



Review

Symbiont genomics, our new tangled bank

M. Medina ^{a,*}, J.L. Sachs ^b^a School of Natural Sciences, University of California, Merced, 5200 N. Lake Road, Merced, CA 95343, USA^b Department of Biology, University of California, Riverside, 1208 Spieth Hall, Riverside, CA 92521, USA

ARTICLE INFO

Article history:

Received 16 September 2009

Accepted 25 December 2009

Available online 4 January 2010

Keywords:

Genomics
Gut flora
Holobiont
Interactome
Metagenomics
Microbiome
Mutualism
Symbiosis
Parasitism
Tree of life

ABSTRACT

Microbial symbionts inhabit the soma and surfaces of most multicellular species and instigate both beneficial and harmful infections. Despite their ubiquity, we are only beginning to resolve major patterns of symbiont ecology and evolution. Here, we summarize the history, current progress, and projected future of the study of microbial symbiont evolution throughout the tree of life. We focus on the recent surge of data that whole-genome sequencing has introduced into the field, in particular the links that are now being made between symbiotic lifestyle and molecular evolution. Post-genomic and systems biology approaches are also emerging as powerful techniques to investigate host–microbe interactions, both at the molecular level of the species interface and at the global scale. In parallel, next-generation sequencing technologies are allowing new questions to be addressed by providing access to population genomic data, as well as the much larger genomes of microbial eukaryotic symbionts and hosts. Throughout we describe the questions that these techniques are tackling and we conclude by listing a series of unanswered questions in microbial symbiosis that can potentially be addressed with the new technologies.

© 2010 Elsevier Inc. All rights reserved.

Contents

Introduction	129
Discussion	130
Symbiotic lifestyles and genome structure	130
Common symbiont genes and pathways	133
Interactomes and meta-interactomics	133
From genotype to ecosystems	134
Conclusions	134
Future directions	134
Acknowledgments	135
References	135

“It is interesting to contemplate an entangled bank, clothed with many plants of many kinds, with birds singing on the bushes, with various insects flitting about, and with worms crawling through the damp earth, and to reflect that these elaborately constructed forms, so different from each other, and dependent on each other in so complex a manner, have all been produced by laws acting around us.” Charles Darwin (1859)

Introduction

Symbiotic microbes thrive on the surfaces and in the tissues of diverse multicellular organisms [1–6]. So while Darwin [7] contemplated ‘an entangled bank’ crowded with plants and animals, modern microbiologists can now envisage these creatures as ‘entangled hosts,’ each populated by their own nested microcommunity. Symbionts are defined as organisms that share intimate interactions with other species, whether beneficial or harmful [8,9], and microbial partnerships with animals and plants dominate this classification. Yet what molecular pathways have allowed these symbioses to originate and diversify? How are the microbial partners or their biochemical functions distributed in space and time and among host phylogenies?

* Corresponding author. Fax: +1 209 228 4053.

E-mail address: mmedina@ucmerced.edu (M. Medina).

Is there a ‘minimal molecular toolbox’ necessary for host association, does this toolbox vary between beneficial symbionts and pathogens, and what are the evolutionary trends this toolbox follows once symbiosis is established? These fundamental questions are important to applied research and basic science as well. The medical and economic relevance of beneficial symbiotic microbes is recently becoming evident; these microbial communities are thought to be critical to human health [10,11] and can be key determinants of crop [12] and livestock production [13]. At a broader level, both beneficial and harmful symbionts can drive the ecology and evolution of metazoan and plant hosts [14] [15], and their in-depth study is revealing a level of complexity and biological diversity that had been previously unexplored.

It might seem like an impossible challenge to make generalizations about the vast diversity of microbial symbionts. One hurdle is a dizzying phylogenetic breadth; beneficial microbial symbionts alone encompass lineages spread throughout the tree of life (see Fig. 1). Microbial partners also instigate a wide array of fitness effects on their multicellular hosts, ranging from life-sustaining nutrient provisioning [16] to deadly pathogenesis [17]. There is also variation in the degree of ‘partner fidelity’ between symbiont and host partners [18,19]: while some microbes evolve to become locked in-step within host lineages – like bacteria-derived organelles [20] and obligate endosymbionts [21] that have lost their independence – other microbes infect or inhabit hosts transiently or with little apparent specialization to host genotype [22,23]. For decades, biologists have focused on resolving the particular roles and complex biology of individual symbiont taxa [24]. More recently, the advent of whole-genome sequencing is inspiring a broader, systems view of microbial biology [25]. This approach of analyzing and comparing whole-genome data sets promises to provide new insight into the origins, diversity, and phylogenetic distribution of microbial symbionts and their host-associated functions.

The study of host-associated microbes likely began with Van Leeuwenhoek’s first microscopic observations of *Selenomonas* spp. (bacteria) from the human mouth in 1676 [26]. The development of *in vitro* microbial culturing blossomed during the early 20th century and

introduced key experimental approaches that complemented microscopic techniques. For the first time, microbes could be extracted from hosts, grown clonally, and experimentally infected into new hosts to test their effects. Yet, while microscopy and culturing both remain powerful methods in microbiology, both have limitations. Microscope-based taxonomy has often failed to uncover cryptic diversity in the taxa under study, even classifying diverse lineages as single species [6]. While novel culture protocols are still being actively developed, roughly 90–99% of microbes remain uncultured [27], and it is likely that most microbes will remain so. Molecular approaches of the last two decades have offered powerful alternatives for microbial identification initially revolutionized by 16s rDNA sequencing, e.g., [28,29], but also complemented by other techniques such as fluorescence in situ hybridization (FISH), restriction fragment length polymorphisms (RFLPs), and denaturing gradient gel electrophoresis (DGGE) [30]. More recently, direct sequencing of environmental samples has revealed additional unforeseen microbial diversity [31]. Here, we review key aspects of this recent progress, especially the surge of information being generated by whole-genome sequencing data. While most eukaryotic organisms likely host a complex community of symbionts, only a few such systems have been studied extensively and even fewer have genomic resources currently available (Fig. 1). We focus largely on beneficial microbes, with the understanding that the line between beneficial and harmful microbial host–associates can often be blurred by context dependence [32] and rapid evolution [33,34]. We begin by highlighting broad patterns that have emerged to date, and subsequently, we move beyond the current research frontier to suggest how newly emerging data sets might be analyzed.

Discussion

Symbiotic lifestyles and genome structure

The small genome size of many bacterial endosymbionts (that live within the bodies or cells of hosts) has facilitated a better

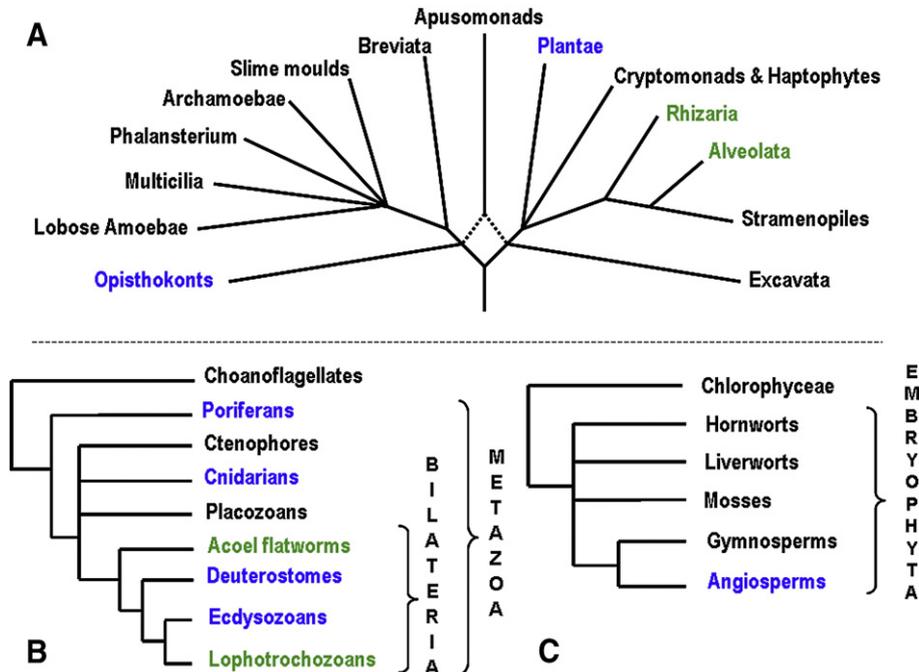


Fig. 1. Phylogenetic distribution of microbial symbionts on the tree of life with genomic resources. (A) Main eukaryotic lineages and color-coded are those with microbial symbioses that are currently being investigated at the genomic level. (B and C) Major relationships within animals and plants. Green represents those lineages known to have symbioses with unicellular eukaryotes. Blue represents lineages with bacterial and eukaryotic symbiont genomic resources. Most eukaryotic symbioses have available only EST resources as opposed to whole-genome sequences (with the exception of the *Laccaria* fungal symbiont of Poplar trees). The bottom panels depict the animal and plant lineages that have microbial symbioses for which there are genomic resources. Phylogenetic trees from Roger and Simpson [93] and Medina [25].

