School of Doctoral Studies in Biological Sciences

University of South Bohemia in České Budějovice
Faculty of Science

Origins of bacterial endosymbionts in arthropods

Ph.D. Thesis

RNDr. Tomáš Chrudimský

Supervisor: prof. RNDr. Václav Hypša, CSc.
University of South Bohemia in České Budějovice, Faculty of Science

České Budějovice 2014

This thesis should be cited as:

Chrudimský T., 2014: Origins of bacterial endosymbionts in arthropods. Ph.D. Thesis. University of South Bohemia, Faculty of Science, School of Doctoral Studies in Biological Sciences, České Budějovice, Czech Republic, 78 pp.

Annotation

Current bioinformatic methods such as molecular phylogenetics and phylogenomics provide us with good insight to symbiont evolution. Though modern science evolves rapidly, accelerates speed of acquiring novel discoveries and improves their quality, there is still endless row of questions waiting to be answered. This thesis focuses on origins of symbiosis between insects and Enterobacteria, and the mechanisms promoting association of bacteria with arthropods. The main emphasis is put on the secondary symbionts of the genus *Sodalis* (Enterobacteriaceae) and the pathogenic *Anaplasma phagocytophilum* (Anaplasmataceae) that seems to be undergoing first steps to become hereditary mutualist.

Declaration [in Czech]

Prohlašuji, že svoji disertační práci jsem vypracoval samostatně pouze s použitím pramenů a literatury uvedených v seznamu citované literatury.

Prohlašuji, že v souladu s § 47b zákona č. 111/1998 Sb. v platném znění souhlasím se zveřejněním své disertační práce, a to v nezkrácené podobě elektronickou cestou ve veřejně přístupné části databáze STAG provozované Jihočeskou univerzitou v Českých Budějovicích na jejích internetových stránkách, a to se zachováním mého autorského práva k odevzdanému textu této kvalifikační práce. Souhlasím dále s tím, aby touto elektronickou cestou byly v souladu s uvedeným ustanovením zákona č. 111/1998 Sb. zveřejněny posudky školitele a oponentů práce i záznam o průběhu a výsledku obhajoby kvalifikační práce. Rovněž souhlasím s porovnáním textu mé kvalifikační práce s databází kvalifikačních prací Theses.cz provozovanou Národním registrem vysokoškolských kvalifikačních prací a systémem na odhalování plagiátů.

České Budějovice, 21. 2. 2014 .....

Tomáš Chrudimský

This thesis originated from a partnership of Faculty of Science, University of South Bohemia, and Institute of Parasitology, Biology Centre of the ASCR, supporting doctoral studies in the Molecular and Cell Biology and Genetics study programme.

Financial support

Grant Agency of Academy of Sciences of the Czech Republic

grant P505/10/1401

Ministry of Education, Youth and Sports of the Czech Republic

grants LC06073, LM2010005 and MSM6007665801

Grant agency of the University of south Bohemia in České Budějovice

grant 097/2010/P

Acknowledgements

I would like to express my gratitude to Professor Václav Hypša and members of the Laboratory of Molecular Phylogeny and Evolution of Parasites as well as to other people who deserve it. Special thanks belong to my current employer for allowing me to finish this thesis in working hours. I'm also very grateful to my parents for their support of all kinds. Last but not least I want to thank Terezka for her love and patience.

List of papers and author's contribution

The thesis is based on the following papers (listed chronologically):

I. Husník F., Chrudimský T., Hypša V., 2011. Multiple origins of endosymbiosis within the Enterobcateriaceae (γ-Proteobacteria): convergence of complex phylogenetic approaches. BMC Biology Dec 28;9:87. doi: 10.1186/1741-7007-9-87. (IF=6.53) . *Tomáš Chrudimský compiled and analyzed the AT/GC reduced matrices.*

II. Chrudimský T., Husník F., Nováková E., Hypša V., 2012. *Candidatus* Sodalis melophagi sp. nov.: Phylogenetically Independent Comparative Model to the Tsetse Fly Symbiont *Sodalis glossinidius*. PLoS One. 2012;7(7):e40354. doi: 10.1371/journal.pone.0040354. Epub 2012 Jul 17. (IF=3.73) *Tomáš Chrudimský annotated and analyzed T3SS sequences and participated on writing the paper, phylogenetic analyses, isolation and culturing of the S. melophagi and light and electron microscopy.*

III. Víchová B., Majláthová V., Nováková M., Stanko M., Hviščová I., Pangrácová L., Chrudimský T., Čurlík J., Peťko B., 2013. *Anaplasma* infections in ticks and reservoir hosts from Slovakia, Central Europe. Infection, Genetics and Evolution. In press. (IF=2.768) *Tomáš Chrudimský performed phylogenetic analyses.*

Contents

1. Introduction and results

1.1 General introduction

The term Symbiosis was coined in sense as “living together of unlike organisms” for the first time in 1879 by the founder of modern mycology and plant pathology Heinrich Anton de Bary in his book “Die Erscheinung der Symbiose”. Symbiosis is really widespread phenomenon that played and plays a critical role in Earth’s biosphere. Without symbiosis the life would not be as we know it since there would be no eukaryotes [1].
Symbiosis as a long term association of different organisms constitutes whole spectrum of ecological interactions (i.e. mutualism, commensalism and parasitism), however, the term symbiosis is often used only in connection to mutualists and commensals.

The phylum arthropoda comprises broad range of life forms and is the most numerous part of the kingdom animalia [2]. Arthropods host variety of bacteria of which many are of medical importance for animals and humans, and use parasitic arthropods as vectors MS3. In this thesis, I will mainly focus on the endosymbionts in insects MS 1,2 and ticks MS3 since intracellular symbiotic bacteria are rarely studied in other arthropods.
Endosymbiotic bacteria are usually treated as two different types: so called P (primary) and S (secondary) symbionts. This division is mainly based on physiological and ecological aspects and is unrelated to the bacterial phylogeny/taxonomy.

1.1.1 P symbionts

P symbionts are intracellular bacteria (except for the symbionts of Platasipidae [3]), strictly vertically transmitted from mother to offspring. These symbioses are usually ancient and the host-bacterium association undergoes millions of years of coevolution [4]. P symbionts are not distributed randomly in their host’s bodies. The insect hosts usually keep them restricted to specialized cells called bacteriocytes, often forming an organ called bacteriome. The role of P symbionts seems to be nutritional. They, for example, provide essential amino acids [5], vitamins [6-9] or recycle uric acid [10]. P symbionts appear to be essential for hosts survival and/or normal development. The loss

of P symbiont causes sterility, shortened lifespan or growth and developmental disorders [11,12].

The long term association of P symbionts with their hosts led to severe genome reductions and modifications. All unnecessary genes and metabolic pathways were eliminated, leaving the bacteria in state of “host-controlled metabolic factories”. This is probably the main reason, why all known P symbionts are uncultivable *in vitro*. The absolute record in genome reduction holds *Nasuia deltocephalinicola* (Betaproteobacteria), symbiont of the leafhopper *Nephotettix cincticeps* with 112 kbps genome encoding for 137 proteins [13] which has recently beaten another genome dwarf *Tremblaya princeps* (Betaproteobacteria) with 139 kbps genome that encodes just for 116 proteins [14]. These numbers are indeed comparable with characteristics of plastid organellar genomes [15] and stress the question about borders between organism and organelle.

1.1.2 S symbionts

Unlike obligate P symbionts, S symbionts are facultative and their loss is not fatal for their hosts; the hosts just may lose some ecological advantage [16]. Besides beneficial microbes this group also includes commensals and reproductive manipulators. The role of beneficial symbiont is usually protection against infections [17] and heat stress [18] or complementing missing parts of the P symbionts metabolic pathways [19,20]. In some cases they also can replace the P symbiont and overtake its function in the host e. g. [21,22]. Transmission of the S symbionts is not exclusivelly vertical, as is case of the P symbionts, since they can also spread horizontally [23-28]. Mechanisms of these horizontal transfers remain unknown. Since S symbionts are not so tightly associated with their hosts metabolisms, their *in vitro* cultivation may be successful in some cases [29-33]; and MS2).

1.2 Diversity of endosymbiotic bacteria in ticks and insects

1.2.1 Endosymbiotic bacteria of insects

It is supposed that majority of insects harbor symbiotic bacteria (see Table 1) that often increase their fitness and improve adaptation to the environment. Symbionts are also

known as important driving force of insect evolution since they represent an extra source of genetic information with mutation potential [16]. A good example are aphids and their mutualistic symbionts. Aphids are widespread sap-feeding insects that typically (but see [34,35]) host P symbiont called *Buchnera aphidicola*. *Buchnera* provides essential amino acids [5] and riboflavin [9] that are missing in sap and which the aphids cannot synthesize. Of the S symbionts, these aphids can host *Hamiltonella defensa* and *Regiella insecticola* which protect their hosts against parasitic wasps [36]. Though insect-bacteria symbiosis is extensively studied, there are still groups of insects, where symbiotic bacteria (P and S) were only identified by microscopic observation [37] and were never examined by modern biological methods. Identity of these morphologically described symbionts is therefore still unknown. These insect groups are largely coleopterans (namely Bostrychidae, Cerambycidae, Chrysomelidae, Lyctidae, Nosodendridae, Throscidae), but also Amblycera (chewing lice), partially Membracidae (tree hoppers) and Ceratopogonidae (biting midges). As part of my work I tried to obtain and examine insect samples from these groups. However, I was successful just in Throscidae. By using PCR based 16S rDNA screening and temperature gradient gel electrophoreis that allows for separation of fragments of the same size but different GC content, I identified three symbiotic bacteria in *Trixagus meybohmi* (Coleoptera, Throscidae). First of them is a novel P symbiont (according to very low GC content) that appears to be closely related to the P symbiont of sharpshooters *Sulcia muelleri* (Bacteroidetes). In addition, I have identified novel species of the genus *Sodalis* (Gammaproteobacteria) and one species of the genus *Wolbachia* (Alphaproteobacteria). These data are unpublished.

Insect order	Symbiotic bacteria (unnamed symbionts are ommited)
Blattaria	*Blattabacterium cuenoti*
Coleoptera	*Curculioniphilus buchneri, Macropleicola, Nardonella, Sodalis*
Collembola	*Wolbachia*
Diptera	*Arsenophonus, Aschnera chinzeii, Wigglesworthia glossinidia, Sodalis*
Hemiptera	*Asaia, Arsenophonus, Benitsuchiphilus tojoi, Brownia*

	rhizoecola, Baumannia cicadellinicola, Buchnera aphidicola, Cardinium, Carsonella ruddii, Ecksteinia, Gillettellia cooleyia, Hamiltonella defensa, Hodgkinia cicadicola, Ishikawaella capsulata, Kleidoceria schneideri, Liberibacter, Moranella endobia, Nasuia deltocephalinicola, Portiera aleyrodidarum, Profftia, Purcelliella pentastirinorum, Regiella insecticola, Rickettsia, Rickettsiella, Rohrkolberia cinguli, Rosenkranzia clausaccus, Serratia symbiotica, Schneideria nysicola, Sodalis, Sulcia muelleri, Spiroplasma, Tremblaya, Uzinura diaspidicola, Vallotia, Vidania fulgoroideae, Wolbachia, Zinderia insecticola
Hymenoptera	*Bartonella* – like, *Blochmannia, Sodalis, Streptomyces philanthi*
Isoptera	*Blattabacterium, Wolbachia*
Lepidoptera	*Arsenophonus, Wolbachia*
Phthiraptera	*Legionella, Riesia pediculicola, , Sodalis*
Protura	*Cardinium*-like, *Rickettsia* –like
Psocoptera	*Rickettsia*
Siphonaptera	*Rickettsia*,*Wolbachia*
Thysanoptera	*Stammerula tephritidis*

Table 1: List of insect endosymbionts

1.2.2 Endosymbiotic bacteria of ticks

Ticks are obligate hematophagous ectoparasites of mammals, birds, and reptiles. They can carry and transmit a wide range of pathogens, such as bacteria, protozoa, viruses, and nematodes that can cause various diseases of humans and animals [38].

There are several routes, by which ticks may become infected with bacteria, namely by feeding on bacteremic reservoirs and by transstadial or transovarial transmission. Some bacterial species are able to use all these transmission routes [39].

Among the epidemiologically most important bacteria are spirochetes from the *Borrelia burgdorferi* sensu lato complex, members of the family Anaplasmataceae, *Bartonella* spp., *Rickettsia* spp., *Coxiella burnetii* and *Francisella tularensis*.

Most of these bacteria are maintained in natural cycles that involve ticks and various species of reservoir hosts, and are responsible for diseases which are recognized as zoonozes. For each bacterial disease, one or several tick vectors and reservoir hosts may exist in the natural foci [39].

As the tick-transmitted bacteria present risk to human and animal health, research of ticks and their bacterial flora is dominantly focused on this perspective. Of the tick associated bacteria, I have worked on the genus Anaplasma. *Anaplasma phagocytophilum* is a Gramnegative bacterium that causes granulocytic anaplasmosis in humans and animals. It multiplies in a phagosome of infected neutrophils [40,41]. Ticks of the genus *Ixodes* (but there are also other *Anaplasma* transmitting genera) serve as vectors of the disease. In MS3 we mainly tried to explore epizoonotiological situation in Slovakia and map reservoir species involved in *A. phagocytophilum* circulation.

Little is known about possible benefits that the ticks obtain from the hosted bacterial flora. Study on a *Coxiella*-like bacterium in *Amblyomma* showed that antibiotic treated aposymbiotic ticks hadlower fitness than the untreated ticks [42]. In 2006, an intra-mitochondrial bacterium now called *Midichloria mitochondrii* was described from the hard tick *Ixodes ricinus* [43]. It was shown that bacterium consumes mitochondria in oocytes. However, despite bacteria destroying many mitochondria, oocytes develop normally. Although the genome sequence did not clearly answer the question whether *Midichloria* is parasite or mutualist, some genomic features indicate that it might be more beneficial than detrimental for its host. For example it can theoretically synthesize heme (ticks have to get it from blood meal) and B vitamins, and is able to provide additional ATP to the host [44]. Other research group showed that *Anaplasma phagocytophylum* enhances freeze tolerance of the *Ixodes scapularis* ticks helping them to survive in the cold weather [45].

Is the association of *Midichloria* with its host parasitism or mutualism? Does *Midichloria* pay for its transmission by synthesizing essential nutrients, which help ticks to survive starvation? What are biological roles of other bacteria associated with ticks? These and many other questions need to be answered. Fortunately, the pathogen-oriented research focus seems to be little bit shifting at least to bacterium-bacterium interactions and their consequences for a tick host, which is noticeable in recent reviews [46,47].

1.3 Widely distributed and emerging lineages of symbionts

1.3.1 The genus *Wolbachia*

Wolbachia is a genus of Alphaproteobacteria that belongs together with pathogenic *Rickettsia*, *Anaplasma* and *Ehrlichia* to the order Rickettsiales. Based on phylogenetic reconstructions, *Wolbachia* splits into so called "supergroups" A-H. *Wolbachia* is widely distributed bacterium that unlike other members of Rickettsiales does not directly infect vertebrates. Supergroups A, B, F and H are associated with arthropods and have been reported from insects [48], mites [49,50], spiders [51] and crustaceans [52].

The role of *Wolbachia* in arthropods is usually negative. They are reproductive parasites that change sex ratio of the offspring to gain maximal fitness. Generally, little is known about possible positive effects of otherwise detrimental bacteria on their hosts, but even such effects cannot be ruled out. There is just a little step from parasitism to mutualism and "helping bad guy" might be a state of transition between these two ecological strategies. *Wolbachia* was for example shown to protect *Drosophila* species against RNA viruses [53-56] and reduce viral infection rate in mosquitoes [57,58]. Furthermore, *Wolbachia* largely behave as mutualist in filarial worms [59] and there are also several arthropod groups where it developed into obligate mutualistic symbionts, e.g in bedbugs [60] and parasitic wasps [61].

1.3.2 The genus *Cardinium*

Cardinium belongs to the Bacteroidetes group and was discovered for the first time in tick cell line [62]. Unlike *Wolbachia*, no *Cardinium* lineage has proved positive effect on their arthropod hosts. They are all reproductive manipulators [63]. *Cardinium* prevalence is rather low in all arthropods, numbers range from 4.4 to 7.2% [63,64] , however, in spiders and mites prevalence is much higher 22 – 31.6% [63,65]. Research on *Cardinium* is relatively short and we can perhaps expect discovery of mutualistic *Cardinium* lineage(s) too.

1.3.3 The genus *Spiroplasma*

Spiroplasmas (Mollicutes) are descendants of Grampositive bacteria that lost their cell walls. They are infectious agents of plants and arthropods. In arthropods, they are

largely commensals, however some appear to be reproductive manipulators [66,67] and several mutualistic lineages have also been reported [68,69].

1.3.4 The genus *Sodalis*

The genus *Sodalis* belongs to the family Enterobacteriaceae, Gammaproteobacteria. Though many of its isolates are called "*Sodalis*-like endosymbiont of..", it is usually considered a regular monophyletic genus. Until recently (MS2), there was only one cultivable and described species – S symbiont from the tsetse flies, *Sodalis glossinidius* [30]. Ecology of *Sodalis* spans from commensals to obligate mutualists [70,71], however, their role is largely uncharacterized in other hosts. *Sodalis* was identified in wide spectrum of insects comprising tsetse flies [30], keds MS2 [72], psyllids [26], mealybugs [27], chewing lice [73], ants [74], aphids [75], scale insects [76], stinkbugs [77,78], cerambycid beetles [79], weevils [80-83].

1.3.5 the genus *Arsenophonus*

Arsenophonus is a species-rich genus of Enterobacteria infecting wide range of arthropod hosts and it was also identified as a plant pathogen. The genus is a monophyletic assemblage of various isolates, sometimes described as different genera [84]. Consequences of the *Arsenophonus* symbioses for the hosts are as variable as the *Arsenophonus* diversity itself. The ecology spans from reproductive manipulators [85-87] [88] and plant pathogens [89,90] to countless S symbionts of unknown roles e. g. [91] and P symbionts [92-94]. Particularly interesting is a reported putative defensive role of *Arsenophonus* in psyllid *Glycaspis brimblecombei* [95].

1.3.6 The genus *Rickettsia*

Rickettsia are generally known as pathogens of vertebrates that are transmitted by arthropod vectors and are causative agents of diseases such as spotted fever or typhus [39]. Although the research is mainly focused on *Rickettsia* with medical importance and their blood-feeding vectors, there are also species that play various known biological roles in their hosts. They may be obligate mutualists [96], facultative mutualists e.g. [97] or reproductive manipulators [98-101], however, the effect of *Rickettsia* on its host is unknown in most cases.

1.4 Biology versus phylogeny and phylogenetic problems

1.4.1 Phylogeny and biological interpretation

Inferring robust phylogeny is a crucial step toward biological and evolutionary interpretations (for example for identification of a strict symbiont-host coevolution found in P symbionts, or detection of host switches). An example can be provided by the genus *Sodalis* with wide distribution among insects. In MS2 we describe novel *Sodalis* species *Candidatus* Sodalis melophagi found in the sheep ked *Melophagus ovinus* (Hippoboscidae). Our phylogenies demonstrate that *Candidatus* Sodalis melophagi, a typical S symbiont, established its symbiosis independently of other two hippoboscoid-derived *Sodalis* lineages (including the *S. glossinidus* lineage known from tsetse fies). This incongruence reflects an early stage of the symbiosis and horizontal transfer, similar to some cases of coleopteran and homopteran hosts.

1.4.2 Phylogenetic obstacles in symbiont research

Correct phylogenetic inference of relationships among symbiotic bacteria appears to be a nut to crack at least by current phylogenetic and phylogenomic methods. This is especially true for P symbionts with highly degraded genomes, high base compositional bias and fast evolving sites [16].

In general, intracellular lifestyle and vertical transmission lead to severe bottlenecks and prevent exchange of genetic information, which speeds up the Muller's rachet [102] by which mutations are being accumulated and fixed [103]. Vertical transmission means small population size, which results in increased genetic drift and weak purifying selection. This accelerates sequence evolution of all genes in the genome. At the beginning of symbiotic state there is a burst of phage and transposable activity, which results in genome remodeling. Non-essential genes are pseudogenized and lost, portions of the genomes are relocated and large deletions take place. Therefore typical S symbiont possesses numerous pseudogenes, transposons and phages/phage derived sequences; whereas typical P symbiont has tiny genome without transposons and phages and has high coding density [104-106]. Bacteria have a natural tendency to higher GC content in their DNA, while mutations are universally biased toward A/T [107,108]. In reduced genomes of obligate P symbionts, there is a general trend towards high A/T contents. For example, the most A/T rich known genome has been found in *Zinderia*

insecticola. The genome size is 209 kbp and A/T content is 86,5% [109]. This is probably due to loss of genes involved in DNA reparation [16] in combination with weak purifying selection as part of Muller's rachet [110]. As usual, every rule has its exceptions. Not all P symbionts with reduced genomes have high A/T content. This is true for two Alphaproteobacteria- *Tremblaya princeps* and *Hodgkinia cicadicola*, with A/T content 41.2 and 41.6, respectively [20,111]. There must be some G/C mutation bias in Alphaproteobacteria that persisted in these two reduced genomes [111].

The most symbiont-rich group of bacteria is undoubtedly Enterobacteriaceae (Gammaproteobacteria) that also comprises pathogens of humans (such as *Salmonella* or *Yersinia*) and is therefore intensively studied. The 16S rDNA is a gene of choice (frequently the first and the only choice) with good taxon sampling in Enterobacteriaceae, however, it was shown as useless for inferring reliable phylogeny [16]. Unfortunately, many symbionts are represented only by 16S rDNA in Genbank. It is therefore important to shed light on relationships among Enterobacteria and produce more robust phylogenies. The number of symbiotic events in Enterobacteriaceae and monophyly/polyphyly of P symbionts became center of debates. As mentioned above, in P symbionts the rapid gene evolution and compositional bias causing homoplasies make inferring phylogenies very difficult. About 20 studies [112-131] tried to solve P symbionts monophyly/polyphyly by different approaches prior to our MS1. Of these studies, only few broke monophyly of P symbionts. Generally, the usage of non-homogenous models appeared particularly effective [113,118]. The "standard" models of sequence evolution assume composition homogeneity among taxa [132], which is definitely invalid assumption in case of P symbionts. Other researchers, who broke the P symbiont monophyly, used genomic data and performed break-point and inversion distance phylogeny [117]. And finally, they successfully coped with phylogenetic artifacts using "telescoping multiprotein phylogenetic analysis" [133]. Unfortunately, especialy in the case of older studies there is insufficient taxon sampling due to low genome data availability in Genbank. In 2011 we publihed MS1 where we used most common methods of removing phylogenetic artifacts. We analyzed data on both nucleotides and amino acids levels and looked for similar paterns in resulting phylogenetic trees. Since we obtained very similar results with different methodologies, it makes our conclusions quite robust. While for example widly used Slow-Fast method [134] failed, we retreived good results using nonhomogeneous model and models accounting for site-specific features of protein evolution (CAT and CAT+GTR) [135],

various matrix recodings and our novel AT/GC method. Since Slow-Fast [134] method removes fastest evolving sites, it does not remove large portion of homoplasies that occur from compositional bias. In our novel method, we focused on removing all sites that could be possibly biased. Therefore all sites that contained A or T along with C or G were removed, leaving all information only in A/T and G/C sites. After performing numerous phylogenetic analyses, we assumed, that symbiotic event occured at least 4 times in Enterobacteriaceae.

1.5 Genomic traits and mechanisms important for symbiosis

This part of introduction and results is focused on genetic features and mechanisms that I have studied and which play, in my opinion, important roles in symbiotic relationships. The connection among these phenomena is transfer of genes or gene clusters. Transfer of genetic material is ubiquitous in bacteria and represents important driving force in their evolution. By acquiring foreign information, bacterium can learn how to perform novel processes or synthesize novel compounds with biological activity. This may result in new opportunities and open new niches. Genetic information can be transmitted by several ways: transformation, transduction and conjugation [136].

1.5.1 Phages-bacterial viruses introducing new genetic features

While lytic phages use aggressive strategy of multiplication and destroy the host cell in short time period after infection, lysogenic phages can repress lytic cycle and incorporate into bacterial chromosome. In certain conditions they may decide to multiply and subsequently lyse the host cell. These so called temperate phages additionally prevent hosting bacteria of being colonized by other phages (which is beneficial for both partners) [137]. Along with their detrimental effects on bacteria (e. g. lysis), bacteriophages constitute an important way of genetic transfer and therefore bacterial evolution. They may accidentally pack parts of bacterial genome during capsid assemblage and inject it into different cell. Foreign bacterial DNA may be incorporated into the chromosome by mechanisms of homologous recombination. This process is called the transduction [138]. In addition, temperate phages can bring novel information also in a different way. When sitting in the host cell with repressed lytic cycle they allow for expression of gene cassette they may encode. This cassette contains genes that

are beneficial for the host and improves its fitness. The best documented viruses that behave this way are APSE phages infecting the Gammaproteobacterium *Hamiltonella defensa*.

History of research on APSE phages is relatively short. The first described, isolated and sequenced phage was bacteriophage APSE-1 (the abbreviation means the first bacteriophage of *Acyrthosyphon pisum* secondary endosymbiont) in 1999 [139]. The host bacterium was identified as an agent that protects aphids from being infected by parasitic wasps [140]. However, it was later demonstrated, that it is not *H. defensa* itself but the APSE phage that is responsible for protective phenotype [141]. In 2008, seven APSE phages were partially sequenced. The key difference was in content of so called toxin cassette, where e.g. YD-repeat toxin, Shiga-like toxin or cytolethal distending toxin were identified [142,143]. Content of a toxin cassette influenced degree of aphid protection. For example APSE-3 phage encoding YD-repeat toxin guaranteed more than 85% survival rate, while APSE-2 phage encoding cytolethal distending protein was effective in 40% [144] . Although protective phenotype of *H. defensa* was connected with the phages, the exact mechanism of toxin delivery to the parasitic wasp remains unknown. As every rule has its exceptions and nothing is 100% certain in biological systems, Adam Martinez and colleagues recently discovered two *Hamiltonella*+APSE-2 infected aphid lines that do not possess better protection than the same lines treated with antibiotics – i.e. with *Hamiltonella* eradicated. Furthemore and most interestingly, they described existence of parasitoid-resistant aphid lines that are S symbiont free [145]. Authors are investigating now how these aphid lines can fight the parasitoids without protection of APSE.

