Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Genetics & Development

Genomic changes during evolution of animal parasitism in eukaryotes

Amber Leckenby and Neil Hall

Understanding how pathogens have evolved to survive in close association with their hosts is an important step in unraveling the biology of host–pathogen interactions. Comparative genomics is a powerful tool to approach this problem as an increasing number of genomes of multiple pathogen species and strains become available. The ever-growing catalog of genome sequences makes comparison of organisms easier, but it also allows us to reconstitute the evolutionary processes occurring at the genomic level that may have led to the acquisition of pathogenic or parasitic mechanisms.

Address
Department of Functional and Comparative Genomics, The University of Liverpool, Biosciences Building, Crown Street, Liverpool L69 7ZB, UK

Corresponding author: Hall, Neil (neilhall@liv.ac.uk)

Current Opinion in Genetics & Development 2015, **35**:86–92

This review comes from a themed issue on **Genomes and evolution**

Edited by **Antonis Rokas** and **Pamela S Soltis**

http://dx.doi.org/10.1016/j.gde.2015.11.001

0959-437/Published by Elsevier Ltd.

This review will focus on comparisons between genomes of pathogenic and free-living organisms to determine what sets apart these pathogenic species from their non-pathogenic relatives. We explore the evolutionary changes that led to the emergence of non-pathogenic organisms to a pathogenic lifestyle. We also emphasise how comparative genomic approaches have led to important discoveries and demonstrate how comparisons within a genus or species can help to elucidate how pathogens evolve genus or species-specific phenotypes that allow them to adapt to a new host [1]. We have summarised the changes that occur in adaptation to parasitism into three processes: gene family expansion, gene loss, and genome rearrangement.

Pathogenic organisms are often seen as distinct derivatives away from a free-living lifestyle. Though crude, this classification helps to elucidate the genetic changes that have occurred during the evolutionary pathway to parasitism or pathogenicity. In some prokaryotes it appears that the ability to infect a host organism can be acquired readily through the horizontal transfer of genes that enable the organism to enter a new niche [2–5]. Hence, a pathogenic strain may be very closely related to a non-pathogenic strain of the same species. In eukaryotic organisms this mechanism of evolution is rare due to the physical barriers such as the nuclear membrane and more complex genetic organization. Although examples of horizontal gene transfer do exist [6,7,8••], the frequency of these events is significantly lower in eukaryotic lineages in comparison to prokaryotes [9,10]. Hence, we can assume that the majority of pathogenicity in eukaryotes has evolved over a long period of coexistence with a host species, and genetic innovation within a species will predominantly occur through neo-functionalization, subfunctionalization or specialization following the duplication of existing genes.

In the process of neo-functionalization, one paralog retains ancestral function while the other acquires a distinct novel function [11]. Under subfunctionalization, degenerative mutations affect different functions in both paralogs resulting in retention of both copies in the genome in order to preserve ancestral functions [12,13], while specialization refers to the process whereby subfunctionalization and neo-functionalization occur together producing two functionally distinct paralogs [14]. Examples of these processes have demonstrated in experiments across a range of model eukaryotes [14–18], and it has been argued that subfunctionalization is merely a transition state on the evolutionary pathway of neo-functionalization [19,20].

Despite the limitations on eukaryotes for gene acquisition, pathogenicity is a very common trait and the sheer number of pathogenic or parasitic species is potentially huge. Poulin and Morand have noted linear or exponential discovery rates of new pathogenic helminth species despite the fact that parts of the globe are sampled very sparsely [21]. It is postulated that parasitic nematodes have independently evolved to parasitize animals many times [22,23]. Hence, the transition from a free-living lifestyle to a pathogenic one has occurred independently many times across a plethora of groups of lineages.

The multitude of genome sequences generated recently, due to the advent of next generation sequencing technologies, has sparked a revolution in the study of pathogen evolution [24,25]. The vast increase in genomic information has improved phylogenies of previous poorly resolved species and is allowing more and more species to be positioned phylogenetically. For example, the use of

mitochondrial genomes in phylogenetic analyses has been used many times [26–32]. Increased sampling has enabled deep phylogenetic comparisons to be carried out between groups of organisms that may not be obviously related from morphological observations. The sequencing of plastid genomes has identified Apicomplexans and Dinoflagellates as sister clades [33] and it has been confirmed that chromerids and colpodellids form a single monophyletic sister group to the Apicomplexans [34,35].