understanding of their molecular evolution given the accessibility to finished genomes from closely related lineages and populations [35–38]. Further, the ability to rapidly shotgun sequence microbial genomes has revealed some fascinating correlates between symbiotic life history and genome evolution. For instance, comparative genomics of obligate endosymbiotic bacteria (that only reproduce within hosts) with free-living relatives have revealed that obligate symbionts tend to evolve reduced genomes, often mediated by gene loss and elevated mutation rates [39–41]. Data from some focused studies of obligate bacterial symbionts are now producing a more detailed view of such reductive genome evolution, as short-read sequencing technologies are facilitating resequencing efforts of multiple bacterial strains. For instance, Moran et al. [37] sequenced seven strains of *Buchnera aphidicola*, the pea aphid endosymbiont. The samples were chosen to be representative of recent colonization events in North America and therefore are of recent divergence (between 135 years ago and 7.5 years ago). A symbiotic–genome erosion model emerged from this study that suggests genome reduction follows a stepwise process. The initial step involves a shift toward higher A + T nucleotide composition, which in turn leads to an excess of homopolymers. These homopolymers have a higher incidence of small indels that finally lead to gene inactivation, and the pseudogenes appear to be removed by larger deletions. This pattern of reductive evolution is likely driven by a combination of selection for genome streamlining (loss of functions that are useless within the rich confines of the host) and genetic drift caused by the small population sizes within hosts [42]. Unlike symbionts that also cycle through environmental stages, obligate endosymbionts have little access to foreign DNA via horizontal gene transfer. Moreover, even if they did have access to foreign genetic material, they often lack genes that allow them to take up and incorporate DNA [4].

In contrast to these obligate interactions, facultative symbionts – that have free-living stages – can often evolve expanded genomes, mostly via gene duplication and horizontal gene transfer [42,43]. Rhizobial bacteria offer some extreme examples. These nitrogen-fixing, root-nodule symbionts of legumes exhibit multiple lineages that have undergone massive genome expansion compared to related taxa [43]. Rhizobial lineages have expanded their genomes both by acquiring many genes via horizontal gene transfer [44,45] and also by duplication and divergence of ancestral loci [46]. These observations still leave the open question of what aspects of rhizobial ecology are responsible for instigating the genome expansion. In the eukaryotic realm, we also observe facultative endosymbionts with massive genomes and evidence of expansion. One striking example is that of the photosynthetic dinoflagellates in the genus *Symbiodinium*, which are facultative symbionts of multiple marine animal taxa as well as other marine eukaryotes. The genome size range among examined *Symbiodinium* species ranges from 1.9 to 4.1 Gb [47]. Another example is the recently sequenced *Laccaria bicolor*, a facultative fungal symbiont of poplar trees. *Laccaria* has emerged as the largest basidiomycete genome sequenced (60 Mb) to date with more genes and larger gene families than other published basidiomycete genomes [48,49]. The high number of fungal genomes sequenced relative to other eukaryotic groups has allowed the identification of common genetic features associated with host infection. Among fungal phytopathogens but not other fungi, certain gene families have undergone marked expansion, thus shedding light on the molecular toolbox necessary for plant infection [50]. As whole-genome data for more eukaryotes becomes accessible, several genomic traits unique to eukaryotes (e.g., alternative splicing or transposable element driven chromosomal rearrangements) can be examined for their roles in the evolution of symbiosis related genes.

Symbiont genome studies have also expanded to consider the effects of host ecology and evolution on microbial partners. Among facultative symbionts such as nitrogen-fixing Actinobacteria in the genus *Frankia*, symbiont genome contraction is correlated with host

plant isolation, whereas symbiont genome expansion is favored during host plant diversification [51]. The evolution of genome reduction might be driven by genetic drift linked to small host plant population sizes, which could favor gene loss in the symbionts [39]. On the other hand, genome expansion is likely driven by selection on the symbiont to exploit a diversity of encountered substrates provided by multiple potential hosts and by their diverse soils [51]. Yet, different symbiont life histories do not always predict divergent genome evolution. A fascinating contrast (that we discuss further below) has been the multiple similarities found between the genomes of beneficial infectious bacteria and pathogenic ones. The list of similarities includes frequent gene rearrangements, lateral gene transfer from distant donors, persistence of mobile elements, and shared genes involved in similar infection mechanisms [42]; see below. One explanatory but untested hypothesis is that the degree to which a bacterium reproduces within versus among hosts is a key determinant of genomic evolution irrespective of the symbiont's fitness effect on any particular host.

With new whole bacterial genome sequences becoming available on a regular basis, it is now possible to dissect the specific evolutionary forces (e.g., duplication, accumulation of selfish elements, horizontal gene transfer, deletion driven by drift or selection, etc.) that shape symbiont genomic structure dependent on lifestyle (e.g., obligate vs facultative symbionts, intracellular vs extra cellular symbionts). Currently, only horizontal transfer is known to rapidly engender bacteria with novel host-associated functions such as symbiotic efficacy [52] or pathogenesis [53]. Given the major role horizontal gene transfer has over evolutionary time in shaping the diversity of bacterial gene content relative to other forces such as gene duplication [54], a clear next step for biologists is to resolve a much fuller breadth of selected functions driven by the integration of novel loci [35]. While horizontal gene transfer can give symbiotic bacteria a massive fitness advantage by allowing them to invade new hosts and host habitats [52], this process must occur in the face of the strong deletional bias that most bacteria exhibit [55]. Some horizontally acquired, host-associated functions have been well characterized including pathogenicity and antibiotic resistance, but considering the high frequency of transferred loci of no known function [35], our understanding of the ecological and evolutionary role of horizontal gene transfer remains highly fragmentary. Finally, there are now a few examples of horizontal gene transfer from bacterial symbiont to eukaryotic hosts in which genes are expressed and functional in the host [56–58]. As more eukaryotic genomes become sequenced, it will

Table 1
Homologous loci among parasitic and mutualistic bacterial symbionts.

Locus	Mechanism in pathogen	Mutualistic taxa and references
Type III secretion systems	Protein delivery into host	<i>Aeromonas</i> [42,63], <i>Parachlamidia</i> [42,63], <i>Rhizobium</i> [61,62,94], <i>Sinorhizobium</i> [43,61,95–97], <i>Mesorhizobium</i> [61,95], <i>Hamiltonella</i> [65]
Type IV secretion systems	Nucleo-protein delivery into host	<i>Parachlamidia</i> [42,63], <i>Wolbachia</i> [42,98], <i>Mesorhizobium</i> [61,95], <i>Sinorhizobium</i> [43]
ATP translocase	Imports ATP from host cytosol	<i>Parachlamidia</i> [42,99]
Urease gene cluster	Virulence factors	<i>Blochmannia</i> [42,100]
Cell surface polysaccharides	Defense, host suppression	<i>Bradyrhizobium</i> [61,95], <i>Sinorhizobium</i> [61,97]
Exo/ChvI system	Two-component regulatory systems	<i>Sinorhizobium</i> [43,101]

Table 2
Ongoing or finished (1) holobiont, (2) symbiont-only, (3) pathogen-only, and (4) host-only genome projects. Summarized from the National Center for Biotechnology Information (NCBI), the Genome News Network (GNN), and Genomes Online (GOLD) databases. Pathogens/parasites are defined as host-associated microbial species that can inflict morbidity and or mortality onto hosts. Mutualists/beneficial symbionts are defined as host-associated microbial species that are asymptomatic and are known or predicted to enhance host health and fitness. This table is intended as an overview of major ongoing projects, but due to the fast pace of genome sequencing, it is by no means exhaustive.