In terms of phage research, I have sequenced, assembled and annotated complete genomes of two APSE-3 phages from two different aphid lineages. In one of these lineages the phage was repeatedly lost in offspring from time to time, but it was steadily inherited in the second aphid lineage. Unfortunately, these phage sequences brought more questions than answers, since they were identical except for three changes in the toxin gene. Since PCR did not detect the “lost phage” both integrated in chromosome and circular in cytosol (i.e. the phage was completely erased from *H. defensa*) it remained rather enigmatic, how and why is a virtually identical phage occasionally eradicated in one aphid lineage and maintained in the other.

1.5.2 Secretion systems, focus on T3SS

Secretion systems are protein machineries that are designed to deliver proteins into surroundings of a bacterial cell. The key player in pathogenesis and symbiosis appears to be the so called Type III secretion system (T3SS) [146]. T3SS typically occur in pathogens such as *Salmonella, Yersinia* or *Chlamydia* that cause disease in humans and animals and is used to transfer bacterial proteins called the effectors into the host cells. These effectors are either encoded directly in the same genomic island as T3SS (these islands are frequently horizontally transferred) or they can be found anywhere in the genome, which complicates detection and research of these interesting molecules. Secretion of effectors typically allows for cell cycle modulation, cytoskeleton control, cell death induction or hiding out from the immune system [147]. Though living in different conditions, plant pathogens use T3SS for very similar purposes [148]. As lots of active molecules have not been discovered so far, it appears that functions of bacterial effectors in host organism is probably more complex. While P symbionts typically does not encode for T3SS, it is found in S symbionts (and not only in arthropods). For example it has been detected in symbioses: *Sodalis glossinidius*/tsetse fly, *Hamiltonella defensa*/aphids. In case of *Sodalis glossinidius*/tsetse fly association, it has been experimentally demonstrated that T3SS is required for successful transmission of a symbiont on insect progeny [149]. This is in perfect agreement with experiment on non-insect symbiosis model – mutualistic *Aeromonas*/leech, where T3SS mutants were not able to colonize the host [150]. This suggests an essential role of T3SS in S symbionts life. In MS2 we compared T3SS of two *Sodalis* species *S. glossinidius* and *S. melophagi*. While *S. glossinidius* encodes for three copies: SSR-1, SSR-2 and SSR-3, *S. melophagi* possesses only SSR-3 copy and SSR-2 is highly degraded. Since SSR-3 is conserved in both species it can be speculated that *S. glossinidius* could use other two specialized copies with their associated secreted proteins for other distinct purposes. Recently, two additional genomes of *Sodalis* species became publicly available, therefore I added their secretion systems to the comparison (see Figure 1). SOPE is a P symbiont of rice weevil *Sitophilus oryzae* [70]. *Sodalis* HS strain was isolated from a wounded man who impaled his hand on a branch of his apple tree. This species represents important discovery, since it demonstrates the way insects may acquire their symbionts, e.i. from plant pathogens [151]. This is noticeable in MS1 where symbiotic bacteria cluster with pathogens (see Discussion and Conclusions part for more details).

Plants also serve as places where plant feeding insects get infected by already established S symbionts [152].

The Figure 1 reflects connection of T3SS degradation with intimacy of the symbiont/host assotiation. All but *Sodalis* HS strain and *S. glossinidius* lost SSR-1 copy of secretion system. These two species retained all three copies. SSR-2 and SSR-3 of SOPE started their degradation and is questionable whether they are still functional or not. This fact is in alignment with state of SOPE that represents very young P symbiont.

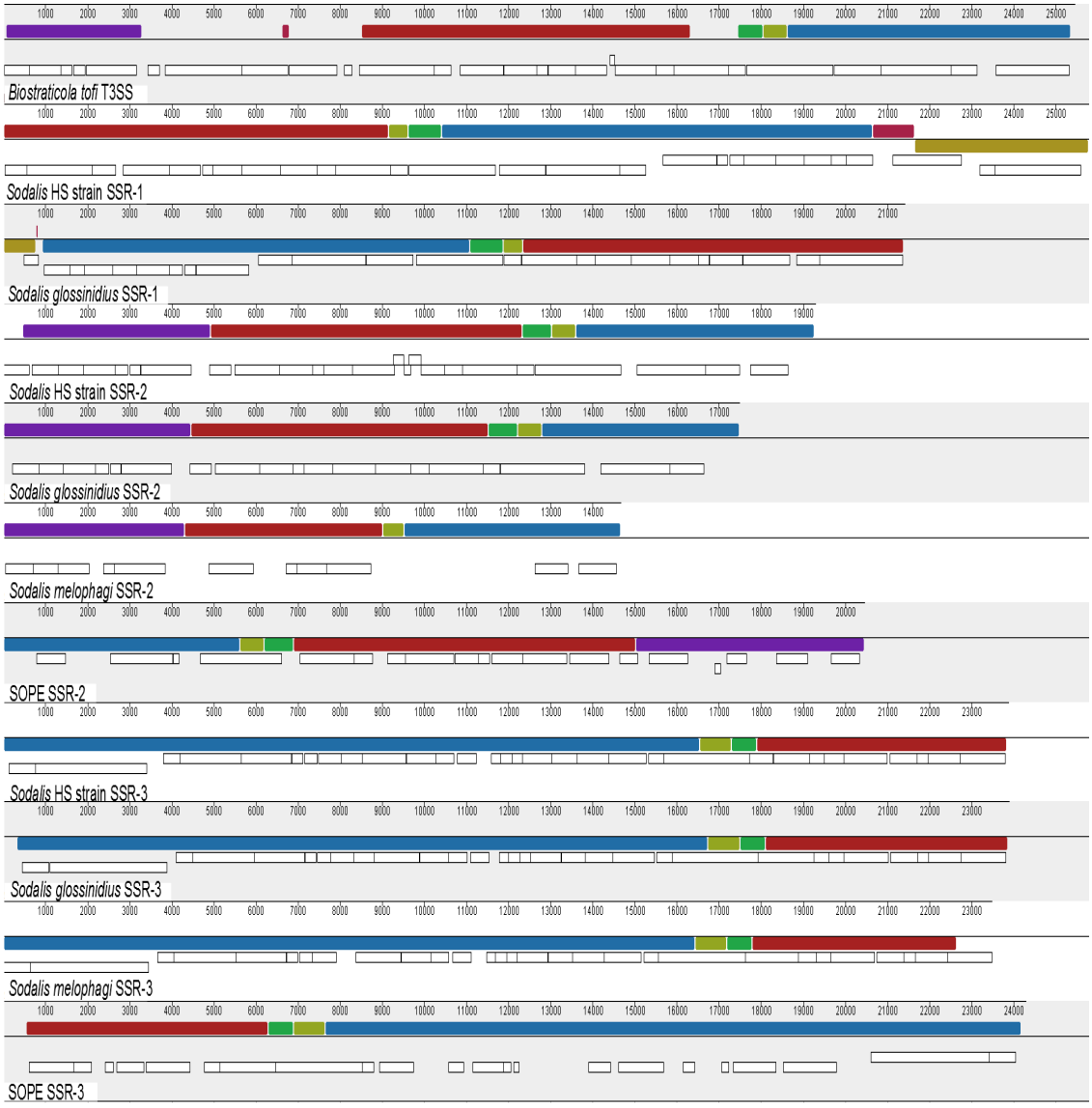


Figure 1: Comparison of T3SS copies using MAUVE [153] software. *Biostraticola tofi* is the closest known relative to the genus *Sodalis*.

1.5.3 Plasmids

Plasmids are dsDNA molecules that replicate independently on chromosomal DNA. They are generally circular, however, for example in case of *Borrelia*, linear forms occur. Their size varies from 846 bp to 1.8 Mbp and they may encode from 1 up to 7281 genes [154,155]. Plasmids usually carry genes that are not essential for survival of bacterium (e.g. housekeeping genes), however, these plasmid encoded genes provide some extra benefits increasing bacterial fitness (e.g. antibiotic resistance [156], bacteriocides [157], ability of conjugation [158], virulence[159], degradation of toxic substances [160], heavy metal resistance [161]).

In order to compare plasmids of *Sodalis melophagi* with *Sodalis glossinidius,* I sequenced and annotated extrachromosomal DNA of *S. melophagi*. Extrachromosomal DNA of *S. glossinidius* comprises three plasmids and one circular phage. The plasmids encodes for iron uptake via siderophors, putative toxins, hemolysins and proteases [162]. Although belonging to the same genus and living in very similar conditions, *S. glossinidius* and *S. melophagi* does dramatically differ in their plasmid contents. *S. melophagi* also possessess four circular extrochromosomal molecules (see Table 2), of which one is the phage (see phage comparison Figure 2), however, their sizes and content are dissimilar. Plasmids of *S. melophagi* encode lots of phage and transposon derived proteins and enzymes located on plasmids are often present also on the chromosome. It is therefore difficult to assign them any clear functions except of one that contains the whole operon for Type 1 fimbriae synthesis.

	pSM1	pSM2	pSM3	phiSM
GC content	46.7%	43.2%	48.4%	50.6%
Size	33.4	28.5	36.2	38.2
Protein coding genes	25	25	38	41

Table 2: Molecular characteristics of extrachromosomal DNA in *Sodalis melophagi*

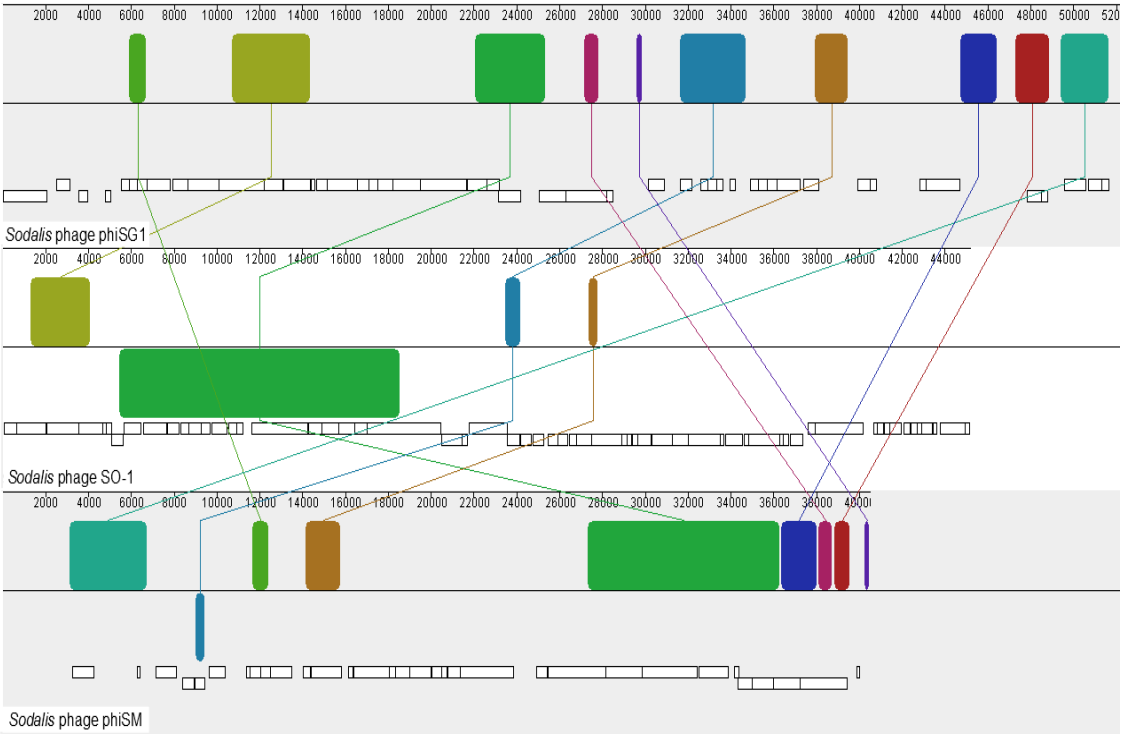


Figure 2: Genomic comparison of the *Sodalis* phages produced using MAUVE [153] software. The bacteriophages phiSG1 and SO-1 occur in the *Sodalis glossinidius*, while phiSM in the *Sodalis melophagi*.

Type 1 fimbriae are bacterial hair-like organelles present on cell surface that are assembled by chaperone-uscher secretion pathway. They occur in Gramnegative bacteria [163]. Among all bacteria, Type 1 fimbriae are most intensively studied in case of uropathogenic *Escherichia coli* causing disease in humans. The tips of Type 1 fimbriae made of FimH lectin are specific to mannosilated glycoproteins and allow for bacterial attachment to tissue cells [164,165]. This promotes bacterial colonization of urinary tract, cell invasion and constitutes the first step in biofilm formation [166]. Interestingly, Type 1 fimbriae also appears to be important player in forming of *E. coli* intracellular communities that are biofilm-like structures protecting bacteria against antibiotics and host immunity system. These protected communities allows for massive *E. coli* multiplication within the urinary bladder [167]. Given *Sodalis melohagi* is S symbiont, it can be speculated that *S. melophagi* uses Type 1 fimbriae for very similar purposes to *E. coli*. That is: 1- attachment to cell surfaces and body colonization and 2 - hiding from the immunity system. The fimbriae operon from *S. melophagi* is related to the operon from another Entorobacterium – *Erwinia*, which is (in most cases) a plant

pathogen that uses Type 1 fimbriae at least for attachment to plant surface [168]. Whether it is a case of horizontal gene transfer, or *Sodalis* is closely related to *Erwinia* cannot be unequivocally decided. According to our phylogenetic analyses MS1, the first case looks more probable.

Being aware of the fact that not all encoded genes are expressed and they therefore constitute cryptic pseudogenes, I verified physical presence of fimbriae experimentally by using yeast agglutination assay with positive result. As Type 1 fimbriae are present in *S. melophagi* on bacteriological plates we can assume that bacterium uses them in the host to.

2 Discussion and Conclusions

As I attempted to cover broad spectrum of arthropod/symbiont associated issues, outcomes of my work split into two main groups. First, the main one, deals with origins of symbiosis and genomic traits that are likely required for symbiosis establishment, while the second covers circulation of arthropod-borne pathogens between host and a vector.

2.1 Topic 1 - Rising of symbioses in Enterobacteria

In the MS1, we performed complex phylogenetic analyses based on modern methodology. Furthermore, we developed novel and powerful way of handlingsequences with extreme compositional bias. Long-branch attraction [169] due to fast evolution of some taxa in a dataset is certainly troublesome artifact, but the fast evolving sites are not necessarily the worst enemies of molecular phylogeneticians. Slowly evolving positions can be tricky too, because of their possible non-homogenous base composition. In our article, we demonstrated the power of this artifact by failure of Slow-Fast analysis [134] that removes fastest evolving sites. Similar results, i.e. failure of creation correct phylogenetic tree after removing fastest evolving sites, were also obtained by Beatricé Roure and Hervé Phillipe [170] in case of animal phylogeny based on mitochondrial genomes. They obtained the correct topology by removal of the so called heteropecillous sites – positions that change their substitution pattern over time. Why fast evolving sites constitute much smaller problem than non-homogenous site composition? That question was answered for example by Hervé Phillipe [171]. While phylogenetic models can (more or less) deal with different rates of evolution over sites via gamma distribution, they fail to recognize violation of model assumptions (like site non-homogeneity). Therefore it is crucial to know if a phylogenetic model you want to use is really suitable for the analyzed data.

Outputs computed from both amino acid and nucleotide datasets strongly suggest at least four independent origins of symbiosis within Enterobacteriaceae. This finding is supported by convergence of results from different ways of data analyses. Beside the

fact that rising of symbiosis was not a single event in Enterobacteriaceae, our topologies also show the putative origins of symbiotic lineages. They seem to recruit from pathogenic or gut associated bacteria. There is more detailed description of four symbiotic clusters in Enterobacteriaceae.

Buchnera clade consists of two different genera: *Buchnera*, which is an intracellular obligate mutualist with long parallel history with aphids [172], and *Ishikawaella capsulata*, gut endosymbiont of plataspid stinkbugs [173]. The placement of *Ishikawaella* together with *Buchnera* might still be an artificial, since other closely related bacteria are missing in the known sample [78,174-176] in the tree. The cluster is associated with the genus *Erwinia* which consist dominantly of plant pathogens. This position is in alignment with current state of knowledge, since *Erwinia* has been already suggested to be a predecessor of *Buchnera* [177].

Arsenophonus clade comprises S symbiont *Arsenophonus*, and *Riesia* [178], P symbiont of lice that is actually *Arsenophonus* misidentified as independent genus [84]. This association of P and S symbiont is not artificial and supports known fact that S symbionts can evolve into P symbionts over time. *Arsenophonus* clade is placed in a neighborhood of genera *Proteus*, *Photorhabdus* and *Xenorhabdus*. All these organisms are connected to insects in various ways. *Photorhabdus* and *Xenorhabdus* are insect pathogens living in association with entomopathogenic nematodes and playing crucial roles in their life cycles [179]. Members of the genus *Proteus* are saprophytes found in natural environments as well in animal and human intestines and feces. They often behave as opportunistic pathogens causing mainly infections of open wounds and urinary tract [180]. One of them, *Proteus myxofaciens*, recently reclassified as a separate genus *Cosenzaea* [181], was originally isolated and described from both living and death larvae of the gypsy moth *Porthetia dispar* [182]. Since ecology of this bacterium in the moth was never clarified, it can only be speculated whether it behave as mutualist, commensal or parasite.

Other symbiotic event is presented by the *Hamiltonella*/*Regiella* clade. These taxa are clustered with the genus *Yersinia* which is definitely closely associated with insects. The linkage of *Yersinia* with insects has negative consequences for the hosts. The most famous example is a causative agent of the plague *Yersinia pestis* that circulates between rats and fleas. The bacterium is well equipped for flea invasion and survival of defence response. It moreover manipulates the host via blocking the midgut and subsequent host starvation to increase the rate of flea biting [183,184]. There are also

other entomopathogenic species in this genus, like *Yersinia entomophaga* recently isolated from diseased larvae of the grass grub *Costelytra zealandica* [185,186].

Sodalis clade is an assemblage of S symbiont *Sodalis* and P symbionts of the genera *Baumannia*, *Blochmannia* and *Wigglesworthia*. Position of *Sodalis* on the base of P symbionts may indicate transformation of the S to the P symbiotic strategy [187]. Exact position of the clade is uncertain, however, it seems to be related to animal and plant pathogens of the genera *Pectobacterium*, *Dickeya* and *Edwardsiella*. Some species of the *Pectobacterium,* which is as a plant pathogen, uses plant feeding insects for spreading. They use other plant-associated invertebrates, like snails, too. There were genes identified in *Pectobacterium,* enabling bacteria to survive attacks of vectors immunity [188]. *Dickeya dadanti,* which behaves as a saprophyte or a plant pathogen, was shown to produce toxins causing septicemia in pea aphids, it is able to survive in the insect host and it is likely that these bacteria are specialized to use aphids for spreading from plant to plant [189,190].

It is probable that there were much more symbiotic events within Enterobacteriaceae, because our analyses comprised only limited number of taxa for which the data is availability. Since there is also lack of known free living relatives of pathogens and symbionts, it is difficult to reconstruct precise stories about origins of particular symbiotic lineages. The costs of genome sequencing became reasonable, which is definitely good for obtaining taxonomic data. We sequenced a genome of a free living bacterium *Biostraticola tofi* (unpublished data), originally isolated from the biofilm in a hard water rivulet [191]. *Biostraticola* is the closest free living relative of the genus *Sodalis*, which was shown in species description article [191] and also in MS2. Interestingly, it possess one copy of T3SS, which differs from *Sodalis* T3SS (see Figure 1) and is of the different origin. According to gene content, T3SS of the *Biostraticola* is complete and in addition, it contains known effector proteins, which suggests some kind of interaction with water-borne eukaryotes, probably pathogenesis.

Are there any prerequisites for a bacterium to become a symbiont? Indeed, yes. Symbionts seem to come from bacterial groups that are in frequent contact with the future host. For example, in cicadas, which spend many years in soil feeding on roots, their P symbiont called *Hodgkinia cicadicola* originated from Rhizobiales, which are bacteria associated with plant roots [192]. Another example is the *Sodalis* cluster. As discussed above, *Sodalis* clade is related with plant pathogens *Pectobacterium* and

Dickeya that learned how to use insects for spreading from plant to plant. Recently, the free living *Sodalis*, named *Sodalis* HS strain, was isolated from a wounded man who impaled his hand on an apple tree branch. Though its ecology is not known, *Sodalis* HS strain can be hypothetized to live as saprophyte or a plant pathogen [151]. This is in perfect alignment with *Pectobacterium* and *Dickeya* ecologies and suggests that *Sodalis* symbionts are probably domesticated plant pathogens as is the case of *Buchnera* in aphids [177].

Besides the first requirement of the symbiosis establishment I have described above, i.e.thr frequent physical contact, there are also some others. Not every free living bacterium that is in contact with insects will automatically become symbiotic. The second requirement can be called a genetic predisposition. Bacteria have to encode for genes and organelles that enables them to survive in insect body before they attempt to get and live inside it. These traits are largely connected with pathogenesis and can be transmitted horizontally. Insect body has evolved to cope with bacteria and is hostile to infection agents. After ingestion, majority of bacteria are killed by insect defence mechanisms or (and) expelled from digestive tract by peristalsis. While it might look difficult to survive there, the survival can be for example assured by single gene. *Erwinia* is protected by Erwinia virulence factor *evf* that still works if transferred to other Gramnegative bacterium [193]. The last steps to symbiosis are obtaining the ability to spread in the host body while evading the defence system and learning transovarial transmission.

Of the many mechanisms evolved to cope with insect defenses [194] I identified two to be present in *Sodalis melophagi*. The first one is T3SS and the second one is plasmid encoded fimbriae.

Though T3SS can be used as a weapon of free living bacteria against predators [195], it is typically associated with both plant and animal pathogenesis [147]. It looks like the T3SS also plays an important role in early stages of symbiosis (S symbionts and entomopathogens). The function of T3SS in insect-bacterium interaction does not seem to be connected with intestine environment (there are no supporting data), it is linked with bacterium spread inside the host body. *Photorhabdus luminescens*, which is an aggressive entomopathogen, uses T3SS to inhibit phagocytosis by insect haemocytes [196]. On the other hand *Sodalis glossinidius*, which is an S symbiont inhabiting tsetse flies, uses T3SS to spread into progeny. Though they were able to survive in a fly, T3SS mutants were unable to transmit from mother to offspring [149]. This difference of

T3SS usage between a pathogen and a commensal (or mutualist) is likely determined by different transmission strategies. *Photorhabdus* primary needs to survive confrontation with host defence, it kills an insect and is transmitted by a nematode, while *Sodalis glossinidius* does not kill the host and rather enter the host larva in uterus. As symbiotic association becomes closer and the S symbiont turns into the P symbiont, the transmission starts to be host-directed and T3SS machinery is therefore useless. In my opinion, such progressive T3SS degradation can be observed in the genus *Sodalis*, where free living *Sodalis* HS strain possesses three probably functional T3SS copies, while SOPE as a P symbiont does not need them anymore. *Sodalis glossinidius* and *Sodalis melophagi* constitute transition between these two states.

The primary function of the Type 1 fimbriae is to attach and hold bacterium on desired surface, biotic or abiotic. In relation to arthropods and their symbionts, fimbriae play an important role for example in life of entomopathogenic *Xenorhabdus*, because they enable the bacterium to colonize the nematode guts [197]. What is the reason for *Sodalis melophagi* to encode functional fimbriae? *S. melophagi* is already established S symbiont, it is already present in host tissues and it is transmitted from mother to offspring. There is therefore no free living life stage outside the host body. So, why *S. melophagi* may need to attach to the host tissue? In my opinion, the main purpose of fimbriae in *S. melophagi* might be direct transmission of the infection from ked to ked, which is a horizontal transfer. Sheep keds often bite one another and suck heamolymph. *S. melphagi* is present in the heamolymph and could be therefore easily transmitted orally from infected to uninfected individual. It uses fimbriae as an anchor to stop in host's intestines, similarly to *Xenorhabdus*, colonize them and spread into the rest of the body. I am not aware of the same behaviour in tsetse. This may provide an explanation why *S. glossinidius* does not need and therefore does not encode for Type 1 fimbriae.

2.1 Topic 2 – Ticks and *Anaplasma* in Slovakia

As discussed above, at the beginning of the symbiogenetic process which ends by P symbiont establishment (or its possible transformation into organelle), bacteria often behave as pathogens and use arthropods as vectors for spreading. *Anaplasma phagocytophilum* can be considered as one of these "biginers" that will (or will not) later switch to pure mutualism. But in my opinion, this process might already started. It

looks like *Anaplasma phagocytophilum* is not able to spread via transovarial transfer in ixodid ticks and rely mainly on ruminats and rodents as reservoir hosts [198-205]. However, the transovarial transfer was documented in *Dermacentor albopictus* [206]. Another step to mutualistic S symbiont is promoting beneficial effects to the host. These effects were indeed reported for *Ixodes scapularis* ticks, where *A. phagocytophilum* helps the ticks to survive in cold weather [45]. However, *A. phagocytophilum* still requires cycling among ticks and their hosts and is responsible for disease in animals and humans. The severity of clinical manifestation varies and is probably linked with variation of genome content among particular strains. Geography-dependent differences are also interesting. In the USA, granulocytic anaplasmosis caused by *A. phagocytophilum* is a common tick-borne disease, however, it is rare in Europe [207]. It seems that American strains and European strains have different characteristics in terms of ecology, host specifity and pathogenicity [208].

