damiani A, Nikolin V, Gruber AD, Grobbel M, Kalthoff D, Höper D, Cziriák GÁ, DeRisi J, Mazzoni CJ, Schüle A, Aue A, East ML, Hofer H, Beer M, Osterrieder N, Greenwood AD (2014) Polar bear encephalitis: establishment of a comprehensive next-generation pathogen analysis pipeline for captive and free-living wildlife. *J Comp Pathol* 150:474–488
- Tettelin H, Riley D, Cattuto C, Medini D (2008) Comparative genomics: the bacterial pan-genome. *Curr Opin Microbiol* 11:472–477
- Tibayrenc M (1998) Beyond strain typing and molecular epidemiology: integrated genetic epidemiology of infectious diseases. *Parasitol Today* 14:323–329
- Tonellato PJ, Crawford JM, Boguski MS, Saffitz JE (2011) A national agenda for the future of pathology in personalized medicine: report of the proceedings of a meeting at the Banbury conference center on genome-era pathology, precision diagnostics, and preemptive care: a stakeholder summit. *Am J Clin Pathol* 135:668–672

- Vineis P, Perera F (2007) Molecular epidemiology and biomarkers in etiologic cancer research: the new in light of the old. *Cancer Epidemiol Biomarkers Prev* 16:1954–1965
- Vogler AJ, Chan F, Wagner DM, Roumagnac P, Lee J, Nera R, Eppinger M, Ravel J, Rahalison L, Rasoamanana BW, Beckstrom-Sternberg SM, Achtman M, Chanteau S, Keim P (2011) Phylogeography and molecular epidemiology of *Yersinia pestis* in Madagascar. *PLoS Negl Trop Dis* 5:e1319
- Walker TM, Ip CLC, Harrell RH, Evans JT, Kapatai G, Dediccoat MJ, Eyre DW, Wilson DJ, Hawkey PM, Crook DW, Parkhill J, Harris D, Walker AS, Bowden R, Monk P, Smith EG, Peto TEA (2013) Whole-genome sequencing to delineate *Mycobacterium tuberculosis* outbreaks: a retrospective observational study. *Lancet Infect Dis* 13:137–146
- Wedlock DN, Denis M, Vordermeier HM, Hewinson RG, Buddle BM (2007) Vaccination of cattle with Danish and Pasteur strains of *Mycobacterium bovis* BCG induce different levels of IFN $\gamma$  post-vaccination, but induce similar levels of protection against bovine tuberculosis. *Vet Immunol Immunopathol* 118:50–58
- Williams ES, Miller MW, Kreeger TJ, Kahn RH, Thorne ET (2002) Chronic wasting disease of Deer and Elk: a review with recommendations for management. *J Wildlife Manag* 66:551–563
- Woodroffe R, Donnelly CA, Johnston WT, Bourne FJ, Cheeseman CL, Clifton-Hadley RS, Cox DR, Gettinby G, Hewinson RG, Le Fevre AM, McInerney JP, Morrison WI (2005) Spatial association of *Mycobacterium bovis* infection in cattle and badgers (*Meles meles*). *J Appl Ecol* 42:852–862
- Woodroffe R, Prager KC, Munson L, Conrad PA, Dubovi EJ, Mazet JAK (2012) Contact with domestic dogs increases pathogen exposure in endangered African wild dogs (*Lycan pictus*). *PLoS One* 7:e30099
- Woolhouse ME, Dye C, Etard J-F, Smith T, Charlwood J, Gamett G, Hagan P, Hii J, Ndhlovu P, Quinnell R (1997) Heterogeneities in the transmission of infectious agents: implications for the design of control programs. *Proc Natl Acad Sci* 94:338–342
- Wylie JL, Cabral T, Jolly AM (2005) Identification of networks of sexually transmitted infection: a molecular, geographic, and social network analysis. *J Infect Dis* 191:899–906
- Xiao L, Ryan UM (2004) Cryptosporidiosis: an update in molecular epidemiology. *Curr Opin Infect Dis* 17:483–490
- Yoshimura A, Koketsu M, Bando H, Saiki E, Suzuki M, Watanabe Y, Kanuka H, Fukumoto S (2014) Phylogenetic comparison of avian Haemosporidian parasites from resident and migratory birds in northern Japan. *J Wildl Dis* 50(2):235–242