A. Whole-genome sequence available for two or more holobiont members		
Host	Vector	Symbiont/Pathogen
Vertebrates		
<i>Homo sapiens</i>	<i>Anopheles gambiae</i> (insect)	<i>Plasmodium falciparum</i> , <i>Plasmodium vivax</i> (Apicomplexan parasite)
<i>Homo sapiens</i>	<i>Biomphalaria glabrata</i> (snail)	<i>Schistosoma japonicum</i> , <i>Schistosoma mansoni</i> (trematode parasite)
<i>Homo sapiens</i>	Flea/tick	<i>Rickettsia conorii</i> , <i>Rickettsia typhi</i> , <i>Rickettsia prowazekii</i> , <i>Rickettsia sibirica</i> (bacterial pathogen)
<i>Homo sapiens</i>		gut microbiome = <i>Bacteroides fragilis</i> , <i>Bacteroides thetaiotaomicron</i> , <i>Bifidobacterium longum</i> , <i>Escherichia coli</i> , <i>Lactobacillus jhonsonii</i> , <i>Lactobacillus plantarum</i> (bacterial symbionts)
<i>Homo sapiens</i>		<i>Leptospira interrogans</i> , <i>Listeria monocytogenes</i> , <i>Salmonella enterica</i> , <i>Salmonella typhimurium</i> , <i>Shigella flexneria</i> , <i>Burholderia pseudomallei</i> , <i>Chlamydia trachomatis</i> , <i>Chlamydomydia pneumoniae</i> , <i>Clostridium perfringens</i> , <i>Clostridium tetani</i> , <i>Corynebacterium diphtheriae</i> , <i>Fusobacterium nucleatum</i> , <i>Haemophilus ducreyi</i> , <i>Haemophilus influenzae</i> , <i>Helicobacter pylori</i> , <i>Legionella pneumophila</i> , <i>Mycobacterium leprae</i> , <i>Mycobacterium tuberculosis</i> , <i>Mycoplasma genitalium</i> , <i>Mycoplasma penetrans</i> , <i>Mycoplasma pneumoniae</i> , <i>Neisseria meningitidis</i> , <i>Nocardia farcinica</i> , <i>Porphyromonas gingivalis</i> , <i>Propionibacterium acnes</i> , <i>Pseudomonas aeruginosa</i> , <i>Staphylococcus aureus</i> , <i>Streptococcus agalactiae</i> , <i>Streptococcus mutans</i> , <i>Streptococcus pneumoniae</i> , <i>Streptococcus pyogenes</i> , <i>Treponema denticola</i> , <i>Treponema pallidum</i> , <i>Tropheryma whippelli</i> , <i>Ureaplasma realyticum</i> , <i>Vibrio cholerae</i> , <i>Vibrio parahaemolyticus</i> , <i>Vibrio vulnificus</i> , <i>Yersinia pseudotuberculosis</i> , <i>Yersinia pestis</i> (bacterial pathogens)
<i>Homo sapiens</i>		<i>Candida glabrata</i> (fungal pathogen)
<i>Homo sapiens</i>		<i>Cryptosporidium parvum</i> (Apicomplexan parasite)
<i>Homo sapiens</i>		<i>Encephalitozoon cuniculi</i> (Microsporidian parasite)
Non-human primates		<i>Plasmodium knowlesi</i>
<i>Pan troglodytes</i>		<i>Plasmodium reichenowi</i>
<i>Bos taurus</i>		<i>Mannheimia succiniciproducens</i> , <i>Wolinella succinogenes</i> (bacterial symbionts)
<i>Bos taurus</i>		<i>Mycobacterium bovis</i> , <i>Mycobacterium paratuberculosis</i> , <i>Mycoplasma mycoides</i> , <i>Pasteurella multocida</i> (bacterial pathogens)
Rodents		<i>Plasmodium berghei</i> , <i>Plasmodium chabaudi</i> , <i>Plasmodium yoelii</i>
Multiple mammals (including humans)		<i>Bartonella henselae</i> , <i>Bartonella quintana</i> , <i>Bordetella bronchiseptica</i> , <i>Bordetella parapertussis</i> , <i>Bordetella pertussis</i> , <i>Borellia burgdorferi</i> , <i>Brucella melitensis</i> , <i>Brucella suis</i> , <i>Burholderia mallei</i> , <i>Campylobacter jejuni</i> , <i>Coxiella burnetii</i> , <i>Pasteurella multocida</i> (bacterial pathogens)
Avian	<i>Aedes aegypti</i> (insect)	<i>Plasmodium gallinaceum</i> (Apicomplexan parasite)
Insects		
<i>Culex quinquefasciatus</i>		<i>Wolbachia</i> sp. (bacterial symbiont)
<i>Drosophila melanogaster</i> , <i>Drosophila ananassae</i> , <i>Drosophila simulans</i> , <i>Drosophila williston</i> , <i>Drosophila americana</i> , <i>Drosophila equinoxialis</i> , <i>Drosophila erecta</i> , <i>Drosophila grimshawi</i> , <i>Drosophila littoralis</i> , <i>Drosophila mauritiana</i> , <i>Drosophila mercatorum</i> , <i>Drosophila mimica</i> , <i>Drosophila miranda</i> , <i>Drosophila mojavensis</i> , <i>Drosophila novamexicana</i> , <i>Drosophila persimilis</i> , <i>Drosophila pseudoobscura</i> , <i>Drosophila replete</i> , <i>Drosophila sechellia</i> , <i>Drosophila silvestris</i> , <i>Drosophila virilis</i>		<i>Wolbachia</i> sp. (bacterial symbiont)
<i>Acyrtosiphon pisum</i>		<i>Buchnera aphidicola</i> , <i>Candidatus Hamiltonella defensa</i> 5AT, <i>Erwinia aphidicola</i> (bacterial symbionts)
Nematodes		
<i>Brugia malayi</i>		<i>Wolbachia</i> sp. (bacterial symbiont)
Plants		
<i>Lotus japonicus</i>		<i>Rhizobium etli</i> (bacterial symbiont)
<i>Medicago truncatula</i>		<i>Rhizobium leguminosaru</i> (bacterial symbiont)
<i>Populus trichocarpa</i>		<i>Laccaria bicolor</i> (fungal symbiont)
B. Pathogen genomes sequenced. Host/vector genome sequence not available yet		
Host	Pathogen	
Plants		
Angiosperm plants	<i>Agrobacterium tumefaciens</i> , <i>Pseudomonas syringae</i> , <i>Ralstonia solanacearum</i> (bacterial)	
Cotton and citrus	<i>Ashbya gossypii</i> (fungal)	
Citrus	<i>Xanthomonas axonopodis</i> , <i>Xanthomonas campestris</i> (bacterial)	
Potato	<i>Erwinia carotovora</i> (bacterial)	
Sugarcane	<i>Leifsonia xyli</i> (bacterial)	

Table 2 (continued)

B. Pathogen genomes sequenced. Host/vector genome sequence not available yet	
Host	Pathogen
Multiple crops	<i>Xylella fastidiosa</i> (bacterial)
C. Symbiont genomes sequenced. Host/vector genome sequence not available yet	
Host	Symbiont
Mammals	<i>Bacillus anthracis</i> (bacterial)
Carpenter ants	<i>Blochmannia floridanus</i> (bacterial)
Legume plants	<i>Bradyrhizobium japonicum</i> , <i>Mesorhizobium loti</i> , <i>Sinorhizobium meliloti</i> (bacterial)
Tse-tse flies	<i>Wigglesworthia pipientis</i> (bacterial)

be interesting to resolve the extent to which genes have been transferred from symbionts to host.

Common symbiont genes and pathways

Given that gene homology and shared genomic content is relatively low across the bacteria as a whole [59], it is interesting when genes are shared among bacterial symbionts with divergent lifestyles (such as beneficial symbionts and pathogenic ones) (Table 1). For instance, recent work has shown that loci employed by bacterial parasites to take over host cells share homologs in nonvirulent and even mutualist symbionts [42,60,61]. Type III and type IV secretion systems are well-studied examples; these virulence loci encode membrane-associated complexes that inject pathogen-specific protein (type III; Hueck 1998) or nucleo-protein (type IV) effector molecules into hosts, usually with toxic effects. Yet, nonvirulent symbiotic taxa including *Wolbachia*, *Parachlamydia*, and multiple lineages of Rhizobia exhibit homologs. In the Rhizobia, the effector molecules are known as 'nodulation outer proteins' [61] and at least in one case appear to prevent full induction of plant defenses [62]. *Parachlamydia*, a fascinating Amoeba-symbiont related to parasitic *Chlamydia*, exhibits many such virulence loci such as type III secretion systems and also ATP translocase, which is used to import energetic ATP molecules directly from the host cytosol [42,63]. Similarities between pathogens and mutualists can also be found in host-defense-related loci. For instance, bacterial exo-polysaccharides protect pathogens against a host's antimicrobial arsenal, and appear to offer protection against reactive oxygen species in beneficial symbionts such as *Bradyrhizobium* and *Sinorhizobium* [64]. The recently sequenced genome of *Hamiltonella defensa*, which is an occasional endosymbiont of aphids, shows that this organism combines mechanisms known from both symbiotic and pathogenic bacteria [65].