In MS3 we mainly focused on identification of species involved in circulation of *A. phagocytophilum* in Slovakia and searching for clustering patterns according to host species or geography. *A. phagocytophilum* was detected in wide spectrum of tested species and our results suggest strong position of ungulates, especially roe deer, as reservoir hosts. Interestingly, there is a conflict of our *groEL* tree topology with Silaghi et al. [209], where they imply existence of the “roe deer cluster”. This division into two clusters, where one of them comprises *A. phagocytophilum* isolated from roe deer and the other *A. phagocytophilum* isolated from other species, also appear in other European studies [210-213]. It cannot be unambiguously decided, if polyphyly of *A. phagocytophilum* from roe deer was an artifact (for any reason) or the confirmation of a general true that nothing lasts forever. It would require more than one isolate and better quality of input data to clarify this observation. Regarding our results on rodents, it must be highlighted that rodents in Europe are not likely carriers of *A. phagocytophilum* strains infecting humans, which is not true in the USA.

3 References

1. Williams TA, Foster PG, Cox CJ, Embley TM (2013) An archaeal origin of eukaryotes supports only two primary domains of life. Nature 504: 231-236.
2. Thompson JN (1994) The coevolutionary process.: University of Chicago Press. 383 p.
3. Fukatsu T, Hosokawa T (2002) Capsule-transmitted gut symbiotic bacterium of the Japanese common plataspid stinkbug, *Megacopta punctatissima*. Appl Environ Microbiol 68: 389-396.
4. Baumann P (2005) Biology bacteriocyte-associated endosymbionts of plant sap-sucking insects. Annu Rev Microbiol 59: 155-189.
5. Akman Gunduz E, Douglas AE (2009) Symbiotic bacteria enable insect to use a nutritionally inadequate diet. Proc Biol Sci 276: 987-991.
6. Goldberg C, Pierre LL (1969) Tyrosinase activity of the symbionts and fat bodies of the cockroach, *Leucophaea maderae*. Can J Microbiol 15: 253-255.
7. Heddi A, Grenier AM, Khatchadourian C, Charles H, Nardon P (1999) Four intracellular genomes direct weevil biology: nuclear, mitochondrial, principal endosymbiont, and *Wolbachia*. Proc Natl Acad Sci U S A 96: 6814-6819.
8. Zientz E, Beyaert I, Gross R, Feldhaar H (2006) Relevance of the endosymbiosis of *Blochmannia floridanus* and carpenter ants at different stages of the life cycle of the host. Appl Environ Microbiol 72: 6027-6033.
9. Nakabachi A, Ishikawa H (1999) Provision of riboflavin to the host aphid, Acyrthosiphon pisum, by endosymbiotic bacteria, *Buchnera*. J Insect Physiol 45: 1-6.
10. Mullins DE, Cochran DG (1972) Nitrogen excretion in cockroaches: uric Acid is not a major product. Science 177: 699-701.
11. Nogge G (1976) Sterility in tsetse flies (Glossina morsitans Westwood) caused by loss of symbionts. Experientia 32: 995-996.
12. Nogge G (1981) Significance of Symbionts for the Maintenance of an Optimal Nutritional State for Successful Reproduction in Hematophagous Arthropods. Parasitology 82: 101-104.
13. Bennett GM, Moran NA (2013) Small, smaller, smallest: the origins and evolution of ancient dual symbioses in a phloem-feeding insect. Genome Biol Evol.
14. Lopez-Madrigal S, Latorre A, Porcar M, Moya A, Gil R (2011) Complete genome sequence of "*Candidatus* Tremblaya princeps" strain PCVAL, an intriguing translational machine below the living-cell status. J Bacteriol 193: 5587-5588.
15. Barbrook AC, Howe CJ, Kurniawan DP, Tarr SJ (2010) Organization and expression of organellar genomes. Philos Trans R Soc Lond B Biol Sci 365: 785-797.
16. Moran NA, McCutcheon JP, Nakabachi A (2008) Genomics and Evolution of Heritable Bacterial Symbionts. Annual Review of Genetics 42: 165-190.
17. Brownlie JC, Johnson KN (2009) Symbiont-mediated protection in insect hosts. Trends Microbiol 17: 348-354.
18. Montllor CB, Maxmen A, Purcell AH (2002) Facultative bacterial endosymbionts benefit pea aphids *Acyrthosiphon pisum* under heat stress. Ecological Entomology 27: 189-195.

19. Lamelas A, Gosalbes MJ, Manzano-Marin A, Pereto J, Moya A, et al. (2011) Serratia symbiotica from the aphid *Cinara cedri*: a missing link from facultative to obligate insect endosymbiont. PLoS Genet 7: e1002357.
20. McCutcheon JP, von Dohlen CD (2011) An interdependent metabolic patchwork in the nested symbiosis of mealybugs. Curr Biol 21: 1366-1372.
21. Koga R, Bennett GM, Cryan JR, Moran NA (2013) Evolutionary replacement of obligate symbionts in an ancient and diverse insect lineage. Environ Microbiol 15: 2073-2081.
22. Koga R, Moran NA (2014) Swapping symbionts in spittlebugs: evolutionary replacement of a reduced genome symbiont. ISME J.
23. Fukatsu T, Nikoh N, Kawai R, Koga R (2000) The secondary endosymbiotic bacterium of the pea aphid *Acyrthosiphon pisum* (Insecta : Homoptera). Applied and Environmental Microbiology 66: 2748-2758.
24. Russell JA, Latorre A, Sabater-Munoz B, Moya A, Moran NA (2003) Side-stepping secondary symbionts: widespread horizontal transfer across and beyond the Aphidoidea. Molecular Ecology 12: 1061-1075.
25. Sandstrom JP, Russell JA, White JP, Moran NA (2001) Independent origins and horizontal transfer of bacterial symbionts of aphids. Mol Ecol 10: 217-228.
26. Thao ML, Clark MA, Baumann L, Brennan EB, Moran NA, et al. (2000) Secondary endosymbionts of psyllids have been acquired multiple times. Curr Microbiol 41: 300-304.
27. Thao ML, Gullan PJ, Baumann P (2002) Secondary (gamma-Proteobacteria) endosymbionts infect the primary (beta-Proteobacteria) endosymbionts of mealybugs multiple times and coevolve with their hosts. Appl Environ Microbiol 68: 3190-3197.
28. Thao ML, Baumann P (2004) Evidence for multiple acquisition of *Arsenophonus* by whitefly species (Sternorrhyncha: Aleyrodidae). Curr Microbiol 48: 140-144.
29. Pontes MH, Dale C (2006) Culture and manipulation of insect facultative symbionts. Trends in Microbiology 14: 406-412.
30. Dale C, Maudlin I (1999) *Sodalis* gen. nov. and *Sodalis glossinidius* sp. nov., a microaerophilic secondary endosymbiont of the tsetse fly *Glossina morsitans morsitans*. Int J Syst Bacteriol 49 Pt 1: 267-275.
31. Matthew CZ, Darby AC, Young SA, Hume LH, Welburn SC (2005) The rapid isolation and growth dynamics of the tsetse symbiont *Sodalis glossinidius*. FEMS Microbiol Lett 248: 69-74.
32. Sabri A, Leroy P, Haubruge E, Hance T, Frere I, et al. (2011) Isolation, pure culture and characterization of *Serratia symbiotica* sp. nov., the R-type of secondary endosymbiont of the black bean aphid Aphis fabae. Int J Syst Evol Microbiol 61: 2081-2088.
33. Dale C, Beeton M, Harbison C, Jones T, Pontes M (2006) Isolation, pure culture, and characterization of "*Candidatus* Arsenophonus arthropodicus," an intracellular secondary endosymbiont from the hippoboscid louse fly *Pseudolynchia canariensis*. Appl Environ Microbiol 72: 2997-3004.
34. Vogel KJ, Moran NA (2013) Functional and evolutionary analysis of the genome of an obligate fungal symbiont. Genome Biol Evol 5: 891-904.
35. Koga R, Tsuchida T, Fukatsu T (2003) Changing partners in an obligate symbiosis: a facultative endosymbiont can compensate for loss of the essential endosymbiont *Buchnera* in an aphid. Proc Biol Sci 270: 2543-2550.
36. Scarborough CL, Ferrari J, Godfray HC (2005) Aphid protected from pathogen by endosymbiont. Science 310: 1781.

37. Buchner P (1965) Endosymbiosis of animals with plant microorganisms.: Mycological Society of America.
38. Medlock JM, Hansford KM, Bormane A, Derdakova M, Estrada-Pena A, et al. (2013) Driving forces for changes in geographical distribution of *Ixodes ricinus* ticks in Europe. Parasites & Vectors 6.
39. Raoult D, Roux V (1997) Rickettsioses as paradigms of new or emerging infectious diseases. Clinical Microbiology Reviews 10: 694-&.
40. Dumler JS, Choi KS, Garcia-Garcia JC, Barat NS, Scorpio DG, et al. (2005) Human granulocytic anaplasmosis and *Anaplasma phagocytophilum*. Emerging Infectious Diseases 11: 1828-1834.
41. Carlyon JA, Fikrig E (2003) Invasion and survival strategies of *Anaplasma phagocytophilum*. Cellular Microbiology 5: 743-754.
42. Zhong J, Jasinskas A, Barbour AG (2007) Antibiotic treatment of the tick vector *Amblyomma americanum* reduced reproductive fitness. PLoS One 2: e405.
43. Sassera D, Beninati T, Bandi C, Bouman EA, Sacchi L, et al. (2006) '*Candidatus* Midichloria mitochondrii', an endosymbiont of the tick *Ixodes ricinus* with a unique intramitochondrial lifestyle. Int J Syst Evol Microbiol 56: 2535-2540.
44. Sassera D, Lo N, Epis S, D'Auria G, Montagna M, et al. (2011) Phylogenomic evidence for the presence of a flagellum and cbb(3) oxidase in the free-living mitochondrial ancestor. Mol Biol Evol 28: 3285-3296.
45. Neelakanta G, Sultana H, Fish D, Anderson JF, Fikrig E (2010) *Anaplasma phagocytophilum* induces *Ixodes scapularis* ticks to express an antifreeze glycoprotein gene that enhances their survival in the cold. Journal of Clinical Investigation 120: 3179-3190.
46. Ahantarig A, Trinachartvanit W, Baimai V, Grubhoffer L (2013) Hard ticks and their bacterial endosymbionts (or would be pathogens). Folia Microbiol (Praha).
47. Taylor M, Mediannikov O, Raoult D, Greub G (2012) Endosymbiotic bacteria associated with nematodes, ticks and amoebae. FEMS Immunol Med Microbiol 64: 21-31.
48. Hilgenboecker K, Hammerstein P, Schlattmann P, Telschow A, Werren JH (2008) How many species are infected with *Wolbachia*?--A statistical analysis of current data. FEMS Microbiol Lett 281: 215-220.
49. Breeuwer JA, Jacobs G (1996) *Wolbachia*: intracellular manipulators of mite reproduction. Exp Appl Acarol 20: 421-434.
50. Gotoh T, Noda H, Hong XY (2003) *Wolbachia* distribution and cytoplasmic incompatibility based on a survey of 42 spider mite species (Acari: Tetranychidae) in Japan. Heredity (Edinb) 91: 208-216.
51. Goodacre SL, Martin OY, Thomas CF, Hewitt GM (2006) *Wolbachia* and other endosymbiont infections in spiders. Mol Ecol 15: 517-527.
52. Bouchon D, Rigaud T, Juchault P (1998) Evidence for widespread *Wolbachia* infection in isopod crustaceans: molecular identification and host feminization. Proc Biol Sci 265: 1081-1090.
53. Teixeira L, Ferreira A, Ashburner M (2008) The bacterial symbiont *Wolbachia* induces resistance to RNA viral infections in Drosophila melanogaster. PLoS Biol 6: e2.
54. Hedges LM, Brownlie JC, O'Neill SL, Johnson KN (2008) *Wolbachia* and virus protection in insects. Science 322: 702.
55. Osborne SE, Leong YS, O'Neill SL, Johnson KN (2009) Variation in antiviral protection mediated by different *Wolbachia* strains in *Drosophila simulans*. PLoS Pathog 5: e1000656.

56. Unckless RL, Jaenike J (2012) Maintenance of a male-killing *Wolbachia* in *Drosophila innubila* by male-killing dependent and male-killing independent mechanisms. Evolution 66: 678-689.
57. Moreira LA, Iturbe-Ormaetxe I, Jeffery JA, Lu G, Pyke AT, et al. (2009) A *Wolbachia* symbiont in *Aedes aegypti* limits infection with dengue, Chikungunya, and *Plasmodium*. Cell 139: 1268-1278.
58. Glaser RL, Meola MA (2010) The native *Wolbachia* endosymbionts of *Drosophila melanogaster* and *Culex quinquefasciatus* increase host resistance to West Nile virus infection. PLoS One 5: e11977.
59. Taylor MJ, Bandi C, Hoerauf A (2005) *Wolbachia* bacterial endosymbionts of filarial nematodes. Adv Parasitol 60: 245-284.
60. Hosokawa T, Koga R, Kikuchi Y, Meng XY, Fukatsu T (2010) *Wolbachia* as a bacteriocyte-associated nutritional mutualist. Proc Natl Acad Sci U S A 107: 769-774.
61. Dedeine F, Vavre F, Fleury F, Loppin B, Hochberg ME, et al. (2001) Removing symbiotic *Wolbachia* bacteria specifically inhibits oogenesis in a parasitic wasp. Proc Natl Acad Sci U S A 98: 6247-6252.
62. Kurtti TJ, Munderloh UG, Andreadis TG, Magnarelli LA, Mather TN (1996) Tick cell culture isolation of an intracellular prokaryote from the tick *Ixodes scapularis*. J Invertebr Pathol 67: 318-321.
63. Duron O, Bouchon D, Boutin S, Bellamy L, Zhou LQ, et al. (2008) The diversity of reproductive parasites among arthropods: *Wolbachia* do not walk alone. Bmc Biology 6.
64. Weeks AR, Velten R, Stouthamer R (2003) Incidence of a new sex-ratio-distorting endosymbiotic bacterium among arthropods. Proc Biol Sci 270: 1857-1865.
65. Duron O, Hurst GDD, Hornett EA, Josling JA, Engelstadter J (2008) High incidence of the maternally inherited bacterium *Cardinium* in spiders. Molecular Ecology 17: 1427-1437.
66. Anbutsu H, Fukatsu T (2011) *Spiroplasma* as a model insect endosymbiont. Environmental Microbiology Reports 3: 144-153.
67. Regassa LB, Gasparich GE (2006) Spiroplasmas: evolutionary relationships and biodiversity. Frontiers in Bioscience 11: 2983-3002.
68. Jaenike J, Unckless R, Cockburn SN, Boelio LM, Perlman SJ (2010) Adaptation via Symbiosis: Recent Spread of a Drosophila Defensive Symbiont. Science 329: 212-215.
69. Xie J, Vilchez I, Mateos M (2010) *Spiroplasma* bacteria enhance survival of *Drosophila hydei* attacked by the parasitic wasp *Leptopilina heterotoma*. PLoS One 5: e12149.
70. Heddi A, Charles H, Khatchadourian C, Bonnot G, Nardon P (1998) Molecular characterization of the principal symbiotic bacteria of the weevil *Sitophilus oryzae*: a peculiar G + C content of an endocytobiotic DNA. J Mol Evol 47: 52-61.
71. Maltz MA, Weiss BL, O'Neill M, Wu YN, Aksoy S (2012) OmpA-Mediated Biofilm Formation Is Essential for the Commensal Bacterium *Sodalis glossinidius* To Colonize the Tsetse Fly Gut. Applied and Environmental Microbiology 78: 7760-7768.
72. Novakova E, Hypsa V (2007) A new *Sodalis* lineage from bloodsucking fly *Craterina melbae* (Diptera, Hippoboscoidea) originated independently of the tsetse flies symbiont *Sodalis glossinidius*. FEMS Microbiol Lett 269: 131-135.

73. Fukatsu T, Koga R, Smith WA, Tanaka K, Nikoh N, et al. (2007) Bacterial endosymbiont of the slender pigeon louse, *Columbicola columbae*, allied to endosymbionts of grain weevils and tsetse flies. Appl Environ Microbiol 73: 6660-6668.
74. Sameshima S HE, Kitade O, Minaka N, Matsumoto T (1999) Phylogenetic comparison of endosymbionts with their host ants based on molecular evidence. Zoolog Sci 16: 993-1000.
75. Burke GR, Normark BB, Favret C, Moran NA (2009) Evolution and diversity of facultative symbionts from the aphid subfamily Lachninae. Appl Environ Microbiol 75: 5328-5335.
76. Gruwell ME, Hardy NB, Gullan PJ, Dittmar K (2010) Evolutionary relationships among primary endosymbionts of the mealybug subfamily phenacoccinae (hemiptera: Coccoidea: Pseudococcidae). Appl Environ Microbiol 76: 7521-7525.
77. Kaiwa N, Hosokawa T, Kikuchi Y, Nikoh N, Meng XY, et al. (2010) Primary gut symbiont and secondary, *Sodalis*-allied symbiont of the Scutellerid stinkbug *Cantao ocellatus*. Appl Environ Microbiol 76: 3486-3494.
78. Kaiwa N, Hosokawa T, Kikuchi Y, Nikoh N, Meng XY, et al. (2011) Bacterial symbionts of the giant jewel stinkbug *Eucorysses grandis* (Hemiptera: Scutelleridae). Zoolog Sci 28: 169-174.
79. Grunwald S, Pilhofer M, Holl W (2010) Microbial associations in gut systems of wood- and bark-inhabiting longhorned beetles [Coleoptera: Cerambycidae]. Syst Appl Microbiol 33: 25-34.
80. Toju H, Hosokawa T, Koga R, Nikoh N, Meng XY, et al. (2010) "*Candidatus* Curculioniphilus buchneri," a novel clade of bacterial endocellular symbionts from weevils of the genus Curculio. Appl Environ Microbiol 76: 275-282.
81. Toju H, Fukatsu T (2011) Diversity and infection prevalence of endosymbionts in natural populations of the chestnut weevil: relevance of local climate and host plants. Mol Ecol 20: 853-868.
82. Conord C, Despres L, Vallier A, Balmand S, Miquel C, et al. (2008) Long-term evolutionary stability of bacterial endosymbiosis in curculionoidea: additional evidence of symbiont replacement in the dryophthoridae family. Mol Biol Evol 25: 859-868.
83. Campbell BC, Bragg, T.S., Turner, C.E. (1992) Phylogeny of symbioticbacteria of fourweevilspecies (coleoptera: curculionidae) based on analysis of 16S ribosomal DNA. Insect Biochem Mol Biol 22: 415-421.
84. Novakova E, Hypsa V, Moran NA (2009) *Arsenophonus*, an emerging clade of intracellular symbionts with a broad host distribution. BMC Microbiol 9: 143.
85. Huger AM, Skinner SW, Werren JH (1985) Bacterial infections associated with the son-killer trait in the parasitoid wasp *Nasonia* (= *Mormoniella*) *vitripennis* (Hymenoptera: Pteromalidae). J Invertebr Pathol 46: 272-280.
86. Werren JH, Skinner SW, Huger AM (1986) Male-killing bacteria in a parasitic wasp. Science 231: 990-992.
87. Gherna RL, Werren JH, Weisburg W, Cote R, Woese CR, et al. (1991) *Arsenophonus-Nasoniae* Gen-Nov, Sp-Nov, the Causative Agent of the Son-Killer Trait in the Parasitic Wasp *Nasonia-Vitripennis*. International Journal of Systematic Bacteriology 41: 563-565.
88. Duron O, Wilkes TE, Hurst GDD (2010) Interspecific transmission of a male-killing bacterium on an ecological timescale. Ecology Letters 13: 1139-1148.

89. Zreik L, Bove JM, Garnier M (1998) Phylogenetic characterization of the bacterium-like organism associated with marginal chlorosis of strawberry and proposition of a Candidatus taxon for the organism, '*Candidatus* Phlomobacter fragariae'. International Journal of Systematic Bacteriology 48: 257-261.
90. Bressan A, Semetey O, Arneodo J, Lherminier J, Boudon-Padieu E (2009) Vector Transmission of a Plant-Pathogenic Bacterium in the *Arsenophonus* Clade Sharing Ecological Traits with Facultative Insect Endosymbionts. Phytopathology 99: 1289-1296.
91. Hypsa V, Dale C (1997) In vitro culture and phylogenetic analysis of "*Candidatus* Arsenophonus triatominarum," an intracellular bacterium from the triatomine bug, Triatoma infestans. Int J Syst Bacteriol 47: 1140-1144.
92. Trowbridge RE, Dittmar K, Whiting MF (2006) Identification and phylogenetic analysis of *Arsenophonus*- and *Photorhabdus*-type bacteria from adult Hippoboscidae and Streblidae (Hippoboscoidea). Journal of Invertebrate Pathology 91: 64-68.
93. Sasaki-Fukatsu K, Koga R, Nikoh N, Yoshizawa K, Kasai S, et al. (2006) Symbiotic bacteria associated with stomach discs of human lice. Appl Environ Microbiol 72: 7349-7352.
94. Perotti MA, Allen JM, Reed DL, Braig HR (2007) Host-symbiont interactions of the primary endosymbiont of human head and body lice. FASEB J 21: 1058-1066.
95. Hansen AK, Jeong G, Paine TD, Stouthamer R (2007) Frequency of secondary symbiont infection in an invasive psyllid relates to parasitism pressure on a geographic scale in California. Applied and Environmental Microbiology 73: 7531-7535.
96. Perotti MA, Clarke HK, Turner BD, Braig HR (2006) *Rickettsia* as obligate and mycetomic bacteria. Faseb Journal 20: 2372-+.
97. Himler AG, Adachi-Hagimori T, Bergen JE, Kozuch A, Kelly SE, et al. (2011) Rapid spread of a bacterial symbiont in an invasive whitefly is driven by fitness benefits and female bias. Science 332: 254-256.
98. Werren JH, Hurst GD, Zhang W, Breeuwer JA, Stouthamer R, et al. (1994) Rickettsial relative associated with male killing in the ladybird beetle (Adalia bipunctata). J Bacteriol 176: 388-394.
99. von der Schulenburg JH, Habig M, Sloggett JJ, Webberley KM, Bertrand D, et al. (2001) Incidence of male-killing *Rickettsia* spp. (alpha-proteobacteria) in the ten-spot ladybird beetle *Adalia decempunctata* L. (Coleoptera: Coccinellidae). Appl Environ Microbiol 67: 270-277.
100. Lawson ET, Mousseau TA, Klaper R, Hunter MD, Werren JH (2001) *Rickettsia* associated with male-killing in a buprestid beetle. Heredity (Edinb) 86: 497-505.
101. Hagimori T, Abe Y, Date S, Miura K (2006) The first finding of a *Rickettsia* bacterium associated with parthenogenesis induction among insects. Current Microbiology 52: 97-101.
102. Muller HJ (1964) The Relation of Recombination to Mutational Advance. Mutat Res 106: 2-9.
103. Moran NA (1996) Accelerated evolution and Muller's rachet in endosymbiotic bacteria. Proc Natl Acad Sci U S A 93: 2873-2878.
104. Mira A, Moran NA (2002) Estimating population size and transmission bottlenecks in maternally transmitted endosymbiotic bacteria. Microb Ecol 44: 137-143.
105. Mira A, Ochman H, Moran NA (2001) Deletional bias and the evolution of bacterial genomes. Trends Genet 17: 589-596.

106. Moran NA, Plague GR (2004) Genomic changes following host restriction in bacteria. Curr Opin Genet Dev 14: 627-633.
107. Hildebrand F, Meyer A, Eyre-Walker A (2010) Evidence of selection upon genomic GC-content in bacteria. PLoS Genet 6: e1001107.
108. Hershberg R, Petrov DA (2010) Evidence that mutation is universally biased towards AT in bacteria. PLoS Genet 6: e1001115.
109. McCutcheon JP, Moran NA (2010) Functional convergence in reduced genomes of bacterial symbionts spanning 200 My of evolution. Genome Biol Evol 2: 708-718.
110. McCutcheon JP, Moran NA (2012) Extreme genome reduction in symbiotic bacteria. Nat Rev Microbiol 10: 13-26.
111. McCutcheon JP, McDonald BR, Moran NA (2009) Origin of an alternative genetic code in the extremely small and GC-rich genome of a bacterial symbiont. PLoS Genet 5: e1000565.
112. Sauer C, Stackebrandt E, Gadau J, Holldobler B, Gross R (2000) Systematic relationships and cospeciation of bacterial endosymbionts and their carpenter ant host species: proposal of the new taxon *Candidatus* Blochmannia gen. nov. International Journal of Systematic and Evolutionary Microbiology 50: 1877-1886.
113. Charles H, Heddi A, Rahbe Y (2001) A putative insect intracellular endosymbiont stem clade, within the Enterobacteriaceae, infered from phylogenetic analysis based on a heterogeneous model of DNA evolution. C R Acad Sci III 324: 489-494.
114. Gil R, Silva FJ, Zientz E, Delmotte F, Gonzalez-Candelas F, et al. (2003) The genome sequence of *Blochmannia floridanus*: comparative analysis of reduced genomes. Proc Natl Acad Sci U S A 100: 9388-9393.
115. Lerat E, Daubin V, Moran NA (2003) From gene trees to organismal phylogeny in prokaryotes: The case of the gamma-proteobacteria. Plos Biology 1: 101-109.
116. Canback B, Tamas I, Andersson SG (2004) A phylogenomic study of endosymbiotic bacteria. Mol Biol Evol 21: 1110-1122.
117. Belda E, Moya A, Silva FJ (2005) Genome rearrangement distances and gene order phylogeny in gamma-Proteobacteria. Mol Biol Evol 22: 1456-1467.
118. Herbeck JT, Degnan PH, Wernegreen JJ (2005) Nonhomogeneous model of sequence evolution indicates independent origins of primary endosymbionts within the enterobacteriales (gamma-Proteobacteria). Mol Biol Evol 22: 520-532.
119. Ciccarelli FD (2006) Toward automatic reconstruction of a highly resolved tree of life (vol 311, pg 1283, 2006). Science 312: 697-697.
120. Delmotte F, Rispe C, Schaber J, Silva FJ, Moya A (2006) Tempo and mode of early gene loss in endosymbiotic bacteria from insects. BMC Evol Biol 6: 56.
121. Luo YQ, Fu C, Zhang DY, Lin K (2006) Overlapping genes as rare genomic markers: the phylogeny of gamma-Proteobacteria as a case study. Trends in Genetics 22: 593-596.
122. Wu D, Daugherty SC, Van Aken SE, Pai GH, Watkins KL, et al. (2006) Metabolic complementarity and genomics of the dual bacterial symbiosis of sharpshooters. PLoS Biol 4: e188.
123. Comas I, Moya A, Gonzalez-Candelas F (2007) From phylogenetics to phylogenomics: the evolutionary relationships of insect endosymbiotic γ-Proteobacteria as a test case. Syst Biol 56: 1-16.