Gene family expansion

Gene family expansions have long been known to be an important factor in parasite lifestyle adaptations to host-specific environments and are heavily documented for Apicomplexan species such as *Plasmodium*, *Toxoplasma* and *Cryptosporidium* [36,37]. Observations of gene family expansions are also seen across parasitic helminth species [38••,39••], in *Giardia* [40–42] and Trypanosomatids [42–45]. The expanded gene families often occur at the telomeres and recombinogenic parts of the genome as these regions help facilitate genome rearrangements allowing for the expansion of gene families [46•,47,48]. Gene family expansions can also occur in non-telomeric and non-subtelomeric regions such as phylum specific expansion of the Apicomplexan ApiAP2 gene family, whose members are associated with the regulation of virulence associated subtelomeric gene families [49–51].

The genomes of parasitic nematodes have been studied using within-species comparative genomics [52] and between parasitic species and the close free-living relative, *Caenorhabditis elegans* [39••,53]. The parasitic nematode, *Brugia malayi*, was the first parasite to be sequenced that has a model organism, *C. elegans*, as a close relative [38••]. The original sequencing effort inferred that the clade III nematode *B. malayi* genome contains between 14,500 and 17,800 genes, fewer than that of its clade V relative *C. elegans* [38••]. Of these genes, 20% were predicted as being *B. malayi* specific. The study revealed unexpected diversity of the ALT (abundant larval transcript) family; 13 ALT genes showed diversity in the *B. malayi* genome and are implicated as virulence factors due to their capacity to modulate macrophage function [54]. The ALT family is one of few gene families that are exclusively expanded in *B. malayi* when compared to *C. elegans*, where one member of the ALT gene family is reported. These observations demonstrate how gene expansion can contribute to *B. malayi* pathogenic processes [38••].

Gene family expansion is able to generate genetic variation as redundancy enables genes to diverge without the deleterious effects that are associated with mutations in single copy genes. Large gene families are therefore often associated with surface antigens [43,55–60] or other aspects of host interaction such as cell invasion [61,62••,63], as these genes are often under strong positive selection from the host immune system and cell surface receptors. This can be argued as an example of neo-functionalization whereby a duplicated copy of a gene experiences reduced purifying selection and ultimately evolves a new function (albeit similar to the original copy) [12,64]. Alternatively, this may be considered a case of subfunctionalization as duplication event allows for differential evolution of both genes enabling further optimization of the original function of the gene [11,64].

Sun *et al.* have observed that the majority of recently duplicated genes in *Giardia lamblia* are VSP (Variable Surface Protein) genes, antigen genes involved in evasion of the host immune system [42,65]. Their study revealed that VSP genes underwent expansion independently in the very closely related parasite *Spironucleus salmonicida*. They conclude that the most probable explanation for independent expansion of the *G. lamblia* VSP genes is that there has been strong selection for genes assisting evasion of the host immune systems. The expansion of the VSP genes is estimated to have occurred the same time as species expansion in the mammals, when the ancestor of placental mammals diverged from the platypus. Therefore, positive selection may have occurred as the expanded VSP genes enabled *G. lamblia* to parasitize a wide range of hosts [42,66].

Gene loss

Gene loss is also commonly observed in parasitic species and Nakjang *et al.* have observed varying degrees of genome reduction across Microsporidia genomes [40]. Although studies have observed expansion of NupG transporter proteins which have enabled Microsporidia species to gain essential macromolecules from the host, the metabolic genes responsible for originally producing these have been lost from the genome [62••]. Loss of these metabolic genes contributes to the group's small genome size and bears resemblance to intracellular bacteria. Other similarities were also witnessed between the Microsporidia and intracellular bacteria such as low coding capacity and a high AT-content, which combined with the small genome sizes witnessed, could lend themselves to the process of reductive evolution mirrored across other eukaryotic parasites [8••,67–74] elsewhere [75]. Nakjang *et al.* suggest that it is not entirely unexpected that the Microsporidia may have been subject to a Muller's ratchet process [76•], similar to that witnessed in intracellular bacteria [77•] that also reproduce asexually and have small effective population sizes.

Gene loss does not always present itself as the loss of single genes. Microsporidia appear to have encountered dramatic reductive evolution on their mitochondrion which is now present as a highly divergent and reduced organelle termed the mitosome [78•]. Mitosomes have so far been found in a range of lineages including the Amoebozoa, Microsporidia and Apicomplexans [70,78•,79,80].