More sequencing and the completion of new host-associated microbial genomes will enhance our understanding of the phylogenetic extent and molecular diversity of such unexpected homologies. Moreover, the combination of molecular-functional data with phylogenetic analyses will allow broader questions about the molecular pathways between beneficial symbiosis and pathogenesis to be addressed. Two specific questions are whether these loci shared between pathogens and symbionts retain similar function in their divergent genomic backgrounds, and whether they can express both parasitic and mutualistic traits. One elegant example of the latter is seen in *Aeromonas veronii*, which exhibits type III secretion systems that enable the establishment of symbiotic infections on leech hosts, but unleash deadly pathogenesis on mammalian hosts [32]. These questions are germane to our understanding of the origins of pathogenic and mutualistic strategies, which could emerge commonly if only a few mutations can transition a microbe from a mutualist to a parasite or vice versa [19,33,66]. The conventional wisdom for bacteria (based on broad scale phylogenetic patterns) is that they are constrained from rapid switches between mutualism and parasitism [67], but there is at least one well-known case in which a

mutualist has evolved directly from a parasitic lineage [34]. In the case of fungal species, there appears to be frequent saprotrophism-mutualism switches [68], but again this trend has not been examined in other eukaryotic lineages. Another question is whether pathogenic microbes commonly evolved from beneficial ancestors, or vice versa? Reconstructing gene genealogies of loci common to mutualists and parasites could be illuminating to the history of host association transitions. If host association loci originated in commensals or mutualists as opposed to pathogens, this would imply a very different selective environment that favored the origins of these traits.

Interactomes and meta-interactomics

Whole-genome sequencing projects have mostly focused on microbial partners, often with significantly smaller genomes than their multicellular hosts. Yet more recent work is generating genomic data for both or (more ambitiously) all members of symbiotic associations (Table 2), allowing the analysis of interactions among multiple genomes. Protein-protein interactions within genomes are important determinants of the evolutionary rate and fitness effects of individual genes [69]. The extension of protein-protein interaction analyses across species boundaries (meta-interactomics) will help uncover the key protein interactions that determine fitness outcomes for host and symbiont partners and might ultimately resolve the molecular bases of symbiotic coevolution. The examination of cell surface molecule interactions between hosts and symbionts is an emerging field that will necessitate a combination of genomic data with postgenomic approaches and functional assays. For instance, coimmunoprecipitation experiments have been useful in examining interactions of *Plasmodium* species with hosts or vectors, such as the identification of a complex of adhesion proteins (*PfCCp*) involved in parasite gametogenesis in host blood cells [70], and a receptor/ligand (*Psv25-calreticulin*) complex in the mosquito midgut [71]. Yeast two-hybrid screens have also proven informative to examine protein-protein interactions in parasitic systems. For example, this latter method identified the receptor in *Anopheles gambiae* that binds two homologous surface binding proteins of the parasite *Plasmodium berghei* [72]. In a study on *Bacteroides thetaiotaomicron*, Sonnenburg et al. [73] combined gene expression microarray analysis and a yeast-two hybrid screen to resolve the role of a hybrid two-component system (HTCS) in nutrient sensing and carbon metabolism; these key functions have likely conferred an adaptive advantage to this prominent distal gut species to survive in the vertebrate gut. Meta-proteomics can also be used to investigate host-symbiont interaction, but so far this has only involved microbes and not yet hosts. Verberkmoes et al. [74] used a shotgun meta-proteomic approach of the distal human gut microbiota. Their data revealed a skewed distribution of proteins involved in translation, energy production and carbohydrate metabolism relative to what would be predicted based on the metagenomic data. With the human genome in place, the molecular interactions with the distal-gut microcommunity can begin to be resolved in earnest.

The availability of genome data from multiple plant species that form symbiosis with diverse nitrogen-fixing bacteria has helped uncover conserved genetic programs involved in the onset of symbiosis across phylogenetically distant hosts. These genetic programs are implicated in plant developmental processes but later in evolution were recruited to play roles in root nodule symbiosis [75]. The whole-genome sequencing of the Poplar tree (*Populus trichocarpa*) [76] and its fungal symbiont (*Laccaria bicolor*) [48] allows us to start examining if similar evolutionary processes are shaping eukaryotic symbiotic genomes as those recognized in prokaryotic systems. For instance, horizontal gene transfer does not seem to play as an important role as gene duplication as the driver for specialization in *L. bicolor* [49]. Microarray analysis of onset of *L. bicolor* infection has identified new candidate genes that are likely responsible for a successful establishment of symbiosis [48]. Metabolomics (i.e., high-throughput analysis of metabolic end products) [77] and meta-metabolomics (i.e., 'holobiont' analysis, see below) will also likely provide insight into the metabolism of mutualistic interactions, as well as the examination of chemical–protein network interactions [25].

From genotype to ecosystems

The holobiont is a concept that was initially introduced to refer to the community of organisms that includes a coral host and all its microbial symbionts (including unicellular eukaryotes) [5]. This term can be expanded to any symbiotic interaction between multiple organisms (e.g., a tree and its microbial and fungal endophytes, or the human body and its microbiota). As entire holobionts are sequenced, new perspectives on how we view ecological communities are starting to emerge, for instance how genetic interactions among holobiont members and the surrounding environment can have effects on larger communities by defining ecosystem processes such as nutrient cycling [2,78]. By linking the actual genes involved in mutualistic interactions with the genotypic background of the host, it now seems possible to extend our understanding of ecosystem processes. This understanding can be achieved by the fact that only a few interactions (e.g., a dominant tree species in a forest and its interacting species partners) can determine entire ecosystem communities [78]. Endophytic fungi will be a particularly interesting case to examine to address questions related to genome evolution of the fungi and the host plant as they offer a continuum of lifestyles that encompasses saprotrophism, symbiosis, and pathogenesis [79].

As humans, we are particularly interested in the assembly of our own internal microbial community, which is most dense and diverse in our intestine [80]. Intense study of our own holobiont and those of our mammalian relatives has revealed that the assembly of these micro-

communities appears to be strongly host-dependent. Across our mammalian relatives, host diet and host phylogenetic history are both important factors that influence each species' symbiont community makeup [81]. Yet, studies of human monozygotic twins further show that our own human gut microcommunities are most strongly shaped by host genotype, with diet being secondary in determining community structure [82]. Furthermore, once a human gut microcommunity is established in an individual, it appears to be stable over time [83]. Taken together, these data suggest a 'top-down' model for the assembly of the human gut microcommunity; once symbiotic communities are established, host-driven selection appears to outweigh microbial competition as the driving force shaping symbiotic diversity [84]. Given this hypothesis, research should now focus on elucidating the mechanisms that human hosts use to shape the diversity of our gut symbionts thus maximizing the fitness benefits that we receive. Elucidating these host-driven mechanisms and the microbial ecology that they affect is a first step towards understanding some of the many pathologies that are linked to gut symbiont dysfunction [85].

Conclusions

Future directions

Comparative genomics has already proven to be a critical step in understanding coevolutionary trajectories of host and symbiont genomes. While the focus has initially been on the microbial lineages, similar analyses for host genomes should now be tractable (e.g., mammalian, plant, and insect hosts). Some discernible questions in the near future are highlighted in Box 1.

The majority of these questions can be addressed by molecular evolutionary and comparative genomic approaches to both new whole-genome data from emerging eukaryotic genomes and re-sequencing efforts with 'next-gen' sequencing technologies. The pioneering study of Moran et al. [37] mentioned above, in which recently diverged populations of *Buchnera* were sequenced by the Solexa method, provided a new model of genome evolution at a finer scale simply not possible a couple of years ago. New models will likely soon emerge from re-sequencing approaches of human gut-microbiome in genetically diverse host populations, as well as resequencing of rhizobial symbionts in plants, and *Wolbachia* symbionts in insects among others. While *Buchnera* and *Wolbachia* projects have been 'targeted' symbionts, in many microbiomes, symbionts will have to be identified first. In such cases, single-cell genomics appears an important step in circumventing the issue of low cultivability success [86].