124. Comas I, Moya A, Gonzalez-Candelas F (2007) Phylogenetic signal and functional categories in Proteobacteria genomes. BMC Evol Biol 7 Suppl 1: S7.
125. Ding G, Yu Z, Zhao J, Wang Z, Li Y, et al. (2008) Tree of life based on genome context networks. PLoS One 3: e3357.
126. Jiang LW, Lin KL, Lu CL (2008) OGtree: a tool for creating genome trees of prokaryotes based on overlapping genes. Nucleic Acids Res 36: W475-480.
127. Wu M, Eisen JA (2008) A simple, fast, and accurate method of phylogenomic inference. Genome Biol 9: R151.
128. Gao B, Gupta RS (2012) Phylogenetic framework and molecular signatures for the main clades of the phylum Actinobacteria. Microbiol Mol Biol Rev 76: 66-112.
129. Lopez-Sanchez MJ, Neef A, Pereto J, Patino-Navarrete R, Pignatelli M, et al. (2009) Evolutionary convergence and nitrogen metabolism in *Blattabacterium strain* Bge, primary endosymbiont of the cockroach *Blattella germanica*. PLoS Genet 5: e1000721.
130. Cheng CH, Yang CH, Chiu HT, Lu CL (2010) Reconstructing genome trees of prokaryotes using overlapping genes. BMC Bioinformatics 11: 102.
131. Williams KP, Gillespie JJ, Sobral BWS, Nordberg EK, Snyder EE, et al. (2010) Phylogeny of Gammaproteobacteria. Journal of Bacteriology 192: 2305-2314.
132. Felsenstein J (1988) Phylogenies from Molecular Sequences - Inference and Reliability. Annual Review of Genetics 22: 521-565.
133. Williams KP, Gillespie JJ, Sobral BW, Nordberg EK, Snyder EE, et al. (2010) Phylogeny of gammaproteobacteria. J Bacteriol 192: 2305-2314.
134. Brinkmann H, Philippe H (1999) Archaea sister group of Bacteria? Indications from tree reconstruction artifacts in ancient phylogenies. Mol Biol Evol 16: 817-825.
135. Lartillot N, Philippe H (2004) A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. Molecular Biology and Evolution 21: 1095-1109.
136. Wiedenbeck J, Cohan FM (2011) Origins of bacterial diversity through horizontal genetic transfer and adaptation to new ecological niches. Fems Microbiology Reviews 35: 957-976.
137. Guttman B, Raya R, Kutter E (2005) Basic phage biology. In: Kutter E, Sulakvelidze A, editors. Bacteriophages: Biology and Application
CRC Press, Boca Raton, FL.
138. Lederberg J, Lederberg EM, Zinder ND, Lively ER (1951) Recombination Analysis of Bacterial Heredity. Cold Spring Harbor Symposia on Quantitative Biology 16: 413-443.
139. van der Wilk F, Dullemans AM, Verbeek M, van den Heuvel JF (1999) Isolation and characterization of APSE-1, a bacteriophage infecting the secondary endosymbiont of *Acyrthosiphon pisum*. Virology 262: 104-113.
140. Oliver KM, Russell JA, Moran NA, Hunter MS (2003) Facultative bacterial symbionts in aphids confer resistance to parasitic wasps. Proc Natl Acad Sci U S A 100: 1803-1807.
141. Degnan PH, Moran NA (2008) Evolutionary genetics of a defensive facultative symbiont of insects: exchange of toxin-encoding bacteriophage. Mol Ecol 17: 916-929.
142. (1904) The Cotton Boll Weevil. Science 20: 475-476.
143. Degnan PH, Moran NA (2008) Diverse phage-encoded toxins in a protective insect endosymbiont. Appl Environ Microbiol 74: 6782-6791.

144. Oliver KM, Degnan PH, Hunter MS, Moran NA (2009) Bacteriophages encode factors required for protection in a symbiotic mutualism. Science 325: 992-994.
145. Martinez AJ, Weldon SR, Oliver KM (2013) Effects of parasitism on aphid nutritional and protective symbioses. Mol Ecol.
146. Preston GM (2007) Metropolitan microbes: Type III secretion in multihost symbionts. Cell Host & Microbe 2: 291-294.
147. Coburn B, Sekirov I, Finlay BB (2007) Type III secretion systems and disease. Clinical Microbiology Reviews 20: 535-+.
148. Grant SR, Fisher EJ, Chang JH, Mole BM, Dangl JL (2006) Subterfuge and manipulation: Type III effector proteins of phytopathogenic bacteria. Annual Review of Microbiology 60: 425-449.
149. Dale C, Jones T, Pontes M (2005) Degenerative evolution and functional diversification of type-III secretion systems in the insect endosymbiont *Sodalis glossinidius*. Mol Biol Evol 22: 758-766.
150. Silver AC, Kikuchi Y, Fadl AA, Sha J, Chopra AK, et al. (2007) 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: 9481-9486.
151. Clayton AL, Oakeson KF, Gutin M, Pontes A, Dunn DM, et al. (2012) A Novel Human-Infection-Derived Bacterium Provides Insights into the Evolutionary Origins of Mutualistic Insect-Bacterial Symbioses. Plos Genetics 8.
152. Caspi-Fluger A, Inbar M, Mozes-Daube N, Katzir N, Portnoy V, et al. (2012) Horizontal transmission of the insect symbiont *Rickettsia* is plant-mediated. Proc Biol Sci 279: 1791-1796.
153. Darling AC, Mau B, Blattner FR, Perna NT (2004) Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res 14: 1394-1403.
154. Yu JS, Noll KM (1997) Plasmid pRQ7 from the hyperthermophilic bacterium Thermotoga species strain RQ7 replicates by the rolling-circle mechanism. Journal of Bacteriology 179: 7161-7164.
155. Medema MH, Trefzer A, Kovalchuk A, van den Berg M, Muller U, et al. (2010) The Sequence of a 1.8-Mb Bacterial Linear Plasmid Reveals a Rich Evolutionary Reservoir of Secondary Metabolic Pathways. Genome Biology and Evolution 2: 212-224.
156. Carattoli A (2013) Plasmids and the spread of resistance. International Journal of Medical Microbiology 303: 298-304.
157. Hardy KG (1975) Colicinogeny and Related Phenomena. Bacteriological Reviews 39: 464-515.
158. Gordon S, Rech J, Lane D, Wright A (2004) Kinetics of plasmid segregation in *Escherichia coli*. Molecular Microbiology 51: 461-469.
159. Johnson TJ, Nolan LK (2009) Pathogenomics of the Virulence Plasmids of *Escherichia coli*. Microbiology and Molecular Biology Reviews 73: 750-774.
160. Afendra AS, Parapouli M, Drainas C (2011) Catabolic Plasmids and Mobile Genetic Elements Involved in the Degradation of Non-metal Xenobiotic Compounds. Microbial Bioremediation of Non-Metals: Current Research: 197-216.
161. Endo G, Ji GY, Silver S (1996) Heavy metal resistance plasmids and use in bioremediation. Environmental Biotechnology: 47-62.
162. Darby AC, Lagnel J, Matthew CZ, Bourtzis K, Maudlin I, et al. (2005) Extrachromosomal DNA of the symbiont *Sodalis glossinidius*. J Bacteriol 187: 5003-5007.

163. Fronzes R, Remaut H, Waksman G (2008) Architectures and biogenesis of non-flagellar protein appendages in Gram-negative bacteria. Embo Journal 27: 2271-2280.
164. Baumler AJ, Gilde AJ, Tsolis RM, vanderVelden AWM, Ahmer BMM, et al. (1997) Contribution of horizontal gene transfer and deletion events to development of distinctive patterns of fimbrial operons during evolution of Salmonella serotypes. Journal of Bacteriology 179: 317-322.
165. Barnhart MM, Sauer FG, Pinkner JS, Hultgren SJ (2003) Chaperone-subunit-usher interactions required for donor strand exchange during bacterial pilus assembly. Journal of Bacteriology 185: 2723-2730.
166. Pratt LA, Kolter R (1998) Genetic analysis of *Escherichia coli* biofilm formation: roles of flagella, motility, chemotaxis and type I pili. Molecular Microbiology 30: 285-293.
167. Wright KJ, Seed PC, Hultgren SJ (2007) Development of intracellular bacterial communities of uropathogenic Escherichia coli depends on type 1 pili. Cellular Microbiology 9: 2230-2241.
168. Koczan JM, Lenneman BR, McGrath MJ, Sundin GW (2011) Cell Surface Attachment Structures Contribute to Biofilm Formation and Xylem Colonization by *Erwinia amylovora*. Applied and Environmental Microbiology 77: 7031-7039.
169. Bergsten J (2005) A review of long-branch attraction. Cladistics 21: 163-193.
170. Roure B, Philippe H (2011) Site-specific time heterogeneity of the substitution process and its impact on phylogenetic inference. Bmc Evolutionary Biology 11.
171. Philippe H, Roure B (2011) Difficult phylogenetic questions: more data, maybe; better methods, certainly. Bmc Biology 9.
172. Moran NA, Munson MA, Baumann P, Ishikawa H (1993) A Molecular Clock in Endosymbiotic Bacteria Is Calibrated Using the Insect Hosts. Proceedings of the Royal Society B-Biological Sciences 253: 167-171.
173. Hosokawa T, Kikuchi Y, Nikoh N, Shimada M, Fukatsu T (2006) Strict host-symbiont cospeciation and reductive genome evolution in insect gut bacteria. PLoS Biol 4: e337.
174. Kikuchi Y, Hosokawa T, Nikoh N, Meng XY, Kamagata Y, et al. (2009) Host-symbiont co-speciation and reductive genome evolution in gut symbiotic bacteria of acanthosomatid stinkbugs. BMC Biol 7: 2.
175. Hosokawa T, Kikuchi Y, Nikoh N, Meng XY, Hironaka M, et al. (2010) Phylogenetic position and peculiar genetic traits of a midgut bacterial symbiont of the stinkbug *Parastrachia japonensis*. Appl Environ Microbiol 76: 4130-4135.
176. Prado SS, Almeida RP (2009) Phylogenetic placement of pentatomid stink bug gut symbionts. Curr Microbiol 58: 64-69.
177. Harada H, Oyaizu H, Ishikawa H (1996) A consideration about the origin of aphid intracellular symbiont in connection with gut bacterial flora. Journal of General and Applied Microbiology 42: 17-26.
178. Allen JM, Reed DL, Perotti MA, Braig HR (2007) Evolutionary relationships of "*Candidatus* Riesia spp.," endosymbiotic Enterobacteriaceae living within hematophagous primate lice. Applied and Environmental Microbiology 73: 1659-1664.
179. Koppenhofer HS, Gaugler R (2009) Defensive mutualism in microbial symbiosis.: CRC Press.

180. Berger SA (1985) *Proteus* Bacteremia in a General-Hospital 1972-1982. Journal of Hospital Infection 6: 293-298.
181. Giammanco GM, Grimont PAD, Grimont F, Lefevre M, Giammanco G, et al. (2011) Phylogenetic analysis of the genera *Proteus*, *Morganella* and *Providencia* by comparison of rpoB gene sequences of type and clinical strains suggests the reclassification of *Proteus myxofaciens* in a new genus, *Cosenzaea* gen. nov., as *Cosenzaea myxofaciens* comb. nov. International Journal of Systematic and Evolutionary Microbiology 61: 1638-1644.
182. Cosenza BJ, Podgwait.Jd (1966) A New Species of *Proteus* Isolated from Larvae of Gypsy Moth *Porthetria Dispar* (L). Antonie Van Leeuwenhoek Journal of Microbiology and Serology 32: 187-&.
183. Bacot AW, Martin CJ (1914) LXVII. Observations on the mechanism of the transmission of plague by fleas. J Hyg (Lond) 13: 423-439.
184. Bacot AW (1915) LXXXI. Further notes on the mechanism of the transmission of plague by fleas. J Hyg (Lond) 14: 774-776 773.
185. Hurst MRH, Becher SA, Young SD, Nelson TL, Glare TR (2011) *Yersinia entomophaga* sp. nov., isolated from the New Zealand grass grub *Costelytra zealandica*. International Journal of Systematic and Evolutionary Microbiology 61: 844-849.
186. Hurst MRH, Jones SA, Binglin T, Harper LA, Jackson TA, et al. (2011) The Main Virulence Determinant of Yersinia entomophaga MH96 Is a Broad-Host-Range Toxin Complex Active against Insects. Journal of Bacteriology 193: 1966-1980.
187. Wernegreen JJ, Kauppinen SN, Brady SG, Ward PS (2009) One nutritional symbiosis begat another: Phylogenetic evidence that the ant tribe Camponotini acquired *Blochmannia* by tending sap-feeding insects. Bmc Evolutionary Biology 9.
188. Glasner JD, Marquez-Villavicencio M, Kim HS, Jahn CE, Ma B, et al. (2008) Niche-Specificity and the Variable Fraction of the *Pectobacterium* Pan-Genome. Molecular Plant-Microbe Interactions 21: 1549-1560.
189. Costechareyre D, Balmand S, Condemine G, Rahbe Y (2012) *Dickeya dadantii*, a Plant Pathogenic Bacterium Producing Cyt-Like Entomotoxins, Causes Septicemia in the Pea Aphid Acyrthosiphon pisum. Plos One 7.
190. Grenier AM, Duport G, Pages S, Condemine G, Rahbe Y (2006) The phytopathogen *Dickeya dadantii* (*Erwinia chrysanthemi* 3937) is a pathogen of the pea aphid. Applied and Environmental Microbiology 72: 1956-1965.
191. Verbarg S, Fruhling A, Cousin S, Brambilla E, Gronow S, et al. (2008) *Biostraticola tofi* gen. nov., spec. nov., a novel member of the family Enterobacteriaceae. Curr Microbiol 56: 603-608.
192. McCutcheon JP, McDonald BR, Moran NA (2009) Convergent evolution of metabolic roles in bacterial co-symbionts of insects. Proc Natl Acad Sci U S A 106: 15394-15399.
193. Basset A, Tzou P, Lemaitre B, Boccard F (2003) A single gene that promotes interaction of a phytopathogenic bacterium with its insect vector, *Drosophila melanogaster*. Embo Reports 4: 205-209.
194. Vallet-Gely I, Lemaitre B, Boccard F (2008) Bacterial strategies to overcome insect defences. Nature Reviews Microbiology 6: 302-313.
195. Abd H, Wretlind B, Saeed A, Idsund E, Hultenby K, et al. (2008) *Pseudomonas aeruginosa* utilises its type III secretion system to kill the free-living amoeba *Acanthamoeba castellanii*. Journal of Eukaryotic Microbiology 55: 235-243.

196. Brugirard-Ricaud K, Duchaud E, Givaudan A, Girard PA, Kunst F, et al. (2005) Site-specific antiphagocytic function of the *Photorhabdus luminescens* type III secretion system during insect colonization. Cell Microbiol 7: 363-371.
197. Chandra H, Khandelwal P, Khattri A, Banerjee N (2008) Type 1 fimbriae of insecticidal bacterium *Xenorhabdus nematophila* is necessary for growth and colonization of its symbiotic host nematode *Steinernema carpocapsiae*. Environmental Microbiology 10: 1285-1295.
198. Alberdi MP, Walker AR, Urquhart KA (2000) Field evidence that roe deer (*Capreolus capreolus*) are a natural host for *Ehrlichia phagocytophila*. Epidemiology and Infection 124: 315-323.
199. Liz JS, Sumner JW, Pfister K, Brossard M (2002) PCR detection and serological evidence of granulocytic ehrlichial infection in roe deer (*Capreolus capreolus*) and chamois (*Rupicapra rupicapra*). Journal of Clinical Microbiology 40: 892-897.
200. Petrovec M, Sixl W, Schweiger R, Mikulasek S, Elke L, et al. (2003) Infections of wild animals with *Anaplasma phagocytophila* in Austria and the Czech Republic. Rickettsiology: Present and Future Directions 990: 103-106.
201. Hulinska D, Langrova K, Pejcoch M, Pavlasek I (2004) Detection of *Anaplasma phagocytophilum* in animals by real-time polymerase chain reaction. Apmis 112: 239-247.
202. Polin H, Hufnagl P, Haunschmid R, Gruber F, Ladurner G (2004) Molecular evidence of *Anaplasma phagocytophilum* in *Ixodes ricinus ticks* and wild animals in Austria. Journal of Clinical Microbiology 42: 2285-2286.
203. Smetanova K, Schwarzova K, Kocianova E (2006) Detection of *Anaplasma phagocytophilum*, *Coxiella burnetii*, *Rickettsia spp.*, and *Borrelia burgdorferi* s. 1. in ticks, and wild-living animals in western and middle Slovakia. Century of Rickettsiology: Emerging, Reemerging Rickettsioses, Molecular Diagnostics, and Emerging Veterinary Rickettsioses 1078: 312-315.
204. de la Fuente J, Ruiz-Fons F, Naranjo V, Torina A, Rodriguez O, et al. (2008) Evidence of *Anaplasma* infections in European roe deer (*Capreolus capreolus*) from southern Spain. Research in Veterinary Science 84: 382-386.
205. Stefanidesova K, Kocianova E, Boldis V, Kostanova Z, Kanka P, et al. (2008) Evidence of *Anaplasma phagocytophilum* and *Rickettsia helvetica* infection in free-ranging ungulates in central Slovakia. European Journal of Wildlife Research 54: 519-524.
206. Baldridge GD, Scoles GA, Burkhardt NY, Schloeder B, Kurtti TJ, et al. (2009) Transovarial Transmission of *Francisella*-Like Endosymbionts and *Anaplasma phagocytophilum* Variants in *Dermacentor albipictus* (Acari: Ixodidae). Journal of Medical Entomology 46: 625-632.
207. Strle F (2004) Human granulocytic ehrlichiosis in Europe. International Journal of Medical Microbiology 293: 27-35.
208. Woldehiwet Z (2010) The natural history of *Anaplasma phagocytophilum*. Veterinary Parasitology 167: 108-122.
209. Silaghi C, Hamel D, Thiel C, Pfister K, Passos LMF, et al. (2011) Genetic Variants of *Anaplasma phagocytophilum* in Wild Caprine and Cervid Ungulates from the Alps in Tyrol, Austria. Vector-Borne and Zoonotic Diseases 11: 355-362.
210. Alberti A, Zobba R, Chessa B, Addis MF, Sparagano O, et al. (2005) Equine and canine *Anaplasma phagocytophilum* strains isolated on the island of Sardinia (Italy) are phylogenetically related to pathogenic strains from the United States. Applied and Environmental Microbiology 71: 6418-6422.

211. Petrovec M, Bidovec A, Sumner JW, Nicholson WL, Childs JE, et al. (2002) Infection with *Anaplasma phagocytophila* in cervids from Slovenia: Evidence of two genotypic lineages. Wiener Klinische Wochenschrift 114: 641-647.
212. von Loewenich FD, Baumgarten BU, Schroppel K, Geissdorfer W, Rollinghoff M, et al. (2003) High diversity of *ankA* sequences of *Anaplasma phagocytophilum* among *Ixodes ricinus* ticks in Germany. Journal of Clinical Microbiology 41: 5033-5040.
213. Rymaszewska A (2008) Divergence within the marker region of the *groESL* operon in *Anaplasma phagocytophilum*. European Journal of Clinical Microbiology & Infectious Diseases 27: 1025-1036.

4 Publications

4.1 Chapter 1

Husník F., Chrudimský T., Hypša V., 2011. Multiple origins of endosymbiosis within the Enterobcateriaceae (γ-Proteobacteria): convergence of complex phylogenetic approaches. BMC Biology Dec 28;9:87. doi: 10.1186/1741-7007-9-87.

Abstract

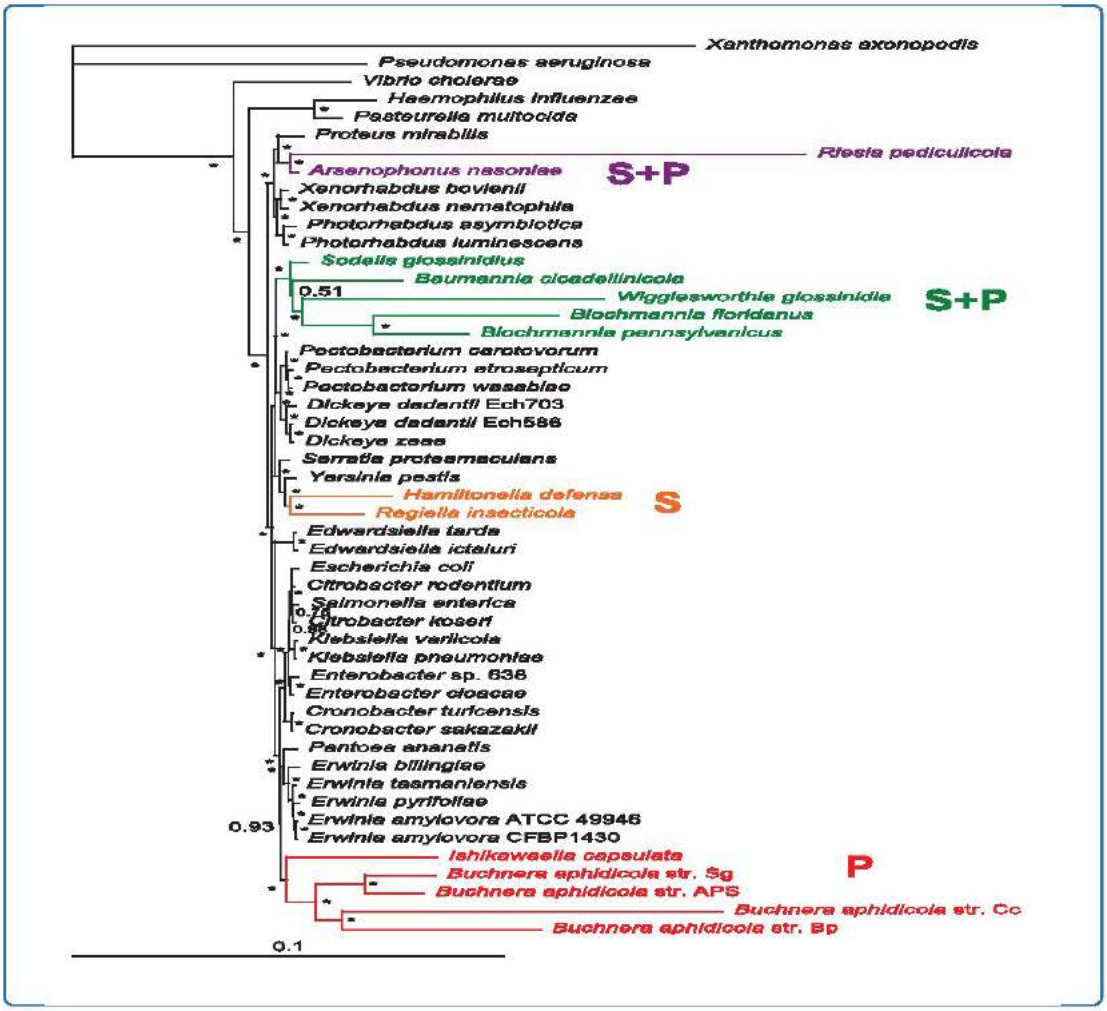
The bacterial family Enterobacteriaceae gave rise to a variety of symbiotic forms, from the loosely associated commensals, often designated as secondary (S) symbionts, to obligate mutualists, called primary (P) symbionts. Determination of the evolutionary processes behind this phenomenon has long been hampered by the unreliability of phylogenetic reconstructions within this group of bacteria. The main reasons have been the absence of sufficient data, the highly derived nature of the symbiont genomes and lack of appropriate phylogenetic methods. Due to the extremely aberrant nature of their DNA, the symbiotic lineages within Enterobacteriaceae form long branches and tend to cluster as a monophyletic group. This state of phylogenetic uncertainty is now improving with an increasing number of complete bacterial genomes and development of new methods. In this study, we address the monophyly versus polyphyly of enterobacterial symbionts by exploring a multigene matrix within a complex phylogenetic framework.

We assembled the richest taxon sampling of Enterobacteriaceae to date (50 taxa, 69 orthologous genes with no missing data) and analyzed both nucleic and amino acid data sets using several probabilistic methods. We particularly focused on the long-branch attraction-reducing methods, such as a nucleotide and amino acid data recoding and exclusion (including our new approach and slow-fast analysis), taxa exclusion and usage of complex evolutionary models, such as nonhomogeneous model and models accounting for site-specific features of protein evolution (CAT and CAT+GTR). Our data strongly suggest independent origins of four symbiotic clusters; the first is formed

by *Hamiltonella* and *Regiella* (S-symbionts) placed as a sister clade to *Yersinia*, the second comprises *Arsenophonus* and *Riesia* (S- and P-symbionts) as a sister clade to *Proteus*, the third *Sodalis*, *Baumannia*, *Blochmannia* and *Wigglesworthia* (S- and P-symbionts) as a sister or paraphyletic clade to the *Pectobacterium* and *Dickeya* clade and, finally, *Buchnera* species and *Ishikawaella* (P-symbionts) clustering with the *Erwinia* and *Pantoea* clade.

The results of this study confirm the efficiency of several artifact-reducing methods and strongly point towards the polyphyly of P-symbionts within Enterobacteriaceae. Interestingly, the model species of symbiotic bacteria research, *Buchnera* and *Wigglesworthia*, originated from closely related, but different, ancestors. The possible origins of intracellular symbiotic bacteria from gut-associated or pathogenic bacteria are suggested, as well as the role of facultative secondary symbionts as a source of bacteria that can gradually become obligate maternally transferred symbionts.