To broaden our understanding of how host–symbiont interactions are actually taking place at the cellular level, a whole suite of

Box 1

Untested hypotheses that can be addressed with genomic and postgenomic data

- To what extent are overall host genome repertoires shaped by the gene contents of their microbial symbionts?
- Are similar evolutionary forces (e.g., horizontal gene transfer, deletion, gene duplication, etc.) shaping bacterial and eukaryotic endosymbiont genomes?
 - o Is host genome size and content affected by the type of symbiont or symbiont community?
- How is natural selection shaping the evolution of host-associated symbiont loci?
 - o What loci are under purifying or positive selection in symbionts?
 - o Are positively selected loci ever consistent across symbiotic lineages?
- Are there core sets of conserved genes common to microbes associated with any particular host taxon?
 - o How do these genes evolve when there are switches to pathogenic relationships?
- Are there any 'universal' loci required for all infectious symbionts? Or are these loci phylogenetically restricted to certain clades?
- Are common host-associated loci ancestral to the evolution of host infection and thus co-opted for host-associated processes or are they horizontally transmitted among many genomic backgrounds?
- How much of the host–symbiont interaction is controlled by regulatory processes rather than genomic differences? Are these different in host and symbionts?

postgenomic techniques will have to be invoked. Microarray data have already provided some insight into how hosts and their respective symbionts may be communicating. For instance, in the case of the onset of coral symbiosis in larval stages, only unsuccessful dinoflagellate symbionts (i.e., those failing to establish symbiosis) caused a radically different transcriptome response (i.e., by microarray analysis) in the host, pointing to regulation of apoptosis and the immune system in the host as likely mediators in the symbiosis [87]. In contrast, the successful symbionts establish symbiosis in a stealth manner with minimal change in gene expression profiles in the host [87]. A similar microarray study in the poplar-*Laccaria* symbiosis showed that several genes encoding small, secreted peptides are expressed during symbiosis [48]. Gene expression analysis has also recently provided insight into the interaction of the human facultative pathogen *Salmonella* when living in transient symbiosis with the free-living amoeba *Acanthamoeba rhysodes*, showing that the genes from the *Salmonella* pathogenicity island and virulence plasmid are upregulated during infection inducing apoptosis-like cell death in this transitory environmental host [88].

Protein networks have already been successfully elucidated for several model organisms [89–91], and the technology could be expanded to examine holobiont communities. A recent proteomic study of cadmium stress alleviation in *Medicago truncatula* by arbuscular mycorrhizal symbiosis demonstrated the usefulness of this approach. Although only Expressed Sequenced Tags (EST) data are currently available for *M. truncatula*, about 30 proteins involved in metal stress alleviation were identified in symbiotic plants [92]. Therefore, it is quite feasible that proteome data from organisms and communities with whole-genome and metagenome data can soon be fully integrated with high-throughput functional analyses such as gene expression or proteomic assays. This can potentially help identify genes and pathways involved in partner interactions as well as multiple receptor/ligand complexes involved in the overall regulation of the establishment and maintenance of host-symbiont relationships.

In conclusion, the time is ripe for the large-scale examination of host-symbiont interactions in different parts of the tree of life at multiple levels of biological organization. A comparative approach that will certainly enable us to reconstruct Darwin's 'entangled bank.'

Acknowledgments

We would like to thank John Quackenbush for the invitation to write this review. We would also like to thank Pilar Francino, Christian Voolstra, Shini Sunagawa, Mickey DeSalvo, Ryan Skophammer, and two anonymous reviewers for their helpful comments on the manuscript. Benoît Dayrat helped with figure design. M.M. was supported by NSF grants BE-GEN 0313708 and IOS 0644438. J.S. was supported by NSF DEB 0816663.

References

- J.H. Andrews, R.F. Harris, The ecology and biogeography of microorganisms of plant surfaces, *Annu. Rev. Phytopathol.* 38 (2000) 145–180.
- P.B. Eckburg, E.M. Bik, C.N. Bernstein, E. Purdom, L. Dethlefsen, M. Sargent, S.R. Gill, K.E. Nelson, D.A. Relman, Diversity of the human intestinal microbial flora, *Science* 308 (2005) 1635–1638.
- U. Hentschel, J. Hopke, M. Horn, A.B. Friedrich, M. Wagner, J. Hacker, B.S. Moore, Molecular evidence for a uniform microbial community in sponges from different oceans, *Appl. Environ. Microbiol.* 68 (2002) 4431–4440.
- N.A. Moran, J.P. McCutcheon, A. Nakabachi, Genomics and evolution of heritable bacterial symbionts, *Annu. Rev. Genet.* (2008) 42.
- F. Rohwer, F. Azam, N. Knowlton, Diversity and distribution of coral-associated bacteria, *Abstr. Gen. Meet. Am. Soc. Microbiol.* 102 (2002) 334.
- R. Rowan, D.A. Powers, Ribosomal-RNA sequences and the diversity of symbiotic dinoflagellates (Zooxanthellae), *Proc. Natl. Acad. Sci. U. S. A.* 89 (1992) 3639–3643.
- C. Darwin, *On the Origin of Species*, London, 1859.
- A.E. Douglas, *Symbiotic Interactions*, Oxford University Press, 1994.
- S. Paracer, V. Ahmadian, *Symbiosis, An Introduction to Biological Associations*, Oxford University Press, 2000.
- M. Kalliomaki, S. Salminen, H. Arvilommi, P. Kero, P. Koskinen, E. Isolauri, Probiotics in primary prevention of atopic disease: a randomised placebo-controlled trial, *Lancet* 357 (2001) 1076–1079.
- J. Xu, J.L. Gordon, Honor thy symbionts, *Proc. Natl. Acad. Sci. U. S. A.* 100 (2003) 10452–10459.
- E.T. Kiers, S.A. West, R.F. Denison, Mediating mutualisms: farm management practices and evolutionary changes in symbiont co-operation, *J. App. Ecol.* 39 (2002) 745–754.
- H.H. Musa, S.L. Wu, C.H. Zhu, H.I. Seri, G.Q. Zhu, The potential benefits of probiotics in animal production and health, *J. Anim. Vet. Adv.* (2009) 8.
- R. Rowan, N. Knowlton, A. Baker, J. Jara, Landscape ecology of algal symbionts creates variation in episodes of coral bleaching, *Nature* 388 (1997) 265–269.
- M.G.A.v.d. Heijden, J.N. Klironomos, M. Ursic, P. Moutoglis, R. Streitwolf-Engel, T. Boller, A. Wiemken, I.R. Sanders, Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity, *Nature* 396 (1998) 69–72.
- A.E. Douglas, Nutritional interactions in insect-microbial symbioses: aphids and their symbiotic bacteria *Buchnera*, *Ann. Rev. Entomol.* 43 (1998) 17–37.
- J.K. Taubenberger, D.M. Morens, 1918 influenza: the mother of all pandemics, *Emerg. Infect. Dis.* 12 (2006) 15–22.
- J.L. Sachs, U.G. Mueller, T.P. Wilcox, J.J. Bull, The evolution of cooperation, *Q. Rev. Biol.* 79 (2004) 135–160.
- J.L. Sachs, E.L. Simms, Pathways to mutualism breakdown, *Trends Ecol. Evol.* 21 (2006) 585–592.
- L. Margulis, *Symbiosis in Cell Evolution*, Freeman, 1981.
- D.J. Funk, L. Helbling, J.J. Wernegreen, N.A. Moran, Intraspecific phylogenetic congruence among multiple symbiont genomes, *Proc. Biol. Sci.* 267 (2000) 2517–2521.
- M. Roy, M.P. Dubois, M. Proffitt, L. Vincenot, E. Desmarais, M.A. Selosse, Evidence from population genetics that the ectomycorrhizal basidiomycete *Laccaria amethystina* is an actual multihost symbiont, *Mol. Ecol.* 17 (2008) 2825–2838.
- E. Martinez-Romero, Coevolution in *Rhizobium*-legume symbiosis? *DNA Cell Biol.* 28 (2009) 361–370.
- J. Sapp, *Evolution by Association. A History of Symbiosis*, Oxford University Press, New York, 1994.
- M. Medina, Genomes, phylogeny, and evolutionary systems biology, *Proc. Natl. Acad. Sci. U. S. A.* 102 (2005) 6630–6635.
- C. Dobell, Antony van Leeuwenhoek and His "Little Animals", New York, 1932.
- J.R. Leadbetter, Cultivation of recalcitrant microbes: cells are alive, well and revealing their secrets in the 21st century laboratory, *Curr. Opin. Microbiol.* 6 (2003) 274–281.
- N.R. Pace, G.J. Olsen, C.R. Woese, Ribosomal RNA phylogeny and the primary lines of evolutionary descent, *Cell* 45 (1986) 325–326.
- C.R. Woese, O. Kandler, M.L. Wheelis, Towards a natural system of organisms: proposal for the domains archaea, bacteria, and eucarya, *Proc. Natl. Acad. Sci. U. S. A.* 87 (1990) 4576–4579.
- C.E. Morris, M. Bardin, O. Berge, P. Frey-Klett, N. Fromin, H. Girardin, M.H. Guinebreteiere, P. Lebaron, J.M. Thiery, M. Troussellier, Microbial biodiversity: approaches to experimental design and hypothesis testing in primary scientific literature from 1975 to 1999, *Microbiol. Mol. Biol. Rev.* 66 (2002) 592–616 table of contents.
- J.C. Venter, K. Remington, J.F. Heidelberg, A.L. Halpern, D. Rusch, J.A. Eisen, D.Y. Wu, I. Paulsen, K.E. Nelson, W. Nelson, D.E. Fouts, S. Levy, A.H. Knap, M.W. Lomas, K. Nealson, O. White, J. Peterson, J. Hoffman, R. Parsons, H. Baden-Tillson, C. Pfannkoch, Y.H. Rogers, H.O. Smith, Environmental genome shotgun sequencing of the Sargasso Sea, *Science* 304 (2004) 66–74.
- A.C. Silver, Y. Kikuchi, A.A. Fadl, J. Sha, A.K. Chopra, J. Graf, Interaction between innate immune cells and a bacterial type III secretion system in mutualistic and pathogenic associations, *Proc. Natl. Acad. Sci. U. S. A.* 104 (2007) 9481–9486.
- J.L. Sachs, T.P. Wilcox, A shift to parasitism in the jellyfish symbiont *Symbiodinium microadriaticum*, *Proc. R. Soc. Lond. B. Biol. Sci.* 273 (2006) 425–429.
- A.R. Weeks, M. Turelli, W.R. Harcombe, K.T. Reynolds, A.A. Hoffmann, From parasite to mutualist: Rapid evolution of *Wolbachia* in natural populations of *Drosophila*, *Plos Biol.* 5 (2007) 997–1005.
- C.H. Kuo, H. Ochman, The fate of new bacterial genes, *FEMS Microbiol. Rev.* 33 (2009) 38–43.
- E. Lerat, V. Daubin, H. Ochman, N.A. Moran, Evolutionary origins of genomic repertoires in bacteria, *PLoS Biol.* 3 (2005) e130.
- N.A. Moran, H.J. McLaughlin, R. Sorek, The dynamics and time scale of ongoing genomic erosion in symbiotic bacteria, *Science* 323 (2009) 379–382.
- H. Ochman, S.R. Santos, Eyeing bacterial genomes, *Curr. Opin. Microbiol.* 6 (2003) 109–113.
- S.G.E. Andersson, C.G. Kurland, Reductive evolution of resident genomes, *Trends Microbiol.* 6 (1998) 263–268.
- A.I. Nilsson, S. Koskiniemi, S. Eriksson, E. Kugelberg, J.C. Hinton, D.I. Andersson, Bacterial genome size reduction by experimental evolution, *Proc. Natl. Acad. Sci. U. S. A.* 102 (2005) 12112–12116.
- H. Ochman, Genomes on the shrink, *Proc. Natl. Acad. Sci. U. S. A.* 102 (2005) 11959–11960.
- J.J. Wernegreen, For better or worse: genomic consequences of intracellular mutualism and parasitism, *Curr. Opin. Genet. Dev.* 15 (2005) 572–583.
- J. Batut, S.G.E. Andersson, D. O'Callaghan, The evolution of chronic infection strategies in the alpha-proteobacteria, *Nat. Rev. Microbiol.* 2 (2004) 933–945.
- W.R. Streit, R.A. Schmitz, X. Perret, C. Staehelin, W.J. Deakin, C. Raasch, H. Liesegang, W.J. Broughton, An evolutionary hot spot: the pNGR234b replicon of *Rhizobium* sp. strain NGR234, *J. Bacteriol.* 186 (2004) 535–542.