BMC Biology
incorporating Journal of Biology



Multiple origins of endosymbiosis within the Enterobacteriaceae (γ-Proteobacteria): convergence of complex phylogenetic approaches

Husník *et al.*

BioMed Central

Husník *et al. BMC Biology* 2011, 9:87
http://www.biomedcentral.com/1741-7007/9/87 (28 December 2011)

Husník et al. BMC Biology 2011, 9:87
http://www.biomedcentral.com/1741-7007/9/87

BMC Biology
Incorporating Journal of Biology

RESEARCH ARTICLE Open Access

Multiple origins of endosymbiosis within the Enterobacteriaceae (γ-Proteobacteria): convergence of complex phylogenetic approaches

Filip Husník[1*], Tomáš Chrudimský[1] and Václav Hypša[1,2*]

Abstract

Background: The bacterial family Enterobacteriaceae gave rise to a variety of symbiotic forms, from the loosely associated commensals, often designated as secondary (S) symbionts, to obligate mutualists, called primary (P) symbionts. Determination of the evolutionary processes behind this phenomenon has long been hampered by the unreliability of phylogenetic reconstructions within this group of bacteria. The main reasons have been the absence of sufficient data, the highly derived nature of the symbiont genomes and lack of appropriate phylogenetic methods. Due to the extremely aberrant nature of their DNA, the symbiotic lineages within Enterobacteriaceae form long branches and tend to cluster as a monophyletic group. This state of phylogenetic uncertainty is now improving with an increasing number of complete bacterial genomes and development of new methods. In this study, we address the monophyly versus polyphyly of enterobacterial symbionts by exploring a multigene matrix within a complex phylogenetic framework.

Results: We assembled the richest taxon sampling of Enterobacteriaceae to date (50 taxa, 69 orthologous genes with no missing data) and analyzed both nucleic and amino acid data sets using several probabilistic methods. We particularly focused on the long-branch attraction-reducing methods, such as a nucleotide and amino acid data recoding and exclusion (including our new approach and slow-fast analysis), taxa exclusion and usage of complex evolutionary models, such as nonhomogeneous model and models accounting for site-specific features of protein evolution (CAT and CAT+GTR). Our data strongly suggest independent origins of four symbiotic clusters; the first is formed by *Hamiltonella* and *Regiella* (S-symbionts) placed as a sister clade to *Yersinia*, the second comprises *Arsenophonus* and *Riesia* (S- and P-symbionts) as a sister clade to *Proteus*, the third *Sodalis*, *Baumannia*, *Blochmannia* and *Wigglesworthia* (S- and P-symbionts) as a sister or paraphyletic clade to the *Pectobacterium* and *Dickeya* clade and, finally, *Buchnera* species and *Ishikawaella* (P-symbionts) clustering with the *Erwinia* and *Pantoea* clade.

Conclusions: The results of this study confirm the efficiency of several artifact-reducing methods and strongly point towards the polyphyly of P-symbionts within Enterobacteriaceae. Interestingly, the model species of symbiotic bacteria research, *Buchnera* and *Wigglesworthia*, originated from closely related, but different, ancestors. The possible origins of intracellular symbiotic bacteria from gut-associated or pathogenic bacteria are suggested, as well as the role of facultative secondary symbionts as a source of bacteria that can gradually become obligate maternally transferred symbionts.

* Correspondence: filip.husnik@gmail.com; vacatko@paru.cas.cz
[1]Faculty of Science, University of South Bohemia, Branišovská 31, České Budějovice 37005, Czech Republic
Full list of author information is available at the end of the article

Background

One of the most fundamental evolutionary questions concerning insect-bacteria symbiosis is the origin and phylogenetic relationships of various symbiotic lineages. This knowledge is necessary for understanding the dynamics and mechanisms of symbiosis establishment and maintenance within the host. For instance, close relationships between symbionts and pathogenic bacteria suggests a transition from pathogenicity to symbiosis; polyphyly of the symbionts within a single host group is evidence of their multiple independent origins and close relationships among symbionts of different biology indicate high ecological flexibility within a given symbiotic group [1-6]. These implications are particularly important within Enterobacteriaceae, the group containing a broad spectrum of symbiotic lineages and forms described from various groups of insects. Their biology varies from loosely associated facultative symbionts (often called Secondary (S) symbionts) to obligatory mutualists of a highly derived nature, called Primary (P) symbionts [7-9]. However, the concept of the P- and S-symbionts and the associated terminology are a major oversimplification and they become inadequate for the description of the ever increasing complexity of the symbiotic system within Enterobacteriaceae. This complexity is manifested by such phenomena as the presence of multiple symbionts in a single host [10], occurrence of intermediate symbiotic forms and the replacement of symbionts within a host [11-14] or close phylogenetic relationships between typical S- and P-symbionts revealing their high ecological versatility [15]. A good example of such a complex system is provided by the occurrence of multiple obligate symbionts within Auchenorrhyncha [10], universally harboring *Sulcia muelleri* (Bacteroidetes) [16] with either *Hodgkinia cicadicola* (α-Proteobacteria) in cicadas, *Zinderia insecticola* (β-Proteobacteria) in spittlebugs or *Baumannia cicadellinicola* (γ-Proteobacteria) in sharpshooters. All of these latter symbionts are obligate and have been cospeciating with their hosts for millions of years [17-21]. A close phylogenetic relationship between typical S- and P-symbionts has been so far demonstrated in two well defined and often studied groups, the enterobacterial genera *Arsenophonus* and *Sodalis* [5,22,23]. The general capability of S-symbionts to supplement the metabolic functions of P-symbionts or even replace them was demonstrated experimentally by replacement of *Buchnera* with *Serratia* in aphids [24].

It is obvious that all these fascinating processes can only be studied on a reliable phylogenetic background [9,25-28]. Unfortunately, under current conditions, the phylogeny within Enterobacteriaceae and the placement of various symbiotic lineages are very unstable. Particularly, the P-symbionts present an extremely difficult challenge to phylogenetic computation due to their strongly modified genomes [9]. There are several root problems that are responsible for this dissatisfactory state. Traditionally, 16S rDNA was frequently used as an exclusive molecular marker for the description of a new symbiont. Many lineages are thus represented only by this gene, which has been shown within Enterobacteriaceae to be inadequate for inferring a reliable phylogeny [29]. In addition, it is notoriously known that the phylogenetic information of symbiotic bacteria is often seriously distorted due to the conditions associated with the symbiotic lifestyle. The effect of strong bottlenecks accompanied by reduced purifying selection and the overall degeneration of symbiotic genomes have been thoroughly discussed in many studies [30-33]. As a result of these degenerative processes, symbiotic lineages may experience parallel changes of their DNAs and these convergences produce the main source of phylogenetic artifacts. Among the most important features are biased nucleotide composition favoring adenine-thymine bases and rapid sequence evolution. While the compositional bias leads to the introduction of homoplasies at both nucleotide and amino acid levels, the accelerated evolution is a well known source of the long-branch attraction phenomenon [34,35]. Due to these circumstances, symbionts almost always appear as long branches in phylogenetic trees and tend to cluster together [36].

Various methodological approaches have been tested to overcome these difficulties (Additional file 1). They are based mainly on the concatenation of a large number of genes through the whole genome [37-39], the supertree and the consensus approach [37], exclusion of amino acids (FYMINK: phenylalanine, tyrosine, methionine, isoleucine, asparagine and lysine) most affected by nucleotide bias [37], modifications of sequence evolution models [11,12,36,40] and use of the genome structure as a source of phylogenetic data [41]. Phylogenomic studies based on large concatenated sets frequently imply monophyly of the typical P-symbionts (Additional file 1). However, due to the limited number of available genomes, these studies are usually based on inadequate taxon sampling. For example, secondary symbionts and plant pathogens that were shown to break the P-symbiont monophyly in the analysis using a nonhomogeneous model [40] could not be included into these phylogenomic studies. It is important to note that P-symbionts are probably only distantly related to the *Escherichia/Salmonella/Yersinia* clade. Therefore, the monophyly of P-symbionts derived from such a phylogenomic dataset is logically inevitable, but does not carry any evolutionary information.

The non-monophyletic nature of P-symbionts has been recently suggested in several studies. Perhaps the most inspiring is a study based on a nonhomogeneous model that separates P-symbionts into two independent lineages [40]. As an alternative, a paraphyletic arrangement of these symbionts in respect to several free-living taxa has been revealed from gene-order analysis based on break-point and inversion distances [41]. Most recently, Williams *et al.* [42] performed a 'telescoping' multiprotein phylogenomic analysis of 104 γ-Proteobacterial genomes. The phylogeny of Enterobacteriaceae endosymbionts was difficult to resolve, although it appeared that there were independent origins of at least the *Sodalis* and *Buchnera* lineages.

Thus, there is now a spectrum of hypotheses on the phylogeny of insect symbionts, ranging from complete polyphyly with multiple independent origins to complete monophyly with one common origin. In this study, we take advantage of current progress in computational methods to investigate phylogenetic relationships among the symbiotic lineages. One of the promising recent methodological advances is the introduction of a site-heterogeneous non-parametric mixture CAT model that allows for site-specific features of protein evolution [43]. This model was shown to solve the long-branch attraction (LBA) artifacts and outperform the previous models [44-47]. Similarly, the slow-fast method based on removal of the fastest evolving sites was shown to reduce phylogenetic artifacts [48-54], as well as purine/pyrimidine (RY) data recoding [55-58] or amino acid data recoding [59,60]. We used these methods as the core of a complex approach and tried to investigate series of methods, models and parameters to detect common trends in changes of the topologies. To do this, we applied two parallel approaches, one based on the application of recently developed algorithms and the other on the removal or recoding of the positions most affected by rapid sequence evolution and/or compositional (AT) bias. In addition, we paid particular attention to the sampling and used as much of a complete set of both symbiotic and free-living lineages as possible. This approach is particularly important to avoid interpretation uncertainties due to the absence of phylogenetically important lineages.

Results

The complete methodological design of this study and the resulting topologies are depicted in Figure 1. All matrices, alignments and phylogenetic trees are available in the TreeBASE database http://purl.org/phylo/treebase/phylows/study/TB2:S11451, as supplementary material, or on the webpage http://users.prf.jcu.cz/husnif00.

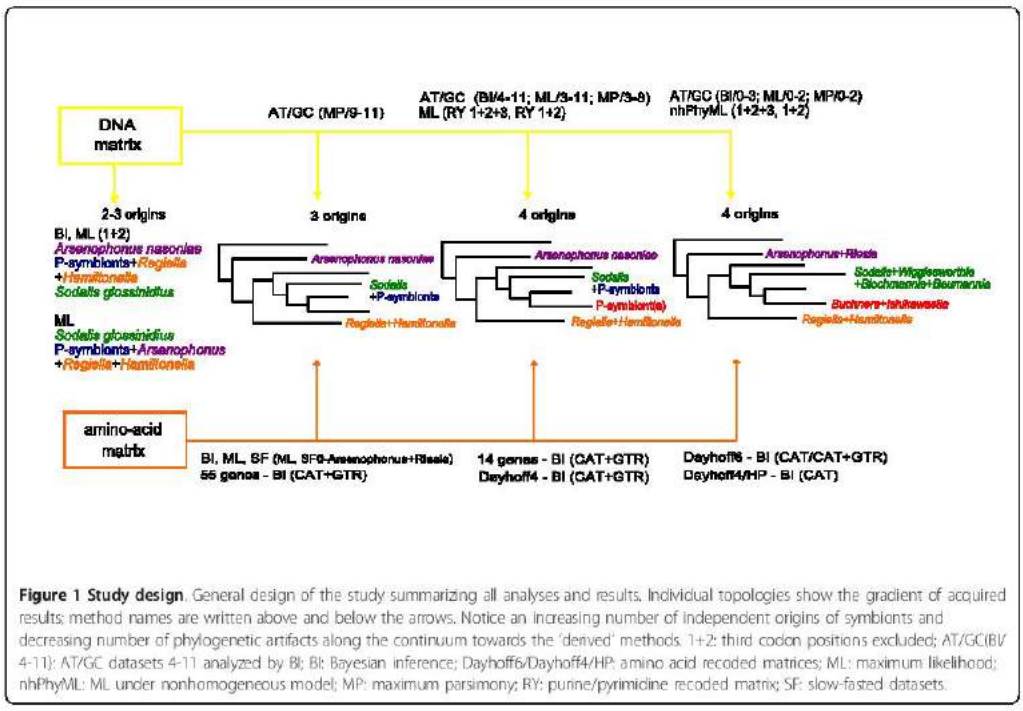


Figure 1 Study design. General design of the study summarizing all analyses and results. Individual topologies show the gradient of acquired results; method names are written above and below the arrows. Notice an increasing number of independent origins of symbionts and decreasing number of phylogenetic artifacts along the continuum towards the 'derived' methods. 1+2: third codon positions excluded; AT/GC(BI/4-11): AT/GC datasets 4-11 analyzed by BI; BI: Bayesian inference; Dayhoff6/Dayhoff4/HP: amino acid recoded matrices; ML: maximum likelihood; nhPhyML: ML under nonhomogeneous model; MP: maximum parsimony; RY: purine/pyrimidine recoded matrix; SF: slow-fasted datasets.

Standard maximum likelihood and Bayesian inference

The single gene maximum likelihood (ML) analyses of both nucleic and amino acid data provided an array of mutually exclusive topologies. The majority consensus based on amino acid data (Additional file 2a) groups almost all symbionts into polytomy with only two pairs of sister symbiotic species being resolved (*Buchnera* and *Blochmannia*). Phylogenetic trees inferred by ML and Bayesian inference (BI) from the nucleic acid concatenated data using the General Time Reversible model with an estimated proportion of invariable sites (I) and heterogeneity of evolutionary rates modeled by the four substitution rate categories of the gamma (Γ) distribution with the gamma shape parameter (alpha) estimated from the data (GTR+I+Γ) were apparently affected by phylogenetic artifacts, as demonstrated by placement of *Riesia* and *Wigglesworthia* within the *Buchnera* cluster with high posterior probabilities in the BI tree (Figure 2) and the attraction of two outgroup species (*Haemophilus* and *Pasteurella*) in the ML tree with high bootstrap support (Additional file 2b). Similar topologies were also retrieved from the amino acid concatenate by ML and BI using the LG+I+Γ, WAG+I+Γ and GTR+I+Γ models (Figure 3). Nevertheless, in contrast to the nucleotide-derived results, the monophyly of the *Buchnera* clade was not disrupted and *Hamiltonella* and *Regiella* were unambiguously separated from the other symbionts and clustered with *Yersinia*.

PhyloBayes, non-homogenous PhyML and modified matrices

The phylogenetic trees acquired under the CAT+GTR PhyloBayes model from 14 and 55 concatenated genes (Figure 4 and Additional file 2p) split symbiotic bacteria into four and three independent lineages, respectively. First, *Arsenophonus nasoniae* is a sister species to *Proteus mirabilis*; second, *Hamiltonella* and *Regiella* form a sister clade to *Yersinia pestis*; third, the *Sodalis*, *Baumannia*, *Blochmannia*, *Wigglesworthia*, *Riesia* and *Buchnera* clade form a sister clade to *Dickeya*/*Pectobacterium*. The position of *Ishikawaella* differs between the two datasets. In the 14-gene dataset, *Ishikawaella* forms a sister clade to *Pantoea* (Figure 4) and in the 55-gene dataset, it is attracted to the P-symbiont cluster (Additional file 2p).

A topology with four independent symbiotic clades resulted from the trees derived from dayhoff6 and dayhoff4 recoded amino acid data sets analyzed by CAT and CAT+GTR models (Figure 5, Additional file 2r, q) and partially with the hp (hydrophobic-polar) recoded dataset (Additional file 2c) - which was on the other hand affected by the substantial loss of phylogenetic information. The first clade is *Buchnera*+*Ishikawaella* as a sister clade to the *Erwinia*/*Pantoea* clade, the second clade is *Riesia*+*Arsenophonus* as a sister clade to *Proteus*, the third clade is *Hamiltonella*+*Regiella* as a sister clade to *Yersinia*, and the last clade is composed of *Sodalis*, *Baumannia*, *Blochmannia* and *Wigglesworthia*.

The analyses testing each symbiont independently, using a CAT+GTR model on the dayhoff6 recoded datasets, resulted in topologies supporting multiple origins of endosymbiosis (Additional file 2s). *Arsenophonus* clusters with *Proteus*; *Hamiltonella* clusters with *Yersinia*; *Regiella* clusters with *Yersinia*; and *Sodalis*, *Blochmannia*, *Baumannia*, *Riesia* and *Wigglesworthia* grouped into polytomy with the basal enterobacterial clades. Most importantly, the *Buchnera* clade clusters as a sister clade to the *Erwinia* clade and *Ishikawaella* is placed in polytomy with the *Pantoea* and *Erwinia* clade.

The non-homogenous (nh) PhyML nucleotide analyses with two different starting trees resulted in two different topologies (Figure 6 and Additional file 2d, e, f). When compared by the approximately unbiased (AU) test, the topology with four independent origins of symbiotic bacteria prevailed ($P = 1$) over the topology with monophyly of P-symbionts, which therefore corresponds to a local minimum due to a tree search failure (complete matrix: $P = 2 \times 10^{-67}$; matrix without the third positions: $P = 9 \times 10^{-87}$). The only incongruence in topologies based on the complete matrix (Additional file 2d) and the matrix without the third positions (Figure 6) was the placement of the *Sodalis*+*Baumannia*+*Blochmannia*+*Wigglesworthia* clade as a sister clade to the *Edwardsiella* or *Dickeya*/*Pectobacterium* clades.

Matrices obtained by removing positions according to the AT/GC contents produced trees covering the whole continuum illustrated in Figure 1. The most severe restrictions, that is, removal of all positions that contain both AT and GC categories or relaxing for up to three taxa (see BI trees in Additional file 2g, h, i, j), yielded topologies compatible with the results of the CAT model applied on the recoded amino acid data and of the nhPhyML analysis. Further relaxing the restriction rule led to a variety of trees along the Figure 1 continuum, with a less clear relation between the used parameter and the resulting topology (Additional file 3).

Compared to the ML analysis of all nucleotide positions, the analysis of first plus second positions reduced the obvious artifact of outgroup attraction (Additional file 2k). Nevertheless, it also sorted symbionts according to their branch length. Analysis of the RY recoded nucleotide matrix produced a tree compatible with the results of the CAT+GTR model (Additional file 2l). Analysis of the RY recoded nucleotide matrix without the third positions resulted in a topology with a *Sodalis*+*Baumannia*+*Blochmannia* cluster (as a sister to the *Pectobacterium*/*Dickeya* clade) separated from the rest

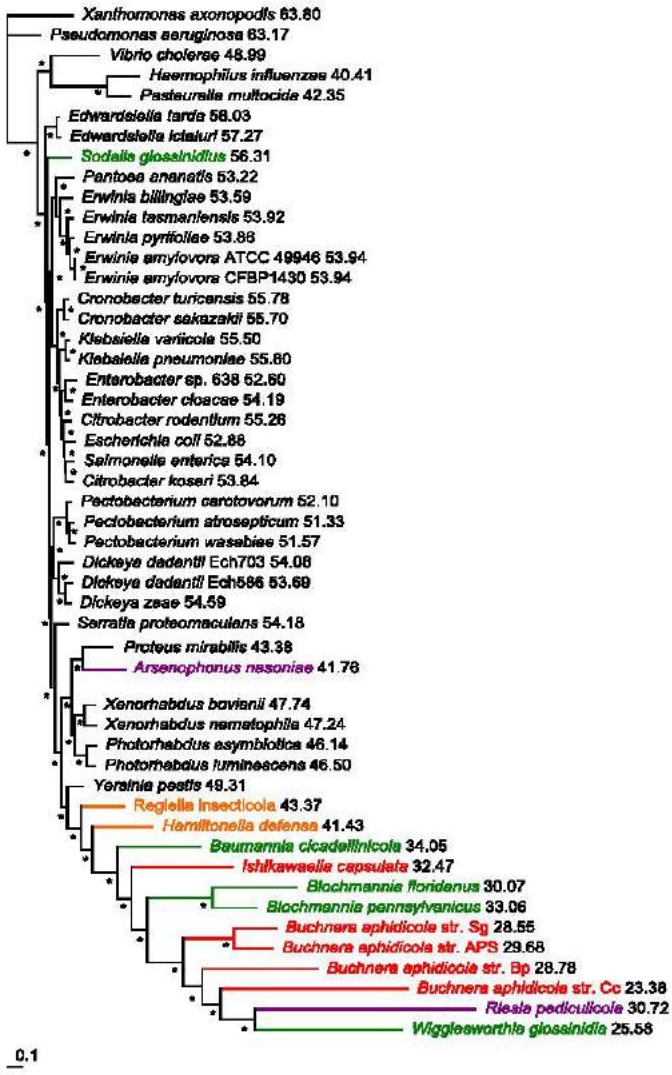


Figure 2 MrBayes phylogram - 69 genes, nucleotide matrix. Phylogenetic tree inferred from the concatenated nucleotide matrix using BI under the GTR+I+Γ model. Asterisks designate nodes with posterior probabilities equal to 1.0, values next to species names represent GC content calculated from the 69-gene dataset, genomic GC content can be found in Additional file 4. BI: Bayesian inference.

of the P-symbionts, which clustered with the *Erwinia*/*Pantoea* clade (Additional file 2m). Slow-fast analyses with gradual reduction of saturated positions did not produce the polyphyly of P-symbionts (Additional file 3; only the first five trees presented, subsequent trees are identical to the fifth tree). However, this analysis shows an increasing effect of LBA artifacts associated with the increasing number of remaining saturated positions, especially *Riesia* attraction and swapping of symbiotic branches according to their length.

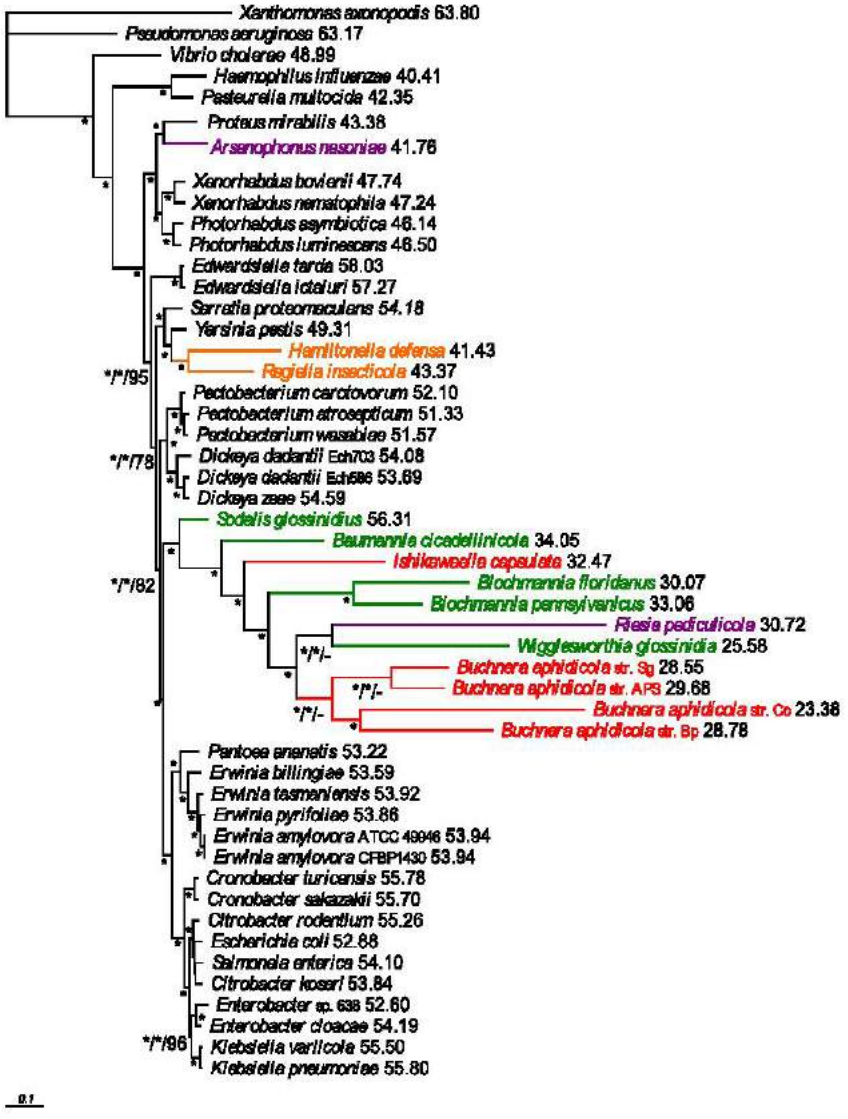


Figure 3 MrBayes phylogram - 69 genes, amino acid matrix. Phylogram inferred from the concatenated amino acid matrix using BI under the WAG+I+Γ model. Values at nodes represent posterior probabilities (WAG+I+Γ model, GTR+I+Γ protein model) and bootstrap supports from ML analysis (LG+I+Γ model). Asterisks designate nodes with posterior probabilities or bootstrap supports equal to 1.0, dashes designate values lower than 0.5 or 50, values next to species names represent GC content calculated from the 69-gene dataset, genomic GC content can be found in Additional file 4. BI: Bayesian inference. ML: maximum likelihood.