- [45] K. Wong, G.B. Golding, A phylogenetic analysis of the pSymB replicon from the *Sinorhizobium meliloti* genome reveals a complex evolutionary history, *Can. J. Microbiol.* 49 (2003) 269–280.
- [46] F. Galibert, T.M. Finan, S.R. Long, A. Puhler, P. Abola, F. Ampe, F. Barloy-Hubler, M.J. Barnett, A. Becker, P. Boistard, G. Bothe, M. Boutry, L. Bowser, J. Buhrmester, E. Cadieu, D. Capela, P. Chain, A. Cowie, R.W. Davis, S. Dreano, N.A. Federspiel, R.F. Fisher, S. Gloux, T. Godrie, A. Goffeau, B. Golding, J. Gouzy, M. Gurjal, I. Hernandez-Lucas, A. Hong, L. Huizar, R.W. Hyman, T. Jones, D. Kahn, M.L. Kahn, S. Kalman, D.H. Keating, E. Kiss, C. Komp, V. Lalaurie, D. Masuy, C. Palm, M.C. Peck, T.M. Pohl, D. Portetelle, B. Purnelle, U. Ramsperger, R. Surzycki, P. Thebault, M. Vandenbol, F.J. Vorholter, S. Weidner, D.H. Wells, K. Wong, K.C. Yeh, J. Batut, The composite genome of the legume symbiont *Sinorhizobium meliloti*, *Science* 293 (2001) 668–672.
- [47] T.C. Lajeunesse, G. Lambert, R.A. Andersen, M.A. Coffroth, D.W. Galbraith, *Symbiodinium* (Pyrrophyta) genome sizes (DNA content) are smallest among dinoflagellates, *J. Phycol.* 41 (2005) 880–886.
- [48] F. Martin, A. Aerts, D. Ahren, A. Brun, E.G. Danchin, F. Duchaussoy, J. Gibon, A. Kohler, E. Lindquist, V. Pereda, A. Salamov, H.J. Shapiro, J. Wuyts, D. Blaudez, M. Buee, P. Brockstein, B. Canback, D. Cohen, P.E. Courty, P.M. Coutinho, C. Delaruelle, J.C. Detter, A. Deveau, S. DiFazio, S. Duplessis, L. Fraissinet-Tachet, E. Lucic, P. Frey-Klett, C. Fourrey, I. Feussner, G. Gay, J. Grimwood, P.J. Hoegger, P. Jain, S. Kilaru, J. Labbe, Y.C. Lin, V. Legue, F. Le Tacon, R. Marmeisse, D. Melayah, B. Montanini, M. Muratet, U. Nehls, H. Niculita-Hirzel, M.P. Oudot-Le Secq, M. Peter, H. Quesneville, B. Rajashekar, M. Reich, N. Rouhier, J. Schmutz, T. Yin, M. Chalot, B. Henrissat, U. Kues, S. Lucas, Y. Van de Peer, G.K. Podila, A. Polle, P.J. Pukkila, P.M. Richardson, P. Rouze, I.R. Sanders, J.E. Stajich, A. Tunlid, G. Tuskan, I.V. Grigoriev, The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis, *Nature* 452 (2008) 88–92.
- [49] F. Martin, M.A. Selosse, The *Laccaria* genome: a symbiont blueprint decoded, *New Phytol.* 180 (2008) 296–310.
- [50] D.M. Soanes, I. Alam, M. Cornell, H.M. Wong, C. Hedeler, N.W. Paton, M. Rattray, S.J. Hubbard, S.G. Oliver, N.J. Talbot, Comparative genome analysis of filamentous fungi reveals gene family expansions associated with fungal pathogenesis, *PLoS One* 3 (2008) e2300.
- [51] P. Normand, P. Lapiere, L.S. Tisa, J.P. Gogarten, N. Alloisio, E. Bagnard, C.A. Bassi, A.M. Berry, D.M. Bickhart, N. Choise, A. Couloux, B. Cournoyer, S. Cruveiller, V. Daubin, N. Demange, M.P. Francino, E. Goltsman, Y. Huang, O.R. Kopp, L. Labarre, A. Lapidus, C. Lavire, J. Marechal, M. Martinez, J.E. Mastrorunzio, B.C. Mullin, J. Niemann, P. Pujic, T. Rawnsley, Z. Rouy, C. Schenowitz, A. Sellstedt, F. Tavares, J.P. Tomkins, D. Vallenet, C. Valverde, L.G. Wall, Y. Wang, C. Medigue, D.R. Benson, Genome characteristics of facultatively symbiotic *Frankia* sp. strains reflect host range and host plant biogeography, *Genome Res.* 17 (2007) 7–15.
- [52] J.T. Sullivan, H.N. Patrick, W.L. Lowther, D.B. Scott, C.W. Ronson, Nodulating strains of *Rhizobium loti* arise through chromosomal symbiotic gene transfer in the environment, *Proc. Natl. Acad. Sci. U. S. A.* 92 (1995) 8985–8989.
- [53] H. Brussow, C. Canchaya, W.D. Hardt, Phages and the evolution of bacterial pathogens: From genomic rearrangements to lysogenic conversion, *Microbiol. Mol. Biol. Rev.* 68 (2004) 560–+.
- [54] E. Lerat, V. Daubin, H. Ochman, N.A. Moran, Evolutionary origins of genomic repertoires in bacteria, *Plos Biol.* 3 (2005) 807–814.
- [55] A. Mira, H. Ochman, N.A. Moran, Deletional bias and the evolution of bacterial genomes, *Trends Genet.* 17 (2001) 589–596.
- [56] E.A. Gladyshev, M. Meselson, I.R. Arkhipova, Massive horizontal gene transfer in bdelloid rotifers, *Science* 320 (2008) 1210–1213.
- [57] J.C. Hotopp, M.E. Clark, D.C. Oliveira, J.M. Foster, P. Fischer, M.C. Torres, J.D. Giebel, N. Kumar, N. Ishmael, S. Wang, J. Ingram, R.V. Nene, J. Shepard, J. Tomkins, S. Richards, D.J. Spiro, E. Ghedin, B.E. Slatko, H. Tettelin, J.H. Werren, Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes, *Science* 317 (2007) 1753–1756.
- [58] N. Nikoh, A. Nakabachi, Aphids acquired symbiotic genes via lateral gene transfer, *BMC Biol.* 7 (2009) 12.
- [59] R.L. Charlebois, W.F. Doolittle, Computing prokaryotic gene ubiquity: rescuing the core from extinction, *Genome Res.* 14 (2004) 2469–2477.
- [60] W.J. Deakin, W.J. Broughton, Symbiotic use of pathogenic strategies: rhizobial protein secretion systems, *Nat. Rev. Microbiol.* 7 (2009) 312–320.
- [61] M.J. Soto, A. Dominguez-Ferreras, D. Perez-Mendoza, J. Sanjuan, J. Olivares, Mutualism versus pathogenesis: the give-and-take in plant–bacteria interactions, *Cell. Microbiol.* 11 (2009) 381–388.
- [62] A.V. Bartsev, W.J. Deakin, N.M. Boukli, B.C. McAlvin, G. Stacey, P. Malnoe, W.J. Broughton, C. Staehelin, Nopl, an effector protein of *Rhizobium* sp. NGR234, thwarts activation of plant defense reactions, *Plant Physiol.* 134 (2004) 871–879.
- [63] M. Horn, A. Collingro, S. Schmitz-Esser, C.L. Beier, U. Purkhold, B. Fartmann, P. Brandt, G.J. Nyakatura, M. Droege, D. Frishman, T. Rattai, H.W. Mewes, M. Wagner, Illuminating the evolutionary history of chlamydiae, *Science* 304 (2004) 728–730.
- [64] W. D'Haese, M. Holsters, Surface polysaccharides enable bacteria to evade plant immunity, *Trends Microbiol.* 12 (2004) 555–561.
- [65] P.H. Degnan, Y. Yu, N. Sisneros, R.A. Wing, N.A. Moran, *Hamiltonella defensa*, genome evolution of protective bacterial endosymbiont from pathogenic ancestors, *Proc. Natl. Acad. Sci. U. S. A.* 106 (2009) 9063–9068.
- [66] J.L. Sachs, E.L. Simms, The origins of uncooperative rhizobia, *Oikos* 117 (2008) 961–966.
- [67] N.A. Moran, J.J. Wernegreen, Lifestyle evolution in symbiotic bacteria: insights from genomics, *Trends Ecol. Evol.* 15 (2000) 321–326.
- [68] D.S. Hibbett, P.B. Matheny, The relative ages of ectomycorrhizal mushrooms and their plant hosts estimated using Bayesian relaxed molecular clock analyses, *BMC Biol.* 7 (2009) 13.
- [69] H.B. Fraser, A.E. Hirsh, L.M. Steinmetz, C. Scharfe, M.W. Feldman, Evolutionary rate in the protein interaction network, *Science* 296 (2002) 750–752.
- [70] N. Simon, S.M. Scholz, C.K. Moreira, T.J. Templeton, A. Kuehn, M.A. Dude, G. Pradel, Sexual stage adhesion proteins form multi-protein complexes in the malaria parasite *Plasmodium falciparum*, *J. Biol. Chem.* 284 (2009) 14537–14546.
- [71] C. Rodriguez Mdel, J. Martinez-Barnette, A. Alvarado-Delgado, C. Batista, R.S. Argotte-Ramos, S. Hernandez-Martinez, L. Gonzalez Ceron, J.A. Torres, G. Margos, M.H. Rodriguez, The surface protein Pvs25 of *Plasmodium vivax* ookinetes interacts with calreticulin on the midgut apical surface of the malaria vector *Anopheles albimanus*, *Mol. Biochem. Parasitol.* 153 (2007) 167–177.
- [72] D. Vlachou, G. Lycett, I. Siden-Kiamos, C. Blass, R.E. Sinden, C. Louis, *Anopheles gambiae* laminin interacts with the P25 surface protein of *Plasmodium berghei* ookinetes, *Mol. Biochem. Parasitol.* 112 (2001) 229–237.
- [73] E.D. Sonnenburg, J.L. Sonnenburg, J.K. Manchester, E.E. Hansen, H.C. Chiang, J.I. Gordon, A hybrid two-component system protein of a prominent human gut symbiont couples glycan sensing in vivo to carbohydrate metabolism, *Proc. Natl. Acad. Sci. U. S. A.* 103 (2006) 8834–8839.
- [74] N.C. Verberkmoes, A.L. Russell, M. Shah, A. Godzik, M. Rosenquist, J. Halfvarson, M.G. Lefsrud, J. Apajalahti, C. Tysk, R.L. Hettich, J.K. Jansson, Shotgun metaproteomics of the human distal gut microbiota, *ISME J.* 3 (2009) 179–189.
- [75] K. Markmann, M. Parniske, Evolution of root endosymbiosis with bacteria: how novel are nodules? *Trends Plant Sci.* 14 (2009) 77–86.
- [76] G.A. Tuskan, S. Difazio, S. Jansson, J. Bohlmann, I. Grigoriev, U. Hellsten, N. Putnam, S. Ralph, S. Rombauts, A. Salamov, J. Schein, L. Sterck, A. Aerts, R.R. Bhalaria, R.P. Bhalaria, D. Blaudez, W. Boerjan, A. Brun, A. Brunner, V. Busov, M. Campbell, J. Carlson, M. Chalot, J. Chapman, G.L. Chen, D. Cooper, P.M. Coutinho, J. Couturier, S. Covert, Q. Cronk, R. Cunningham, J. Davis, S. Degroove, A. DeJardin, C. Depamphilis, J. Detter, B. Dirks, I. Dubchak, S. Duplessis, J. Ehlting, B. Ellis, K. Gendler, D. Goodstein, M. Gribskov, J. Grimwood, A. Groover, L. Gunter, B. Hamberger, B. Heinze, Y. Helariutta, B. Henrissat, D. Holligan, R. Holt, W. Huang, N. Islam-Faridi, S. Jones, M. Jones-Rhoades, R. Jorgensen, C. Joshi, J. Kangasjarvi, J. Karlsson, C. Kelleher, R. Kirkpatrick, M. Kirst, A. Kohler, U. Kalluri, F. Larimer, J. Leebens-Mack, J.C. Leple, P. Locasio, Y. Lou, S. Lucas, F. Martin, B. Montanini, C. Napoli, D.R. Nelson, C. Nelson, K. Nieminen, O. Nilsson, V. Pereda, G. Peter, R. Philippe, G. Pilate, A. Poliakov, J. Razumovskaya, P. Richardson, C. Rinaldi, K. Ritland, P. Rouze, D. Ryaboy, J. Schmutz, J. Schrader, B. Segerman, H. Shin, A. Siddiqui, F. Sterky, A. Terry, C.J. Tsai, E. Uberbacher, P. Unneberg, et al., The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray), *Science* 313 (2006) 1596–1604.
- [77] O. Fiehn, Metabolomics—the link between genotypes and phenotypes, *Plant Mol. Biol.* 48 (2002) 155–171.
- [78] T.G. Whitham, S.P. Difazio, J.A. Schweitzer, S.M. Shuster, G.J. Allan, J.K. Bailey, S.A. Woolbright, Extending genomics to natural communities and ecosystems, *Science* 320 (2008) 492–495.
- [79] R. Rodriguez, R. Redman, More than 400 million years of evolution and some plants still can't make it on their own: plant stress tolerance via fungal symbiosis, *J. Exp. Bot.* 59 (2008) 1109–1114.
- [80] W.B. Whitman, D.C. Coleman, W.J. Wiebe, Prokaryotes: the unseen majority, *Proc. Natl. Acad. Sci. U. S. A.* 95 (1998) 6578–6583.
- [81] R.E. Ley, M. Hamady, C. Lozupone, P.J. Turnbaugh, R.R. Ramey, J.S. Birchler, M.L. Schlegel, T.A. Tucker, M.D. Schrenzel, R. Knight, J.I. Gordon, Evolution of mammals and their gut microbes, *Science* 320 (2008) 1647–1651.
- [82] E.G. Zoetendal, The host genotype affects the bacterial community in the human gastrointestinal tract, *Microb. Ecol. Health Dis.* (2001) 129–134.
- [83] E.G. Zoetendal, A.D.L. Akkermans, W.M. De Vos, Temperature gradient gel electrophoresis analysis of 16S rRNA from human fecal samples reveals stable and host-specific communities of active bacteria, *App. Env. Microbiol.* 64 (1998) 3854–3859.
- [84] F. Backhed, R.E. Ley, J.L. Sonnenburg, D.A. Peterson, J.I. Gordon, Host–bacterial mutualism in the human intestine, *Science* 307 (2005) 1915–1920.
- [85] F. Guarner, J.R. Malagelada, Gut flora in health and disease, *Lancet* 361 (2003) 512–519.
- [86] T. Ishoye, T. Woyke, R. Stepanauskas, M. Novotny, R.S. Lasken, Genomic sequencing of single microbial cells from environmental samples, *Curr. Opin. Microbiol.* 11 (2008) 198–204.
- [87] C.R. Voolstra, J.A. Schwarz, J. Schnetzler, S. Sunagawa, M.K. Desalvo, A.M. Szmant, M.A. Coffroth, M. Medina, The host transcriptome remains unaltered during the establishment of coral–algal symbioses, *Mol. Ecol.* 18 (2009) 1823–1833.
- [88] Y. Feng, Y.H. Hsiao, H.L. Chen, C. Chu, P. Tang, C.H. Chiu, Apoptosis-like cell death induced by *Salmonella* in *Acanthamoeba rhysodes*, *Genomics* 94 (2009) 132–137.
- [89] L. Giot, J.S. Bader, C. Brouwer, A. Chaudhuri, B. Kuang, Y. Li, Y.L. Hao, C.E. Ooi, B. Godwin, E. Vitols, G. Vijayadamar, P. Pochart, H. Machineni, M. Welsh, Y. Kong, B. Zerhusen, R. Malcol, Z. Varrone, A. Collis, M. Minto, S. Burgess, L. McDaniel, E. Stimpson, F. Spriggs, J. Williams, K. Neurath, N. Ioime, M. Agee, E. Voss, K. Furtak, R. Renzulli, N. Aenansen, S. Carrolla, E. Bickelhaup, Y. Lazovatsky, A. DaSilva, J. Zhong, C.A. Stanyon, R.L. Finley Jr., K.P. White, M. Braverman, T. Jarvie, S. Gold, M. Leach, J. Knight, R.A. Shimkets, M.P. McKenna, J. Chant, J.M. Rothberg, A protein interaction map of *Drosophila melanogaster*, *Science* 302 (2003) 1727–1736.
- [90] S. Li, C.M. Armstrong, N. Bertin, H. Ge, S. Milstein, M. Boxem, P.O. Vidalain, J.D. Han, A. Chesneau, T. Hao, D.S. Goldberg, N. Li, M. Martinez, J.F. Rual, P. Lamesch, L. Xu, M. Tewari, S.L. Wong, L.V. Zhang, G.F. Bertiz, L. Jacotot, P. Vaglio, J. Reboul, T. Hirozane-Kishikawa, Q. Li, H.W. Gabel, A. Elewa, B. Baumgartner, D.J. Rose, H. Yu, S. Bosak, R. Sequerra, A. Fraser, S.E. Mango, W.M. Saxton, S. Strome, S. Van Den Heuvel, F. Piano, J. Vandenhaute, C. Sartet, M. Gerstein, L. Doucette-Stamm, K.C. Gunsalus, J.W. Harper, M.E. Cusick, F.P. Roth, D.E. Hill, M. Vidal, A map of the interactome network of the metazoan *C. elegans*, *Science* 303 (2004) 540–543.