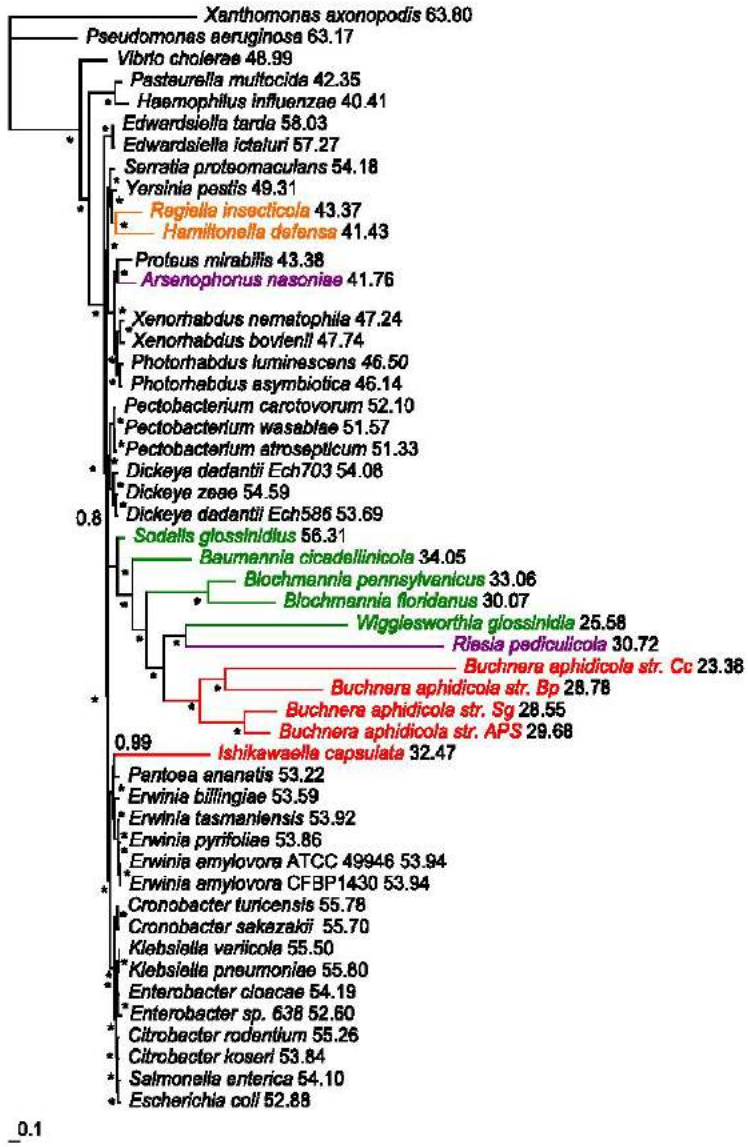


Figure 4 PhyloBayes phylogram - 14 genes, amino acid matrix. Phylogram derived from concatenation of 14 genes (*AceE*, *ArgS*, *AspS*, *EngA*, *GidA*, *GlyS*, *InfB*, *PheT*, *Pgi*, *Pnp*, *RpoB*, *RpoC*, *TrmE* and *YidC*) using PhyloBayes under the CAT+GTR model. Asterisks designate nodes with posterior probabilities equal to 1.0, values next to species names represent GC content calculated from the 69-gene dataset, genomic GC content can be found in Additional file 4.

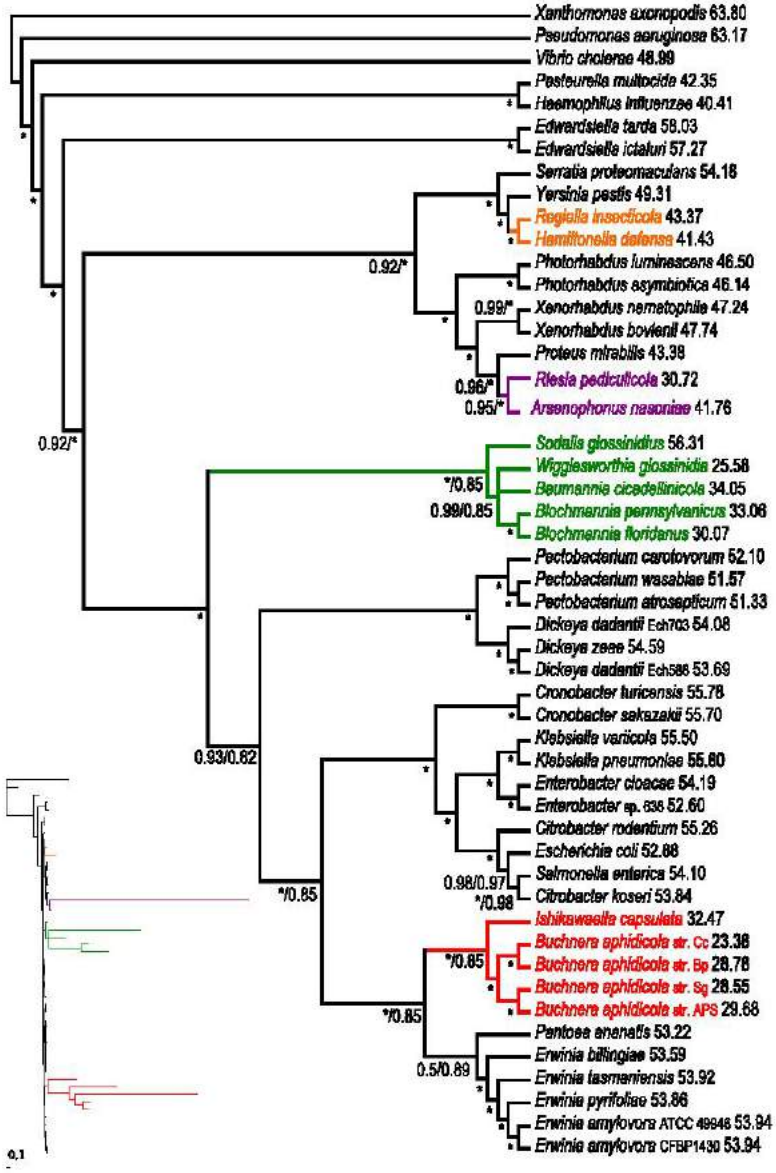


Figure 5 PhyloBayes cladogram - 69 genes, Dayhoff6 amino acid recoded matrix. Cladogram inferred from amino acid matrix recoded with Dayhoff6 scheme using PhyloBayes with the CAT and CAT+GTR model. Because of the length of symbiotic branches, phylogram is presented only as a preview (original phylogram can be found in Additional trees on our website). Values at nodes represent posterior probabilities from CAT and CAT+GTR analyses, respectively (asterisks designate nodes with posterior probabilities equal to 1.0). Values next to species names represent GC content calculated from the 69-gene dataset; genomic GC content can be found in Additional file 4.

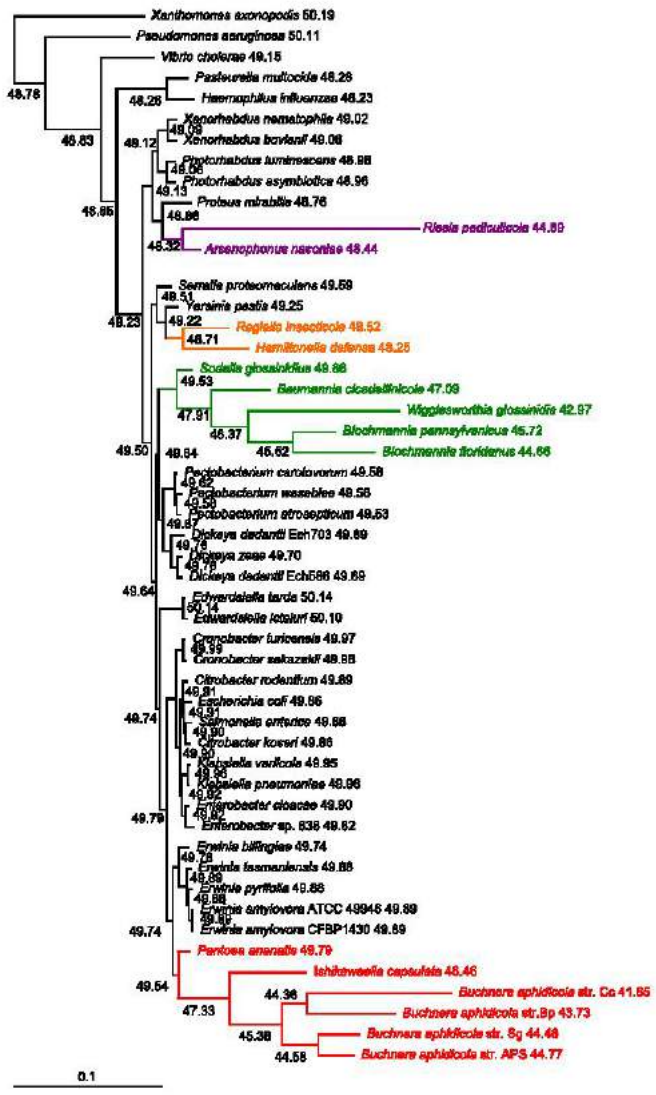


Figure 6 nhPhyML phylogram - 69 genes, nucleotide matrix, third positions excluded. Phylogram inferred from the concatenated nucleotide matrix without third codon positions using the nonhomogeneous model of evolution as implemented in nhPhyML. Values at nodes and branches represent GC content.

Discussion

Performance of the methods: convergence towards non-monophyly

The results obtained in this study strongly indicate that the frequently retrieved monophyly of P-symbionts is an artifact caused by their highly modified genomes. None of the most widely used methods, that is, ML and BI with different models used on nucleic (GTR+I+Γ) and amino acid (GTR/LG/WAG+I+Γ) data, were capable of resolving deep phylogenetic relationships and correct placement of the symbiotic taxa. This conclusion is evidenced by obvious artifacts, such as the inclusion of *Riesia* into the P-symbiotic lineage or the even more conspicuous distorted placing of *Wigglesworthia* within the *Buchnera* cluster. The arrangement of such trees suggests that these methods sort the symbionts according to their branch lengths and/or AT contents and attach the whole symbiotic cluster to the longest branch available. While the difficulty with placement of the most aberrant taxa, such as *Riesia*, *Wigglesworthia* and *Buchnera* (*Cinara cedri*) was also observed when using the mixture model accounting for site specific characteristics of protein evolution (Figure 4; Additional files 2p and 5), these artifacts disappeared after amino acid data recoding followed by CAT and CAT+GTR model analysis and the application of a nonhomogeneous model.

Additional support for the non-monophyly view stems from the second, parallel approach based on the restricted matrices. While our newly developed method shares the basic principles with the slow-fast and recoding methods, such as the removal of the positions that are likely to distort the phylogenetic relationships due to their aberrant evolution, it differs in the criteria of their removal and thus produces different input data. It is therefore significant that this method led independently to the same picture, the non-monophyly of the P-symbionts with clustering identical to the above analyses: *Ishikawaella+Buchnera* and *Sodalis+Baumannia+Blochmannia+Wigglesworthia*. The removal of the heteropecillous sites was recently shown to have similar effectiveness as our new method [61], which further supports the results. Moreover, this topology was obtained even under the maximum parsimony (MP) criterion (Additional file 3), which is known to be extremely sensitive to LBA [34]. On the other hand, although slow-fast analysis is generally considered a powerful tool for resolving relationships among taxa with different rates of evolution, we show in our data that the mere exclusion of the fast evolving sites is not sufficient when using empirical models and should be followed by analysis using some of the complex models, such as the CAT-like models. In addition, since this method usually requires an *a priori* definition of monophyletic groups, it should be used and interpreted with caution. Similar to the slow-fast method, RY recoding and exclusion of third codon positions were not sufficient for resolving deep symbiont phylogeny. However, all these methods can remove at least some of the artifacts and provide insight for further analyses.

Summarizing the topologies obtained in this study (Figure 1), a convergence can be detected towards a particular non-monophyletic arrangement of P-symbionts, as revealed under the most 'derived' methods. This result strongly supports the view of multiple origins of insect endosymbionts, as first revealed by the nonhomogeneous model of sequence evolution [40], and is partially congruent with the analyses of gene order [41] and phylogenomics of Gammaproteobacteria [42]. It is also important to note that, apart from multiple symbiont clustering, the arrangement of the non-symbiotic taxa corresponds to most of the phylogenomic analyses using *Escherichia/Salmonella/Yersinia* taxon sampling [37-39].

Biological significance of P-symbionts non-monophyly

Considering that most of the 'artifact-resistant' analyses point towards the non-monophyly of enterobacterial P-symbionts, the questions of how many symbiotic lineages are represented by the known symbiotic diversity and what are their closest free-living relatives now becomes of particular importance. It is not clear whether the split of the original P-symbiotic cluster into two lineages is definite or these two groups will be further divided after yet more sensitive methods and more complete data are available. At the moment there are still several clusters composed exclusively of derived symbiotic forms. In principle, three different processes may be responsible for the occurrence of such clusters: first, horizontal transmission of established symbiotic forms among host species; second, inadequate sampling with missing free-living relatives; or third, phylogenetic artifacts. All of these factors are likely to play a role in the current topological patterns. Being the main issues of this study, the role of methodological artifacts has been discussed above. Horizontal transmission, as the basis of non-artificial symbiotic clusters, is likely to take part at least in some cases. Perhaps the most convincing example is the *Wolbachia* cluster [62]: while within Enterobacteriaceae it may apply to *Arsenophonus*, *Sodalis* and possibly some other S-symbionts.

Recognition of the third cause, the incomplete sampling, and identification of the closest free-living relatives, now becomes a crucial step in future research. It is often assumed that symbionts originate from bacteria common to the environment typical for a given insect group. For example, cicadas spend most of their life cycle underground and feed primarily on plant roots.

Consequently, their α-Proteobacterial symbiont *Hodgkinia cicadicola* originated within Rhizobiales [19]. A similar ecological background can be noticed in yet different hosts, the ixodid and argasid ticks. Several reports have shown that some of the tick-transmitted pathogens are related to their symbiotic fauna [63-65]. Many of the insect taxa associated with symbiotic Enterobacteriaceae are phytophagous, and plant pathogens thus fit well into this hypothesis as hypothetical ancestors of various insect symbionts lineages. The presence of a type III secretion system, which is used in pathogenic bacteria for host cell invasion, in secondary symbionts [66-69] and its remnant in the primary symbiont of *Sitophilus* spp. weevils [70] could further support the theory of pathogenic ancestors of insect symbionts. It can only be speculated that these bacteria first became S-symbiont type and were horizontally transferred to various other insect species. Within some of the infected species, facultative symbionts eventually became obligatory primary symbionts. An identical situation can be observed in symbiotic clades with numerous species, such as *Wolbachia* [71,72], *Sodalis* [23,73,74] or *Arsenophonus* [5].

In our study, we gave particular attention to the sampling of free-living Enterobacteriaceae to provide as complete a background for the symbiotic lineages as possible under the current state of knowledge (that is, the availability of the genomic data). The most consistent picture derived from the presented analyses places the four main symbiotic clusters into the following positions. First, for the *Buchnera* cluster, its previously suggested relationship to *Erwinia* was confirmed. *Erwinia*, as a genus of mostly plant pathogenic bacteria, has been previously suggested to represent an ancestral organism, which upon ingestion by aphids at least 180 million years ago [75] turned into an intracellular symbiotic bacterium [76]. However, it is not known whether it was primarily pathogenic to aphids, similar to *Erwinia aphidicola* [77], or a gut associated symbiotic bacterium as in pentatomid stinkbugs [78], thrips [79,80] or Tephritidae flies [81-83]. *Ishikawaella capsulata*, an extracellular gut symbiont of plataspid stinkbugs [84], was the only symbiotic bacterium that clustered in our 'derived' analyses with the *Buchnera* clade. However, several single-gene studies indicate that this group contains some additional symbiotic lineages for which sequenced genome data is not currently available. These are, in particular, the extracellular symbionts of acanthosomatid stinkbugs [85], parastrachid stinkbugs [86], scutellerid stinkbugs [87,88] and some of the symbionts in pentatomid stinkbugs [78].

The second clade, represented in our analysis by *Sodalis+Baumannia+Blochmannia+Wigglesworthia*, is likely to encompass many other P- and S-symbionts [89-92]. The possible single origin of these symbionts has to be further tested, however the interspersion of both forms, together with basal position of *Sodalis*, seem to support a transition from a secondary to primary symbiotic lifestyle [15]. In our analysis, the whole clade was placed between pathogenic bacteria of plants and animals, the *Edwardsiella* and *Pectobacterium/Dickeya* clades, or as a sister to the latter group. Recently, another symbiotic bacterium (called BEV, *Euscelidius variegatus* host) was shown to be a sister species to *Pectobacterium* [93].

Two additional independent origins of insect symbionts are represented by the *Arsenophonus/Riesia* clade and *Hamiltonella+Regiella*. Both of these clades clustered in our analyses in the positions indicated by previous studies, that is, as related to *Proteus* and *Yersinia*, respectively [5,67,93-97].

While the position of individual symbiotic lineages is remarkably consistent across our 'artifact-resistant' analyses and are well compatible with some of the previous studies, the topology can only provide a rough picture of the relationships within Enterobacteriaceae. To get a more precise and phylogenetically meaningful background for an evolutionary interpretation, the sample of free-living bacteria as a possible source of symbiotic lineages has to be much improved. An illuminating example is provided by the bacterium *Biostraticola tofi*, described from water biofilms. When analyzed using 16S rDNA, this bacterium seemed to be closely related to *Sodalis* [98]. Its position as a sister group to the *Sodalis/Baumannia/Blochmannia/Wigglesworthia* clade was also retrieved in our single-gene analysis (*groEL*, data not shown). If confirmed by more precise multigene approach, *Biostraticola* would represent the closest bacterium to the large symbiotic cluster.

Conclusions

The topologies obtained by several independent approaches strongly support the non-monophyletic view of enterobacterial P-symbionts. Particularly, they show that at least three independent origins led to highly specialized symbiotic forms, the first giving rise to *Sodalis*, *Baumannia*, *Blochmannia* and *Wigglesworthia* (S- and P-symbionts), the second to *Buchnera* and *Ishikawella* and the last to *Riesia* and *Arsenophonus* (S- and P-symbionts). This separation of symbiotic clusters poses an interesting question as to whether the presented disbandment of the P-symbiotic cluster is definite or if it will continue after yet more complete data are available and more realistic evolutionary models [99-101] are applied. One obvious drawback of the current state is that many additional symbiotic lineages already known within Enterobacteriaceae cannot be at the moment included into serious phylogenetic analyses due to the lack of sufficient molecular data and will have to be revisited once complete genomic data are available.

These bacteria include symbionts of mealybugs [89,102], psyllids [90,103], lice [2,91], weevils [11,12,92], reed beetles [104,105], true bugs [78,84-88,106,107] and symbionts of leeches [108,109]. Similarly, the importance of free-living bacteria and variety of S-symbionts as possible ancestors of P-symbionts should not be underestimated when assembling datasets for phylogenetic analyses. The shift from polymerase chain reaction-based gene-centered sequencing towards high-throughput next-generation sequencing may soon provide sufficient data for more complete analyses of the Enterobacteriaceae phylogeny.

Methods

Matrices and multiple sequence alignments

The genes used in this study were extracted from 50 complete genome sequences of γ-Proteobacteria available in GenBank (Additional file 4), including 14 endosymbiotic Enterobacteriaceae. We did not include *Carsonella ruddii* [110] since this psyllid symbiotic bacterium does not appear to be a member of the Enterobacteriaceae clade [90,111] and is only attracted there by the AT rich taxa. After removal of the AT rich lineages from the analysis, *Carsonella ruddii* clusters with the genus *Pseudomonas* [42]. Also, we did not include *Serratia symbiotica* [95] because its genome only became available after completion of our datasets. However, the phylogenetic position of this symbiotic bacterium within *Serratia* genus is robust and was confirmed in several studies [6,14,112].

To minimize the introduction of a false phylogenetic signal, we compared the genomes of all symbiotic bacteria and selected only single-copy genes present in all of the included symbiotic and free-living taxa. Such strict gene exclusion was also necessary regarding the usage of computationally demanding methods; it was one of our goals to produce a taxonomically representative data set of efficient size with no missing data. Altogether, 69 orthologous genes, mostly involved in translation, ribosomal structure and biogenesis (Additional file 4) were selected according to the Clusters of Orthologous Groups of proteins (COGs) [113,114]. Single-gene nucleotide data sets were downloaded via their COG numbers from a freely available database (MicrobesOnline [115]).

All protein coding sequences were translated into amino acids in SeaView version 4 [116], aligned by the MAFFT version 6 L-INS-i algorithm [117] and toggled back to the nucleotide sequences. Ambiguously aligned positions (codons) were excluded by Gblocks v0.91b [118,119] with the following parameters: minimum number of sequences for a conserved position: 26; minimum number of sequences for a flanking position: 43; maximum number of contiguous nonconserved positions: 8; minimum length of a block: 10; allowed gap positions: with half. The resulting trimmed alignments were checked and manually corrected in BioEdit v7.0.5 [120]. Alignments were concatenated in SeaView. The 69 gene concatenate resulted in an alignment of 63, 462 nucleic acid positions with 42, 481 parsimony-informative and 48, 527 variable sites and 21, 154 amino acid positions with 12, 735 parsimony-informative and 15, 986 variable sites.

Phylogenetic analyses

We used two different approaches to deal with the distortions caused by the highly modified nature of symbiotic genomes, which are the main source of the phylogenetic artifacts in phylogenetic analyses.

First, we applied complex models of molecular evolution. Using PhyloBayes 3.2f [121], we applied non-parametric site heterogeneous CAT and CAT+GTR models [43]. For all PhyloBayes analyses, we ran two chains with an automatic stopping option set to end the chain when all discrepancies were lower than 0.3 and all effective sizes were larger than 100. Under the CAT and CAT+GTR models, the four independent PhyloBayes runs were stuck in a local maximum (maxdiff = 1) even after 25, 000 and 10, 000 cycles, respectively, and we were not able to reach Markov Chain Monte Carlo (MCMC) convergence. Therefore, we present these trees only as supplementary material (although they mostly point toward multiple origins of symbiosis; Additional file 5) and we ran the CAT+GTR analyses with the reduced dataset based on 14 genes with the number of parsimony-informative amino acid positions higher than 300 (*AceE*, *ArgS*, *AspS*, *EngA*, *GidA*, *GlyS*, *InfB*, *PheT*, *Pgi*, *Pnp*, *RpoB*, *RpoC*, *TrmE* and *YidC*). To check for compatibility of these arbitrary selected 14 genes with the rest of the data, we also analyzed, in a separate analysis, the remaining 55-gene dataset under the CAT +GTR model. Using nhPhyML [122], we applied a non-homogeneous nonstationary model of sequence evolution [123,124], which can deal with artifacts caused by compositional heterogeneity [40,125,126]. We used two different starting trees (Additional file 2n) and ran the analyses with and without the third codon positions. The resulting trees were evaluated by an AU test in CONSEL [127].

The second approach relies on the selective restriction of the data matrix. We used four previously established methods of data weighting and/or exclusion (see Background): RY data recoding, amino acid data recoding, exclusion of third codon positions and slow-fast analysis, and developed one additional method: since transition from G/C to A/T at many positions is a common homoplasy of symbiotic genomes, we removed from the matrix all positions containing both the G/C and A/T

states. All substitutions considered in the subsequent analyses thus included exclusively transversions within the A/T or G/C categories. To analyze an effect of this restriction on the reduction of the data, we prepared 11 matrices with a partially relaxed rule (removing all positions with AT+GC, allowing for one taxon exception, two taxa exception, and so on, up until a 10 taxa exception). Since this method has never been tested, we analyzed the restricted matrices by the BI, ML (parameters as for standard analyses) and MP using PAUP* 4.0b10 with the tree bisection and reconnection algorithm [128]. Four other types of data weighting and/or exclusion were used to increase the phylogenetic signal to noise ratio and determine the robustness of our results. First, the third codon positions were removed in SeaView. Second, RY recoding was performed on all and first plus second positions. Third, saturated positions were excluded from the concatenated data sets by SlowFaster [129]. To assign substitutional rates to individual positions, unambiguously monophyletic groups were chosen on a polytomic tree (Additional file 2o), positions with the highest rates were gradually excluded and 21 restricted matrices were produced. These weighted data sets were analyzed by ML. Fourth, amino acid data recoding was performed in PhyloBayes with hp (A, C, F, G, I, L, M, V, W) (D, E, H, K, N, P, Q, R, S, T, Y), dayhoff4 (A, G, P, S, T) (D, E, N, Q) (H, K, R) (F, Y, W, I, L, M, V) (C = ?) and dayhoff6 (A, G, P, S, T) (D, E, N, Q) (H, K, R) (F, Y, W) (I, L, M, V) (C) recoding schemes. In addition, we have prepared 10 dayhoff6 recoded matrices to test individual symbiotic lineages without the presence of other symbionts. Amino acid recoded matrices were analyzed using the CAT and CAT+GTR models, which are more immune to phylogenetic artifacts than one-matrix models.

To allow for comparison of the results with previously published studies, as well as to separate the effect of newly used models and methods from changes due to the extended sampling, we also used standard procedures of phylogenetic inference, ML and BI. The following programs, algorithms and parameters were used in the ML and BI analyses. ML was applied to single-gene and concatenated alignments of both nucleotides and amino acids using PhyML v3.0 [130] with the subtree pruning and regrafting tree search algorithm. BI was performed in MrBayes 3.1.2 [131] with one to five million generations and tree sampling every 100 generations. Exploration of MCMC convergence and burn-in determination was performed in AWTY and Tracer v1.5 [132,133]. Evolutionary substitution models for proteins were selected by ProtTest 2.4 [134] and for DNA by jModelTest 0.1.1 [135] according to the Akaike Information Criterion. For DNA sequences, the GTR+I+Γ model was used [136-138]. Transition and transversion models [139] were used with I+Γ under ML for the first two AT/GC datasets. LG+I+Γ [140], WAG+I+Γ [141] and GTR+I+Γ models were used for amino acid data. A cross-validation method implemented in PhyloBayes [121,142] was used to estimate the fit of CAT-like models. For both datasets, the 14 selected genes as well as the complete 69 genes set, the cross-validation was performed according to the PhyloBayes manual in 10 replicates each with 1, 100 cycles. The CAT-Poisson model had significantly better fit to the data than the GTR model (Δl 157.37 ± 56.9379 for the 14-gene matrix and Δl 3923.9 ± 1963.5 for the 69-gene matrix); of the CAT-like models, the CAT+GTR model was found to be significantly better than the CAT-Poisson model (Δl 536.71 ± 32.8341 for the 14-gene matrix and Δl 1633.4 ± 123.482 for the 69-gene matrix) in all 10 replicates.

Additional material

Additional file 1: Summary of 20 studies on symbionts phylogeny.

Additional file 2: Additional phylogenetic trees.

Additional file 3: All phylogenetic trees derived from AT-GC and SF datasets. A rar file of all phylogenetic trees obtained under BI, ML and MP from 11 AT/GC datasets, and under ML from five slow-fasted datasets. Trees are in phylip and nexus formats and can be viewed, for example, in TreeView http://taxonomy.zoology.gla.ac.uk/rod/treeview.html or Mesquite http://mesquiteproject.org/mesquite/mesquite.html.