- [91] P. Uetz, L. Giot, G. Cagney, T.A. Mansfield, R.S. Judson, J.R. Knight, D. Lockshon, V. Narayan, M. Srinivasan, P. Pochart, A. Qureshi-Emili, Y. Li, B. Godwin, D. Conover, T. Kalbfleisch, G. Vijayadamar, M. Yang, M. Johnston, S. Fields, J.M. Rothberg, A comprehensive analysis of protein–protein interactions in *Saccharomyces cerevisiae*, *Nature* 403 (2000) 623–627.
- [92] A. Aloui, G. Recorbet, A. Gollotte, F. Robert, B. Valot, V. Gianinazzi-Pearson, S. Aschi-Smiti, E. Dumas-Gaudot, On the mechanisms of cadmium stress alleviation in *Medicago truncatula* by arbuscular mycorrhizal symbiosis: a root proteomic study, *Proteomics* 9 (2009) 420–433.
- [93] A.J. Roger, A.G. Simpson, Evolution: revisiting the root of the eukaryote tree, *Curr. Biol.* 19 (2009) R165–167.
- [94] P. Skorpil, M.M. Saad, N.M. Boukli, H. Kobayashi, F. Ares-Orpel, W.J. Broughton, W.J. Deakin, NopP, a phosphorylated effector of *Rhizobium* sp strain NGR234, is a major determinant of nodulation of the tropical legumes *Flemingia congesta* and *Tephrosia vogelii*, *Mol. Microbiol.* 57 (2005) 1304–1317.
- [95] A. Mithofer, A.A. Bhagwat, M. Feger, J. Ebel, Suppression of fungal beta-glucan-induced plant defence in soybean (*Glycine max* L) by cyclic 1,3-1,6-beta-glucans from the symbiont *Bradyrhizobium japonicum*, *Planta* 199 (1996) 270–275.
- [96] J.A. Rodrigues, F.J. Lopez-Baena, F.J. Ollero, J.M. Vinardell, M.D. Espuny, R.A. Bellogin, J.E. Ruiz-Sainz, J.R. Thomas, D. Sumpton, J. Ault, J. Thomas-Oates, NopM and NopD are rhizobial nodulation outer proteins: Identification using LC-MALDI and LC-ESI with a monolithic capillary column, *J. Proteome Res.* 6 (2007) 1029–1037.
- [97] H. Scheidle, A. Gross, K. Niehaus, The lipid A substructure of the *Sinorhizobium meliloti* lipopolysaccharides is sufficient to suppress the oxidative burst in host plants, *New Phytol.* 165 (2005) 559–565.
- [98] J. Foster, M. Ganatra, I. Kamal, J. Ware, K. Makarova, N. Ivanova, A. Bhattacharyya, V. Kapatral, S. Kumar, J. Posfai, T. Vincze, J. Ingram, L. Moran, A. Lapidus, M. Omelchenko, N. Kyrpides, E. Ghedin, S. Wang, E. Goltsman, V. Joukov, O. Ostrovskaya, K. Tsukerman, M. Mazur, D. Comb, E. Koonin, B. Slatko, The *Wolbachia* genome of *Brugia malayi*: endosymbiont evolution within a human pathogenic nematode, *Plos Biol.* 3 (2005) 599–614.
- [99] S. Schmitz-Esser, N. Linka, A. Collingro, C.L. Beier, H.E. Neuhaus, M. Wagner, M. Horn, ATP/ADP translocases: a common feature of obligate intracellular amoebal symbionts related to chlamydiae and rickettsiae, *J. Bacteriol.* 186 (2004) 683–691.
- [100] P.H. Degnan, A.B. Lazarus, J.J. Wernegreen, Genome sequence of *Blochmannia pennsylvanicus* indicates parallel evolutionary trends among bacterial mutualists of insects, *Genome Res.* 15 (2005) 1023–1033.
- [101] A. Sola-Landa, J. Pizarro-Cerda, M.J. Grillo, E. Moreno, I. Moriyon, J.M. Blasco, J.P. Gorvel, I. Lopez-Goni, A two-component regulatory system playing a critical role in plant pathogens and endosymbionts is present in *Brucella abortus* and controls cell invasion and virulence, *Mol. Microbiol.* 29 (1998) 125–138.

Mónica Medina is a marine biologist studying coral-microbial interactions. Her group is trying to address questions relevant to the ecology and evolution of the endangered coral reef ecosystems using a suite of tools ranging from genomics to field biology. Dr. Medina received her Ph.D. in 1998 from the University of Miami followed by postdoctoral positions at the Marine Biological Laboratory and the California Academy of Sciences. She was a research scientist at the Joint Genome Institute before joining the faculty at the University of California, Merced where she is currently an associate professor in the School of Natural Sciences.

Joel Sachs investigates the evolution of symbiotic bacteria. His research studies the selective forces and molecular mechanisms that shape bacterial cooperation with hosts as well as the origins and evolution of harmful strains. Empirical work focuses on rhizobial bacteria that nodulate the roots of legumes. Dr. Sachs received his Ph.D. in 2004 from the University of Texas and went on to the University of California at Berkeley where he received an NIH NRSA Postdoctoral fellowship. Joel Sachs is currently an assistant professor in the Department of Biology at the University of California, Riverside.