Additional file 4: List of the taxa and orthologous genes used in the study.

Additional file 5: Additional phylogenetic trees inferred from CAT and CAT+GTR unconverged chains.

Acknowledgements

This work was supported by the Grant Agency of Academy of Sciences of the Czech Republic (grant number P505/10/1401 to VH); the Student Grant Agency of the Faculty of Science, University of South Bohemia (grant number SGA2009002 to FH); and the Ministry of Education, Youth, and Sports of the Czech Republic (grant numbers LC06073 and MSM 60076605801 to VH). Access to the MetaCentrum computing facilities provided under the 'Projects of Large Infrastructure for Research, Development, and Innovations' program LM2010005 funded by the Ministry of Education, Youth, and Sports of the Czech Republic is highly appreciated.

Author details

[1]Faculty of Science, University of South Bohemia, Branišovská 31, České Budějovice 37005, Czech Republic. [2]Institute of Parasitology, Biology Centre of ASCR, Branišovská 31, České Budějovice 37005, Czech Republic.

Authors' contributions

FH carried out the sequence alignments and phylogenetic analyses, and participated in the study design, evolutionary interpretation of the results and preparation of the manuscript. TCH compiled and analyzed the AT/GC reduced matrices. VH conceived of the study and participated in its design, evolutionary interpretation of the results and preparation of the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Received: 27 July 2011 Accepted: 28 December 2011
Published: 28 December 2011

References

1. Gottlieb Y, Ghanim M, Gueguen G, Kontsedalov S, Vavre F, Fleury F, Zchori-Fein E: **Inherited intracellular ecosystem: symbiotic bacteria share bacteriocytes in whiteflies.** *FASEB J* 2008, **22(7)**:2591-2599.
2. Hypša V, Křížek J: **Molecular evidence for polyphyletic origin of the primary symbionts of sucking lice (Phthiraptera, Anoplura).** *Microb Ecol* 2007, **54(2)**:242-251.
3. Bordenstein SR, Paraskevopoulos C, Hotopp JC, Sapountzis P, Lo N, Bandi C, Tettelin H, Werren JH, Bourtzis K: **Parasitism and mutualism in *Wolbachia*: what the phylogenomic trees can and cannot say.** *Mol Biol Evol* 2009, **26(1)**:231-241.
4. Caspi-Fluger A, Zchori-Fein E: **Do plants and insects share the same symbionts?** *Isr J Plant Sci* 2010, **58(2)**:113-119.
5. Nováková E, Hypša V, Moran NA: ***Arsenophonus*, an emerging clade of intracellular symbionts with a broad host distribution.** *BMC Microbiol* 2009, **9**:143.
6. Moran NA, Russell JA, Koga R, Fukatsu T: **Evolutionary relationships of three new species of Enterobacteriaceae living as symbionts of aphids and other insects.** *Appl Environ Microbiol* 2005, **71(6)**:3302-3310.
7. Baumann P: **Biology of bacteriocyte-associated endosymbionts of plant sap-sucking insects.** *Annu Rev Microbiol* 2005, **59**:155-189.
8. Douglas AE: **Mycetocyte symbiosis in insects.** *Biol Rev Camb Philos Soc* 1989, **64(4)**:409-434.
9. Moran NA, McCutcheon JP, Nakabachi A: **Genomics and evolution of heritable bacterial symbionts.** *Annu Rev Genet* 2008, **42**:165-190.
10. Takiya DM, Tran PL, Dietrich CH, Moran NA: **Co-cladogenesis spanning three phyla: leafhoppers (Insecta: Hemiptera: Cicadellidae) and their dual bacterial symbionts.** *Mol Ecol* 2006, **15(13)**:4175-4191.
11. Lefevre C, Charles H, Vallier A, Delobel B, Farrell B, Heddi A: **Endosymbiont phylogenesis in the Dryophthoridae weevils: evidence for bacterial replacement.** *Mol Biol Evol* 2004, **21(6)**:965-973.
12. Conord C, Despres L, Vallier A, Balmand S, Miquel C, Zundel S, Lemperiere G, Heddi A: **Long-term evolutionary stability of bacterial endosymbiosis in Curculionoidea: additional evidence of symbiont replacement in the Dryophthoridae family.** *Mol Biol Evol* 2008, **25(5)**:859-868.
13. Perez-Brocal V, Gil R, Ramos S, Lamelas A, Postigo M, Michelena JM, Silva FJ, Moya A, Latorre A: **A small microbial genome: the end of a long symbiotic relationship?** *Science* 2006, **314(5797)**:312-313.
14. Lamelas A, Perez-Brocal V, Gomez-Valero L, Gosalbes MJ, Moya A, Latorre A: **Evolution of the secondary symbiont "*Candidatus* Serratia symbiotica" in aphid species of the subfamily Lachninae.** *Appl Environ Microbiol* 2008, **74(13)**:4236-4240.
15. Wernegreen JJ, Kauppinen SN, Brady SG, Ward PS: **One nutritional symbiosis begat another: phylogenetic evidence that the ant tribe Camponotini acquired *Blochmannia* by tending sap-feeding insects.** *BMC Evol Biol* 2009, **9(1)**:292.
16. Moran NA, Tran P, Gerardo NM: **Symbiosis and insect diversification: an ancient symbiont of sap-feeding insects from the bacterial phylum Bacteroidetes.** *Appl Environ Microbiol* 2005, **71(12)**:8802-8810.
17. Wu D, Daugherty SC, Van Aken SE, Pai GH, Watkins KL, Khouri H, Tallon LJ, Zaborsky JM, Dunbar HE, Tran PL, Moran NA, Eisen JA: **Metabolic complementarity and genomics of the dual bacterial symbiosis of sharpshooters.** *PLoS Biol* 2006, **4(6)**:e188.
18. McCutcheon JP, Moran NA: **Functional convergence in reduced genomes of bacterial symbionts spanning 200 million years of evolution.** *Genome Biol Evol* 2010, **2**:708-718.
19. McCutcheon JP, McDonald BR, Moran NA: **Convergent evolution of metabolic roles in bacterial co-symbionts of insects.** *Proc Natl Acad Sci USA* 2009, **106(36)**:15394-15399.
20. McCutcheon JP, McDonald BR, Moran NA: **Origin of an alternative genetic code in the extremely small and GC-rich genome of a bacterial symbiont.** *Plos Genet* 2009, **5(7)**:e1000565.
21. McCutcheon JP, Moran NA: **Parallel genomic evolution and metabolic interdependence in an ancient symbiosis.** *Proc Natl Acad Sci USA* 2007, **104(49)**:19392-19397.
22. Snyder AK, McMillen CM, Wallenhorst P, Rio RV: **The phylogeny of *Sodalis*-like symbionts as reconstructed using surface-encoding loci.** *FEMS Microbiol Lett* 2011, **317(2)**:143-151.
23. Fukatsu T, Koga R, Smith WA, Tanaka K, Nikoh N, Sasaki-Fukatsu K, Yoshizawa K, Dale C, Clayton DH: **Bacterial endosymbiont of the slender pigeon louse, *Columbicola columbae*, allied to endosymbionts of grain weevils and tsetse flies.** *Appl Environ Microbiol* 2007, **73(20)**:6660-6668.
24. Koga R, Tsuchida T, Fukatsu T: **Changing partners in an obligate symbiosis: a facultative endosymbiont can compensate for loss of the essential endosymbiont *Buchnera* in an aphid.** *Proc R Soc Lond B* 2003, **270(1533)**:2543-2550.
25. Oliver KM, Degnan PH, Burke GR, Moran NA: **Facultative symbionts in aphids and the horizontal transfer of ecologically important traits.** *Annu Rev Entomol* 2010, **55**:247-266.
26. Moya A, Pereto J, Gil R, Latorre A: **Learning how to live together: genomic insights into prokaryote-animal symbioses.** *Nat Rev Genet* 2008, **9(3)**:218-229.
27. Gosalbes MJ, Latorre A, Lamelas A, Moya A: **Genomics of intracellular symbionts in insects.** *Int J Med Microbiol* 2010, **300(5)**:271-278.
28. Toft C, Andersson SGE: **Evolutionary microbial genomics: insights into bacterial host adaptation.** *Nat Rev Genet* 2010, **11(7)**:465-475.
29. Naum M, Brown EW, Mason-Gamer RJ: **Is 16S rDNA a reliable phylogenetic marker to characterize relationships below the family level in the Enterobacteriaceae?** *J Mol Evol* 2008, **66(6)**:630-642.
30. Mira A, Moran NA: **Estimating population size and transmission bottlenecks in maternally transmitted endosymbiotic bacteria.** *Microb Ecol* 2002, **44(2)**:137-143.
31. Wernegreen JJ, Moran NA: **Evidence for genetic drift in endosymbionts (*Buchnera*): analyses of protein-coding genes.** *Mol Biol Evol* 1999, **16(1)**:83-97.
32. Lambert JD, Moran NA: **Deleterious mutations destabilize ribosomal RNA in endosymbiotic bacteria.** *Proc Natl Acad Sci USA* 1998, **95(8)**:4458-4462.
33. Allen JM, Light JE, Perotti MA, Braig HR, Reed DL: **Mutational meltdown in primary endosymbionts: selection limits Muller's ratchet.** *Plos One* 2009, **4(3)**:e4969.
34. Bergsten J: **A review of long-branch attraction.** *Cladistics* 2005, **21(2)**:163-193.
35. Ruano-Rubio V, Fares MA: **Artifactual phylogenies caused by correlated distribution of substitution rates among sites and lineages: the good, the bad, and the ugly.** *Syst Biol* 2007, **56(1)**:68-82.
36. Charles H, Heddi A, Rahbe Y: **A putative insect intracellular endosymbiont stem clade, within the Enterobacteriaceae, infered from phylogenetic analysis based on a heterogeneous model of DNA evolution.** *C R Acad Sci Ser III Sci Vie* 2001, **324(5)**:489-494.
37. Comas I, Moya A, Gonzalez-Candelas F: **From phylogenetics to phylogenomics: the evolutionary relationships of insect endosymbiotic γ-Proteobacteria as a test case.** *Syst Biol* 2007, **56(1)**:1-16.
38. Lerat E, Daubin V, Moran NA: **From gene trees to organismal phylogeny in prokaryotes: the case of the γ-Proteobacteria.** *PLoS Biol* 2003, **1(1)**:e19.
39. Canback B, Tamas I, Andersson SG: **A phylogenomic study of endosymbiotic bacteria.** *Mol Biol Evol* 2004, **21(6)**:1110-1122.
40. Herbeck JT, Degnan PH, Wernegreen JJ: **Nonhomogeneous model of sequence evolution indicates independent origins of primary endosymbionts within the Enterobacteriales (γ-Proteobacteria).** *Mol Biol Evol* 2005, **22(3)**:520-532.
41. Belda E, Moya A, Silva FJ: **Genome rearrangement distances and gene order phylogeny in γ-Proteobacteria.** *Mol Biol Evol* 2005, **22(6)**:1456-1467.
42. Williams KP, Gillespie JJ, Sobral BW, Nordberg EK, Snyder EE, Shallom JM, Dickerman AW: **Phylogeny of Gammaproteobacteria.** *J Bacteriol* 2010, **192(9)**:2305-2314.
43. Lartillot N, Philippe H: **A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process.** *Mol Biol Evol* 2004, **21(6)**:1095-1109.
44. Lartillot N, Brinkmann H, Philippe H: **Suppression of long-branch attraction artefacts in the animal phylogeny using a site-heterogeneous model.** *BMC Evol Biol* 2007, **7**:S4.
45. Lartillot N, Philippe H: **Improvement of molecular phylogenetic inference and the phylogeny of Bilateria.** *Philos Trans R Soc Lond B Biol Sci* 2008, **363(1496)**:1463-1472.
46. Nesnidal MP, Helmkampf M, Bruchhaus I, Hausdorf B: **Compositional heterogeneity and phylogenomic inference of metazoan relationships.** *Mol Biol Evol* 2010, **27(9)**:2095-2104.
47. Philippe H, Brinkman FS, Martinez P, Riutort M, Baguna J: **Acoel flatworms are not platyhelminthes: evidence from phylogenomics.** *Plos One* 2007, **2(8)**:e717.

48. Brinkmann H, Philippe H: **Archaea sister group of Bacteria? Indications from tree reconstruction artifacts in ancient phylogenies.** *Mol Biol Evol* 1999, **16(6)**:817-825.
49. Brochier C, Philippe H: **A non-hyperthermophilic ancestor for bacteria.** *Nature* 2002, **417(6886)**:244.
50. Hampl V, Čepička I, Flegr J, Tachezy J, Kulda J: **Critical analysis of the topology and rooting of the parabasalian 16S rRNA tree.** *Mol Phylogenet Evol* 2004, **32(3)**:711-723.
51. Baptiste E, Susko E, Leigh J, Ruiz-Trillo I, Bucknam J, Doolittle WF: **Alternative methods for concatenation of core genes indicate a lack of resolution in deep nodes of the prokaryotic phylogeny.** *Mol Biol Evol* 2008, **25(1)**:83-91.
52. Philippe H, Lopez P, Brinkmann H, Budin K, Germot A, Laurent J, Moreira D, Muller M, Le Guyader H: **Early-branching or fast-evolving eukaryotes? An answer based on slowly evolving positions.** *Proc R Soc Lond B* 2000, **267(1449)**:1213-1221.
53. Hirt RP, Logsdon JM, Healy B, Dorey MW, Doolittle WF, Embley TM: **Microsporidia are related to Fungi: evidence from the largest subunit of RNA polymerase II and other proteins.** *Proc Natl Acad Sci USA* 1999, **96(2)**:580-585.
54. Philippe H, Lartillot N, Brinkmann H: **Multigene analyses of bilaterian animals corroborate the monophyly of Ecdysozoa, Lophotrochozoa, and Protostomia.** *Mol Biol Evol* 2005, **22(5)**:1246-1253.
55. Phillips MJ, Delsuc F, Penny D: **Genome-scale phylogeny and the detection of systematic biases.** *Mol Biol Evol* 2004, **21(7)**:1455-1458.
56. Delsuc F, Phillips MJ, Penny D: **Comment on "Hexapod origins: monophyletic or paraphyletic?".** *Science* 2003, **301(5639)**:1482.
57. Phillips MJ, Penny D: **The root of the mammalian tree inferred from whole mitochondrial genomes.** *Mol Phylogenet Evol* 2003, **28(2)**:171-185.
58. Gibson A, Gowri-Shankar V, Higgs PG, Rattray M: **A comprehensive analysis of mammalian mitochondrial genome base composition and improved phylogenetic methods.** *Mol Biol Evol* 2005, **22(2)**:251-264.
59. Hrdý I, Hirt RP, Doležal P, Bardonová L, Foster PG, Tachezy J, Embley TM: ***Trichomonas* hydrogenosomes contain the NADH dehydrogenase module of mitochondrial complex I.** *Nature* 2004, **432(7017)**:618-622.
60. Embley TM, van der Giezen M, Horner DS, Dyal PL, Foster P: **Mitochondria and hydrogenosomes are two forms of the same fundamental organelle.** *Philos Trans R Soc Lond B Biol Sci* 2003, **358(1429)**:191-201, discussion 201-203.
61. Philippe H, Roure B: **Site-specific time heterogeneity of the substitution process and its impact on phylogenetic inference.** *BMC Evol Biol* 2011, **11**:17.
62. Raychoudhury R, Baldo L, Oliveira DC, Werren JH: **Modes of acquisition of *Wolbachia*: horizontal transfer, hybrid introgression, and codivergence in the Nasonia species complex.** *Evolution* 2009, **63(1)**:165-183.
63. Perlman SJ, Hunter MS, Zchori-Fein E: **The emerging diversity of *Rickettsia*.** *Proc R Soc B* 2006, **273(1598)**:2097-2106.
64. Noda H, Munderloh UG, Kurtti TJ: **Endosymbionts of ticks and their relationship to *Wolbachia* spp. and tick-borne pathogens of humans and animals.** *Appl Environ Microbiol* 1997, **63(10)**:3926-3932.
65. Sassera D, Beninati T, Bandi C, Bouman EAP, Sacchi L, Fabbi M, Lo N: **'*Candidatus* Midichloria mitochondrii', an endosymbiont of the tick *Ixodes ricinus* with a unique intramitochondrial lifestyle.** *Int J Syst Evol Microbiol* 2006, **56**:2535-2540.
66. Degnan PH, Leonardo TE, Cass BN, Hurwitz B, Stern D, Gibbs RA, Richards S, Moran NA: **Dynamics of genome evolution in facultative symbionts of aphids.** *Environ Microbiol* 2009, **12(8)**:2060-2069.
67. Degnan PH, Yu Y, Sisneros N, Wing RA, Moran NA: ***Hamiltonella defensa*, genome evolution of protective bacterial endosymbiont from pathogenic ancestors.** *Proc Natl Acad Sci USA* 2009, **106(22)**:9063-9068.
68. Toh H, Weiss BL, Perkin SA, Yamashita A, Oshima K, Hattori M, Aksoy S: **Massive genome erosion and functional adaptations provide insights into the symbiotic lifestyle of *Sodalis glossinidius* in the tsetse host.** *Genome Res* 2006, **16(2)**:149-156.
69. Wilkes TE, Darby AC, Choi JH, Colbourne JK, Werren JH, Hurst GD: **The draft genome sequence of *Arsenophonus nasoniae*, son-killer bacterium of *Nasonia vitripennis*, reveals genes associated with virulence and symbiosis.** *Insect Mol Biol* 2010, **19(Sp. Iss. SI Suppl. 1)**:59-73.
70. Dale C, Plague GR, Wang B, Ochman H, Moran NA: **Type III secretion systems and the evolution of mutualistic endosymbiosis.** *Proc Natl Acad Sci USA* 2002, **99(19)**:12397-12402.
71. Hosokawa T, Koga R, Kikuchi Y, Meng XY, Fukatsu T: ***Wolbachia* as a bacteriocyte-associated nutritional mutualist.** *Proc Natl Acad Sci USA* 2010, **107(2)**:769-774.
72. Dedeine F, Vavre F, Fleury F, Loppin B, Hochberg ME, Bouletreau M: **Removing symbiotic *Wolbachia* bacteria specifically inhibits oogenesis in a parasitic wasp.** *Proc Natl Acad Sci USA* 2001, **98(11)**:6247-6252.
73. Heddi A, Charles H, Khatchadourian C, Bonnot G, Nardon P: **Molecular characterization of the principal symbiotic bacteria of the weevil *Sitophilus oryzae*: a peculiar G + C content of an endocytobiotic DNA.** *J Mol Evol* 1998, **47(1)**:52-61.
74. Nováková E, Hypša V: **A new *Sodalis* lineage from bloodsucking fly *Craterina melbae* (Diptera, Hippoboscoidea) originated independently of the tsetse flies symbiont *Sodalis glossinidius*.** *FEMS Microbiol Lett* 2007, **269(1)**:131-135.
75. Moran N, Munson M, Baumann P, Ishikawa H: **A molecular clock in endosymbiotic bacteria is calibrated using the insect hosts.** *Proc R Soc Lond B* 1993, **253(1337)**:167-171.
76. Harada H, Oyaizu H, Ishikawa H: **A consideration about the origin of aphid intracellular symbiont in connection with gut bacterial flora.** *J Gen Appl Microbiol* 1996, **42(1)**:17-26.
77. Harada H, Oyaizu H, Kosako Y, Ishikawa H: ***Erwinia aphidicola*, a new species isolated from pea aphid, *Acyrthosiphon pisum*.** *J Gen Appl Microbiol* 1997, **43(6)**:349-354.
78. Prado SS, Almeida RP: **Phylogenetic placement of pentatomid stink bug gut symbionts.** *Curr Microbiol* 2009, **58(1)**:64-69.
79. Chanbusarakum L, Ullman D: **Characterization of bacterial symbionts in *Frankliniella occidentalis* (Pergande), Western flower thrips.** *J Invertebr Pathol* 2008, **99(3)**:318-325.
80. De Vries EJ, Van der Wurff AWG, Jacobs G, Breeuwer JAJ: **Onion thrips, *Thrips tabaci*, have gut bacteria that are closely related to the symbionts of the western flower thrips, *Frankliniella occidentalis*.** *J Insect Sci* 2008, **8**:1-11.
81. Mazzon L, Martinez-Sañudo I, Simonato M, Squartini A, Savio C, Girolami V: **Phylogenetic relationships between flies of the Tephritinae subfamily (Diptera, Tephritidae) and their symbiotic bacteria.** *Mol Phylogenet Evol* 2010, **56(1)**:312-326.
82. Mazzon L, Piscedda A, Simonato M, Martinez-Sañudo I, Squartini A, Girolami V: **Presence of specific symbiotic bacteria in flies of the subfamily Tephritinae (Diptera Tephritidae) and their phylogenetic relationships: proposal of '*Candidatus* Stammerula tephritidis'.** *Int J Syst Evol Microbiol* 2008, **58(Pt 6)**:1277-1287.
83. Capuzzo C, Firrao G, Mazzon L, Squartini A, Girolami V: **'*Candidatus* Erwinia dacicola', a coevolved symbiotic bacterium of the olive fly *Bactrocera oleae* (Gmelin).** *Int J Syst Evol Microbiol* 2005, **55(Pt 4)**:1641-1647.
84. Hosokawa T, Kikuchi Y, Nikoh N, Shimada M, Fukatsu T: **Strict host-symbiont cospeciation and reductive genome evolution in insect gut bacteria.** *PLoS Biol* 2006, **4(10)**:e337.
85. Kikuchi Y, Hosokawa T, Nikoh N, Meng XY, Kamagata Y, Fukatsu T: **Host-symbiont co-speciation and reductive genome evolution in gut symbiotic bacteria of acanthosomatid stinkbugs.** *BMC Biol* 2009, **7**:2.
86. Hosokawa T, Kikuchi Y, Nikoh N, Meng XY, Hironaka M, Fukatsu T: **Phylogenetic position and peculiar genetic traits of a midgut bacterial symbiont of the stinkbug *Parastrachia japonensis*.** *Appl Environ Microbiol* 2010, **76(13)**:4130-4135.
87. Kaiwa N, Hosokawa T, Kikuchi Y, Nikoh N, Meng XY, Kimura N, Ito M, Fukatsu T: **Primary gut symbiont and secondary *Sodalis*-allied symbiont in the scutellerid stinkbug *Cantao ocellatus*.** *Appl Environ Microbiol* 2010, **76(11)**:3486-3494.
88. Kaiwa N, Hosokawa T, Kikuchi Y, Nikoh N, Meng XY, Kimura N, Ito M, Fukatsu T: **Bacterial symbionts of the giant jewel stinkbug *Eucorysses grandis* (Hemiptera: Scutelleridae).** *Zool Sci* 2011, **28(3)**:169-174.
89. Thao ML, Gullan PJ, Baumann P: **Secondary (γ-Proteobacteria) endosymbionts infect the primary (β-Proteobacteria) endosymbionts of mealybugs multiple times and coevolve with their hosts.** *Appl Environ Microbiol* 2002, **68(7)**:3190-3197.
90. Spaulding AW, von Dohlen CD: **Phylogenetic characterization and molecular evolution of bacterial endosymbionts in psyllids (Hemiptera: Sternorrhyncha).** *Mol Biol Evol* 1998, **15(11)**:1506-1513.
91. Fukatsu T, Hosokawa T, Koga R, Nikoh N, Kato T, Hayama S, Takefushi H, Tanaka I: **Intestinal endocellular symbiotic bacterium of the macaque louse *Pedicinus obtusus*: Distinct endosymbiont origins in anthropoid**

primate lice and the old world monkey louse. *Appl Environ Microbiol* 2009, **75(11)**:3796-3799.

92. Toju H, Hosokawa T, Koga R, Nikoh N, Meng XY, Kimura N, Fukatsu T: "*Candidatus* Curculioniphilus buchneri", a novel clade of bacterial endocellular symbionts from weevils of the genus *Curculio*. *Appl Environ Microbiol* 2009, **76(1)**:275-282.
93. Degnan PH, Bittleston LS, Hansen AK, Sabree ZL, Moran NA, Almeida RP: Origin and examination of a leafhopper facultative endosymbiont. *Curr Microbiol* 2011, **62(5)**:1565-1572.
94. Allen JM, Reed DL, Perotti MA, Braig HR: Evolutionary relationships of "*Candidatus* Riesia spp., " endosymbiotic *Enterobacteriaceae* living within hematophagous primate lice. *Appl Environ Microbiol* 2007, **73(5)**:1659-1664.
95. Burke GR, Moran NA: Massive genomic decay in *Serratia symbiotica*, a recently evolved symbiont of aphids. *Genome Biol Evol* 2011, **3**:195-208.
96. Ghena RL, Werren JH, Weisburg W, Cote R, Woese CR, Mandelco L, Brenner DJ: *Arsenophonus nasoniae* gen. nov, sp. nov, the causative agent of the son-killer trait in the parasitic wasp *Nasonia vitripennis*. *Int J Syst Bacteriol* 1991, **41(4)**:563-565.
97. Dale C, Beeton M, Harbison C, Jones T, Pontes M: Isolation, pure culture, and characterization of "*Candidatus* Arsenophonus arthropodicus, " an intracellular secondary endosymbiont from the hippoboscid louse fly *Pseudolynchia canariensis*. *Appl Environ Microbiol* 2006, **72(4)**:2997-3004.
98. Verbarg S, Fruhling A, Cousin S, Brambilla E, Gronow S, Lunsdorf H, Stackebrandt E: *Blastraticola tofi* gen. nov., spec. nov., a novel member of the family *Enterobacteriaceae*. *Curr Microbiol* 2008, **56(6)**:603-608.
99. Blanquart S, Lartillot N: A site- and time-heterogeneous model of amino acid replacement. *Mol Biol Evol* 2008, **25(5)**:842-858.
100. Blanquart S, Lartillot N: A Bayesian compound stochastic process for modeling nonstationary and nonhomogeneous sequence evolution. *Mol Biol Evol* 2006, **23(11)**:2058-2071.
101. Foster PG: Modeling compositional heterogeneity. *Syst Biol* 2004, **53(3)**:485-495.
102. Kono M, Koga R, Shimada M, Fukatsu T: Infection dynamics of coexisting Beta- and Gammaproteobacteria in the nested endosymbiotic system of mealybugs. *Appl Environ Microbiol* 2008, **74(13)**:4175-4184.
103. Thao ML, Clark MA, Baumann L, Brennan EB, Moran NA, Baumann P: Secondary endosymbionts of psyllids have been acquired multiple times. *Curr Microbiol* 2000, **41(4)**:300-304.
104. Kolsch G, Matz-Grund C, Pedersen BV: Ultrastructural and molecular characterization of endosymbionts of the reed beetle genus *Macroplea* (Chrysomelidae, Donaciinae), and proposal of "*Candidatus* Macropleicola appendiculatae" and "*Candidatus* Macropleicola muticae". *Can J Microbiol* 2009, **55(11)**:1250-1260.
105. Kolsch G, Pedersen BV: Can the tight co-speciation between reed beetles (Col., Chrysomelidae, Donaciinae) and their bacterial endosymbionts, which provide cocoon material, clarify the deeper phylogeny of the hosts? *Mol Phylogenet Evol* 2010, **54(3)**:810-821.
106. Kuchler SM, Dettner K, Kehl S: Molecular characterization and localization of the obligate endosymbiotic bacterium in the birch catkin bug *Kleidocerys resedae* (Heteroptera *Lygaeidae*, Ischnorhynchinae). *FEMS Microbiol Ecol* 2010, **73(2)**:408-418.
107. Kuechler SM, Dettner K, Kehl S: Characterization of an obligate intracellular bacterium in the midgut epithelium of the bulrush bug *Chilacis typhae* (Heteroptera, Lygaeidae, Artheneinae). *Appl Environ Microbiol* 2011, **77(9)**:2869-2876.
108. Kikuchi Y, Fukatsu T: Endosymbiotic bacteria in the esophageal organ of glossiphoniid leeches. *Appl Environ Microbiol* 2002, **68(9)**:4637-4641.
109. Perkins SL, Budinoff RB, Siddall ME: New Gammaproteobacteria associated with blood-feeding leeches and a broad phylogenetic analysis of leech endosymbionts. *Appl Environ Microbiol* 2005, **71(9)**:5219-5224.
110. Nakabachi A, Yamashita A, Toh H, Ishikawa H, Dunbar HE, Moran NA, Hattori M: The 160-kilobase genome of the bacterial endosymbiont *Carsonella*. *Science* 2006, **314(5797)**:267.
111. Thao ML, Baumann P: Evolutionary relationships of primary prokaryotic endosymbionts of whiteflies and their hosts. *Appl Environ Microbiol* 2004, **70(6)**:3401-3406.
112. Burke GR, Normark BB, Favret C, Moran NA: Evolution and diversity of facultative symbionts from the aphid subfamily Lachninae. *Appl Environ Microbiol* 2009, **75(16)**:5328-5335.
113. Tatusov RL, Koonin EV, Lipman DJ: A genomic perspective on protein families. *Science* 1997, **278(5338)**:631-637.
114. Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV, Krylov DM, Mazumder R, Mekhedov SL, Nikolskaya AN, Rao BS, Smirnov S, Sverdlov AV, Vasudevan S, Wolf YI, Yin JJ, Natale DA: The COG database: an updated version includes eukaryotes. *BMC Bioinformatics* 2003, **4**:41.
115. Dehal PS, Joachimiak MP, Price MN, Bates JT, Baumohl JK, Chivian D, Friedland GD, Huang KH, Keller K, Novichkov PS, Dubchak IL, Alm EJ, Arkin AP: MicrobesOnline: an integrated portal for comparative and functional genomics. *Nucleic Acids Res* 2010, , 38 Database: D396-D400.
116. Gouy M, Guindon S, Gascuel O: SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* 2010, **27(2)**:221-224.
117. Katoh K, Toh H: Recent developments in the MAFFT multiple sequence alignment program. *Brief Bioinform* 2008, **9(4)**:286-298.
118. Talavera G, Castresana J: Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst Biol* 2007, **56(4)**:564-577.
119. Castresana J: Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 2000, **17(4)**:540-552.
120. Hall TA: BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 1999, **41**:95-98.
121. Lartillot N, Lepage T, Blanquart S: PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics* 2009, **25(17)**:2286-2288.
122. Boussau B, Gouy M: Efficient likelihood computations with nonreversible models of evolution. *Syst Biol* 2006, **55(5)**:756-768.
123. Galtier N, Gouy M: Inferring pattern and process: maximum-likelihood implementation of a nonhomogeneous model of DNA sequence evolution for phylogenetic analysis. *Mol Biol Evol* 1998, **15(7)**:871-879.
124. Galtier N, Gouy M: Inferring phylogenies from DNA sequences of unequal base compositions. *Proc Natl Acad Sci USA* 1995, **92(24)**:11317-11321.
125. Tamo R, Rodriguez-Trelles F, Ayala FJ: Shared nucleotide composition biases among species and their impact on phylogenetic reconstructions of the Drosophilidae. *Mol Biol Evol* 2001, **18(8)**:1464-1473.
126. Galtier N, Tourasse N, Gouy M: A nonhyperthermophilic common ancestor to extant life forms. *Science* 1999, **283(5399)**:220-221.
127. Shimodaira H, Hasegawa M: CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics* 2001, **17(12)**:1246-1247.
128. Swofford DL: PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4. Sunderland, Massachusetts: Sinauer Associates; 2002.
129. Kostka M, Uzlíková M, Čepička I, Flegr J: SlowFaster, a user-friendly program for slow-fast analysis and its application on phylogeny of *Blastocystis*. *BMC Bioinformatics* 2008, **9**:341.
130. Guindon S, Gascuel O: A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 2003, **52(5)**:696-704.
131. Ronquist F, Huelsenbeck JP: MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 2003, **19(12)**:1572-1574.
132. Nylander JA, Wilgenbusch JC, Warren DL, Swofford DL: AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics* 2008, **24(4)**:581-583.
133. Rambaut A, Drummond AJ: Tracer v1.5. 2009 [http://tree.bio.ed.ac.uk/software/tracer/].
134. Abascal F, Zardoya R, Posada D: ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* 2005, **21(9)**:2104-2105.
135. Posada D: jModelTest: phylogenetic model averaging. *Mol Biol Evol* 2008, **25(7)**:1253-1256.
136. Tavaré S: Some probabilistic and statistical problems in the analysis of DNA sequences. *Lect Math Life Sci* 1986, **17**:57-86.
137. Lanave C, Preparata G, Saccone C, Serio G: A new method for calculating evolutionary substitution rates. *J Mol Evol* 1984, **20(1)**:86-93.
138. Yang Z: Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *J Mol Evol* 1994, **39(3)**:306-314.
139. Posada D: Using MODELTEST and PAUP* to select a model of nucleotide substitution. *Curr Protoc Bioinformatics* 2003, 6.5.1-6.5.14.
140. Le SQ, Gascuel O: An improved general amino acid replacement matrix. *Mol Biol Evol* 2008, **25(7)**:1307-1320.
141. Whelan S, Goldman N: A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol Biol Evol* 2001, **18(5)**:691-699.

142. Smyth P: Model selection for probabilistic clustering using cross-validated likelihood. Stat Comput 2000, 10(1):63-72.

doi:10.1186/1741-7007-9-87
Cite this article as: Husnik et al.: Multiple origins of endosymbiosis within the Enterobacteriaceae (γ-Proteobacteria): convergence of complex phylogenetic approaches. BMC Biology 2011 9:87.

4.2 Chapter 2

Chrudimský T., Husník F., Nováková E., Hypša V., 2012. *Candidatus* Sodalis melophagi sp. nov.: Phylogenetically Independent Comparative Model to the Tsetse Fly Symbiont *Sodalis glossinidius*. PLoS One. 2012;7(7):e40354. doi: 10.1371/journal.pone.0040354. Epub 2012 Jul 17.

Abstract

Bacteria of the genus *Sodalis* live in symbiosis with various groups of insects. The best known member of this group, a secondary symbiont of tsetse flies *Sodalis glossinidius*, has become one of the most important models in investigating establishment and evolution of insect-bacteria symbiosis. It represents a bacterium in the early/intermediate state of the transition towards symbiosis, which allows for exploring such interesting topics as: usage of secretory systems for entering the host cell, tempo of the genome modification, and metabolic interaction with a coexisting primary symbiont. In this study, we describe a new *Sodalis* species which could provide a useful comparative model to the tsetse symbiont. It lives in association with *Melophagus ovinus*, an insect related to tsetse flies, and resembles *S. glossinidius* in several important traits. Similar to *S. glossinidius*, it cohabits the host with another symbiotic bacterium, the bacteriome-harbored primary symbiont of the genus *Arsenophonus*. As a typical secondary symbiont, *Candidatus* Sodalis melophagi infects various host tissues, including bacteriome. We provide basic morphological and molecular characteristics of the symbiont and show that these traits also correspond to the early/intermediate state of the evolution towards symbiosis. Particularly, we demonstrate the ability of the bacterium to live in insect cell culture as well as in cell-free medium. We also provide basic characteristics of type three secretion system and using three reference sequences (16 S rDNA, *groEL* and *spaPQR* region) we show that the bacterium branched within the genus *Sodalis*, but originated independently of the two previously described symbionts of hippoboscoids. We propose the name *Candidatus* Sodalis melophagi for this new bacterium.

OPEN ACCESS Freely available online

PLoS one

Candidatus Sodalis melophagi sp. nov.: Phylogenetically Independent Comparative Model to the Tsetse Fly Symbiont *Sodalis glossinidius*

Tomáš Chrudimský[1]*, Filip Husník[1,2], Eva Nováková[1,2], Václav Hypša[1,2]

1 Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic, **2** Institute of Parasitology, Biology Centre, Academy of Sciences of the Czech Republic, České Budějovice, Czech Republic

Abstract

Bacteria of the genus *Sodalis* live in symbiosis with various groups of insects. The best known member of this group, a secondary symbiont of tsetse flies *Sodalis glossinidius*, has become one of the most important models in investigating establishment and evolution of insect-bacteria symbiosis. It represents a bacterium in the early/intermediate state of the transition towards symbiosis, which allows for exploring such interesting topics as: usage of secretory systems for entering the host cell, tempo of the genome modification, and metabolic interaction with a coexisting primary symbiont. In this study, we describe a new *Sodalis* species which could provide a useful comparative model to the tsetse symbiont. It lives in association with *Melophagus ovinus*, an insect related to tsetse flies, and resembles *S. glossinidius* in several important traits. Similar to *S. glossinidius*, it cohabits the host with another symbiotic bacterium, the bacteriome-harbored primary symbiont of the genus *Arsenophonus*. As a typical secondary symbiont, *Candidatus* Sodalis melophagi infects various host tissues, including bacteriome. We provide basic morphological and molecular characteristics of the symbiont and show that these traits also correspond to the early/intermediate state of the evolution towards symbiosis. Particularly, we demonstrate the ability of the bacterium to live in insect cell culture as well as in cell-free medium. We also provide basic characteristics of type three secretion system and using three reference sequences (16 S rDNA, *groEL* and *spaPQR* region) we show that the bacterium branched within the genus *Sodalis*, but originated independently of the two previously described symbionts of hippoboscoids. We propose the name *Candidatus* Sodalis melophagi for this new bacterium.

Citation: Chrudimský T, Husník F, Nováková E, Hypša V (2012) *Candidatus* Sodalis melophagi sp. nov.: Phylogenetically Independent Comparative Model to the Tsetse Fly Symbiont *Sodalis glossinidius*. PLoS ONE 7(7): e40354. doi:10.1371/journal.pone.0040354

Editor: Adam Driks, Loyola University Medical Center, United States of America

Received November 17, 2011; **Accepted** June 7, 2012; **Published** July 17, 2012

Funding: The study was supported by grants LC06073 and MSM6007665801 (Ministry of Education, Youth and Sports of the Czech Republic) and grant P505/10/1401 (Grant agency of the Czech Republic). Access to the MetaCentrum computing facilities was provided by "Projects of Large Infrastructure for Research, Development, and Innovations" LM2010005 (Ministry of Education, Youth, and Sports of the Czech Republic). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: tomas.chrudimsky@gmail.com

Introduction

The genus *Sodalis* belongs to the symbiotic bacterial lineages that adopted several different types of symbiosis with their hosts, ranging from facultative commensals to obligate mutualists [1,2]. *Sodalis* and closely related bacteria were described from a broad spectrum of insect hosts including tsetse flies [3], weevils [4–7], chewing lice [8], hippoboscid louse flies [9], ants [10], scale insects [11], aphids [12], stinkbugs [13,14], and cerambycid beetles [15]. Also, some "secondary" symbionts of psyllids and mealybugs cluster with the *Sodalis* clade [16,17]. Within symbiotic Enterobacteriaceae, diversity of the *Sodalis* clade is comparable only with the genus *Arsenophonus* [18].

The first described, best known and most frequently investigated member of the genus is *S. glossinidius*, S-symbiont of tsetse flies [3]. Its significance for the host is still not clear, but a possible influence on the host longevity and resistance to trypanosomes has previously been suggested [19]. Several molecular analyses and genetic experiments made *S. glossinidius* an important model for investigating evolution and biology of symbiotic bacteria [20–22]. It has been employed in research of various biological traits, such as the structure and role of secretion systems [21], the function of the iron acquisition system [23] or the usage of the quorum sensing system [24]. Moreover, *S. glossinidius* proved to be among the few symbionts that could be maintained in *in vitro* culture in insect cells as well as in the cell-free media [3,25]. This feature has been attributed to the initial or intermediate state of the *S. glossinidius* shift towards symbiosis. One of the most frequently discussed topics in this respect is the state and function of the type three secretion system (TTSS) in this bacterium. Three different copies of the TTSS (SSR-1, SSR-2 and SSR-3) has been detected in *S. glossinidius* [26,27] and a possible role of the SSR-2 in invading host cells has been proposed [21].

Within the host, *S. glossinidius* constitutes part of a complex bacterial community which also contains P-symbiont *Wigglesworthia glossinidia* [28] and alphaproteobacterium *Wolbachia* [29,30]. Recent investigations show that the whole community may be even richer and contain an array of other bacteria [31]. Such complex host-symbiont systems provide a unique opportunity for comparing genomes in different states/modes of symbiosis and studying processes of their metabolic complementation [32,33]. For the *Sodalis-Wigglesworthia-Glossina* association,

complete genomes of both bacteria have been sequenced and annotated [26,34]. Even though each of the sequenced genomes came from different host species, an interaction between *Sodalis* and *Wigglesworthia* via thiamine synthesis could be detected by their comparison [27] and this view was further corroborated by an experimental approach [35] and sequencing of *Wigglesworthia* lineage from *Glossina morsitans* [36].

An establishment of a more complete picture of *Sodalis* genome evolution will require complete genome sequences for a more diverse array of *Sodalis* isolates, and of similar complex systems involving this bacterium. Although tsetse flies are the most important blood feeding brachycerans, there are several related groups of dipterans that display many similar features such as the feeding strategy, vivipary and transmission of trypanosomes.

A member of the genus *Sodalis*, phylogenetically independent on the tsetse symbiont, has already been described from hippoboscoid species *Craterina melbae* [9]. Here, we characterize a new member of the *Sodalis* lineage, inhabiting gut and other tissues of another hippoboscoid, *Melophagus ovinus*. The presence of several symbiotic bacteria in this species has long been known. According to morphological investigations of several researchers summarized by Paul Buchner [37], *Melophagus ovinus* contains symbiotic bacteria within enlarged epithelial cells of a specialized section of the midgut (bacteriome). This P-symbiont was recently characterized by molecular techniques as a member of *Arsenophonus* clade and is likely to play a role resembling that of *Wigglesworthia* in tsetse flies (Nováková et al., in prep). In addition, some of the authors recognized two other bacteria in the sheep keds. The first is *Bartonella melophagi* (originally described as *Rickettsia melophagi* and *Wolbachia melophagi*), which is localized extracellularly along microvilli of the midgut. The second type of bacterium described in Paul Buchner's work resembles *Candidatus* Sodalis melophagi as presented in this study: *"In the low zones of the midgut epithelium…, there are additional delicate bacteria, sometimes forming rather long filaments, which also must not be confused with the symbionts."* The whole system thus remarkably resembles the *Wigglesworthia-Sodalis* association in tsetse flies and can provide important data for a comparative study. In this study, we present a basic molecular and morphological characterization of the new *Sodalis* linage and overview the composition of its TTSS. We suggest the new name *Candidatus* Sodalis melophagi for this bacterium and extend the available *Sodalis* spectrum with three additional samples which allow for more precise phylogenetic characterization.

Results

Sequence Data

Sequences obtained by PCR for the investigated samples and their accession numbers are summarized in Table S1. The 16 S rDNA sequence from *Candidatus* Sodalis melophagi sp. nov. displayed 98.48% similarity to 16 S rDNA of *S. glossinidius*. Illumina assemblies produced preliminary draft sequences from which only selected gene regions were used here for the formal description and basic phylogenetic characterization. These regions, including all TTSS genes (Table S1) and *groEL* chaperonin, were of a high quality and did not contain any SNPs. In order to avoid assembly artifacts affecting the sequence accuracy of highly similar paralogous regions, a partial sequence for 16 S rRNA gene was obtained through Sanger sequencing as described above and was used for inferring phylogeny. Sanger sequenced *groEL* partial sequence was identical to the sequence acquired from Illumina data and we therefore used the full length *groEL* for phylogenetic reconstruction.

Type Three Secretion System (TTSS)

Candidatus Sodalis melophagi possesses only SSR-2 and SSR-3 copies of the TTSS in its genome, while SSR-1 is completely missing. Extensive BLASTX searches did not produce any significant hits for *S. glossinidius* SSR-1 either in Illumina contigs or raw reads. The gene order and content of SSR-2 and SSR-3 of *Candidatus* Sodalis melophagi is very similar, but not identical to that in *S. glossinidius* (Figure 1). The SSR-2 sequence comprises 11 protein coding genes: *orgAbA*, *prgKIH*, *spaSQPO*, *invF* and *hilA*; and 5 pseudogenes: *prgJ*, *sicA*, *spaR* and *invAG*. In comparison with *S. glossinidius*, it lacks 4 genes: *invB*, *invC* and *spaLMN*. The SSR-3 comprises 29 protein coding genes: *ssrAB*, *ssaBCDEGHIJKLMN OPQRSTUV*, *sseABCDE*, *sscB*, a protein similar to locus SG1296 of *S. glossinidius*, and a single pseudogene: *sscA*. Sequences of *Candidatus* Sodalis melophagi TTSS genes were deposited in GenBank as a part of two annotated contigs for each of the islands (Table S1).

Phylogenetic Analyses

The lengths of individual matrices and numbers of variable positions are summarized in Table S2. All phylogenetic trees clearly indicate that the novel bacterium belongs to the genus *Sodalis* (Figures 2, 3, 4, Figure S1, Figure S2). In the trees derived from 16 S rRNA gene sequence data and amino acid sequence of *groEL*, *Candidatus* Sodalis melophagi clusters within a large polytomy and its precise position within the genus is thus uncertain (Figures 2–3A, Figure S1, Figure S2). However, even this unresolved topology excludes its relationship to any of the two other hippoboscoid-derived *Sodalis* members. This conclusion is further supported by the nucleotide matrix for *groEL* (Figure 3B) with a reduced sampling and the *spaPQR* concatenate (Figure 4). Although the exact topology slightly varies with the methods and parameters of the analysis, the three hippoboscoid lineages always form a polyphyletic/paraphyletic assemblage.

In vitro Culture

Bacterial colonies were clearly visible after 8 days of cultivation. Colonies were white, raised, and circular with entire edges. Their size was irregular ranging from 0.5 to 1 mm. The variable size of the colonies was almost certainly not due to new mutations, since subculturing on fresh plates yielded the same variability for each individual colony. The irregular colony morphology was described also in case of a type strain $M1^T$ of the *Sodalis glossinidius* [3] and is probably population dependent as is the case in other microaerophilic bacteria [38].

The type strain CZ^T was established by isolation of a single bacterial colony and was used for C6/36 cells infection. In the C6/36 cell culture, bacteria were predominantly attached to cell surface or free in the medium, but they were also observed inside the cells. Genomic DNA purified from bacterial colonies of *Candidatus* Sodalis melophagi CZ^T was used as a template for Illumina mate pair sequencing.

Microscopy

Under electron microscopy, *Sodalis* cells within the host tissue corresponded well to their light microscope characteristics (Figure 5A). They appeared as rods reaching from approx. 1 to 4 μm, depending on the angle of the section, and were located mainly at the periphery of the bacteriome, sometimes in close association with the P-symbionts (Figures 5B,C).

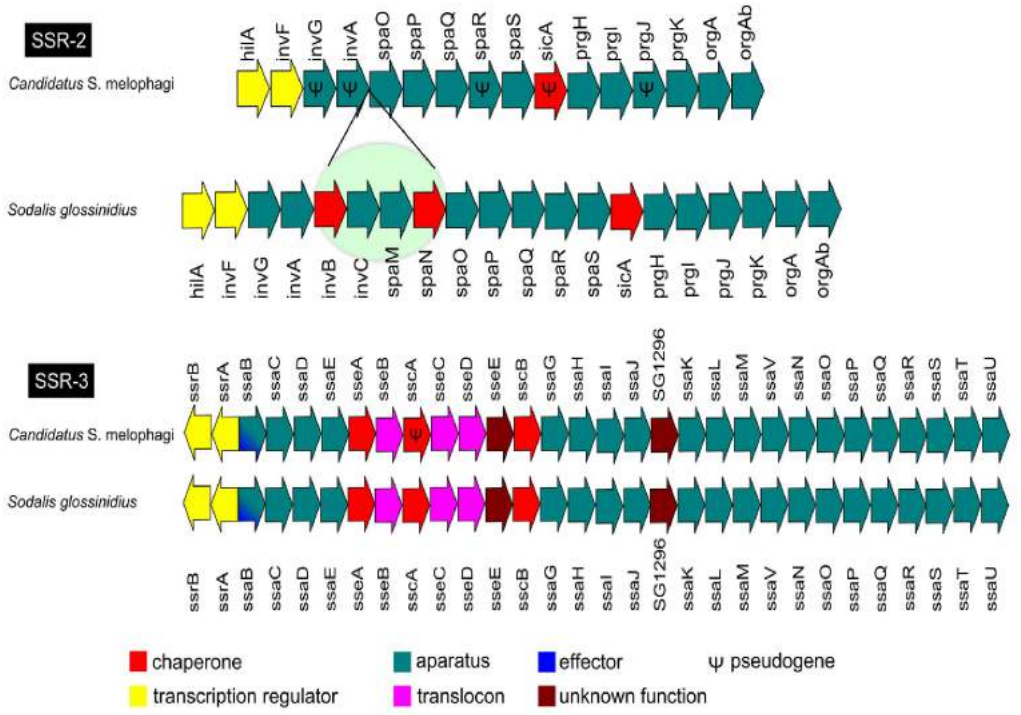


Figure 1. Composition of SSR-2 and SSR-3 copies of *Candidatus* Sodalis melophagi TTSS. Ψ – putative pseudogene (see Material and Methods).
doi:10.1371/journal.pone.0040354.g001

Description of *Candidatus* Sodalis Melophagi sp. nov

Candidatus Sodalis melophagi (me.lo.phag.i; *melophagi* of the genus *Melophagus*).

Candidatus Sodalis melophagi is a Gram-negative bacterium isolated from the sheep ked *Melophagus ovinus*. The bacteria were detected in heamolymph and bacteriome. The cells are rod-shaped non-motile and non-spore forming under laboratory conditions. The average length is 1.9 μm, however, bacteria ranging from 0.6 to 5 μm in length were observed. The bacteria grow intracelullarly and extracellularly in presence of *Aedes albopictus* cell line C6/36 under aerobic conditions. They can also be cultivated axenically both in liquid and solid media containing enzymatically digested proteins as a source of nitrogen, with the addition of blood under the microaerophilic atmosphere (5% O_2 balanced with carbon dioxide) at 27°C.

Discussion

Remarkably, the main biological traits of *Candidatus* Sodalis melophagi and its relation to the host resemble those of *Sodalis glossinidus* in tsetse flies. In both cases the hosts are exclusively blood feeding organisms reproducing by vivipary and their *Sodalis* plays a role of an "accessory" symbiont accompanying phylogenetically distant P-symbiont (*Wigglesworthia* in tsetse flies and *Arsenophonus* in *Melophagus*). This status of *Candidatus* Sodalis melophagi is strongly indicated by phylogenetic characteristics as well as location and morphology revealed by electron microscopy. Similar to *S. glossinidius*, the rod-shaped cells of *Candidatus* Sodalis melophagi are not restricted to the bacteriocytes but they infect various non-specialized cells and can be cultivated from haemolyph. It is therefore interesting to note that according to phylogenetic analyses its symbiosis with the host originated independently of the *S. glossinidius* as well as of the third hipoboscoid lineage, the symbiont of *Craterina melbae*. Although exact position of *Candidatus* Sodalis melophagi within the *Sodalis* clade varies with method and taxon sampling, it never clusters with any of the two other hippoboscoid-derived lineages. Within the *Sodalis* clade, such host-symbiont incongruence is not unique to the hippoboscoid-derived lineages. It is also expressed by the distribution of several other samples, particularly those from coleopteran and homopteran hosts (Figure 2, Figure S1) and clearly demonstrates the capability of *Sodalis* to spread by horizontal transfers. In contrast, the close relationships of the two new 16 S rDNA sequences obtained from *Ornithomya* with the *Sodalis* previously described from *Craterina melbae* suggests that these two hippoboscoid genera may share the same symbiotic lineage inherited from their common ancestor. However, the sampling of *Sodalis* lineages available from other insect groups is highly incomplete. In addition, the position of some sequences, such as *Sodalis* from *Sitophilus rugicollis* is affected by long-branch attraction due to the high AT content; instability of this lineage has been found in previous analyses [9]. Thus, although we extended the sampling with additional three sequences, any interpretations of the modes and mechanism of *Sodalis* transmission can under the current circumstances be only working hypotheses rather than serious conclusions based on the data.

While the draft genomic data of *Candidatus* Sodalis melophagi is currently under investigation, we preliminarily analyzed the

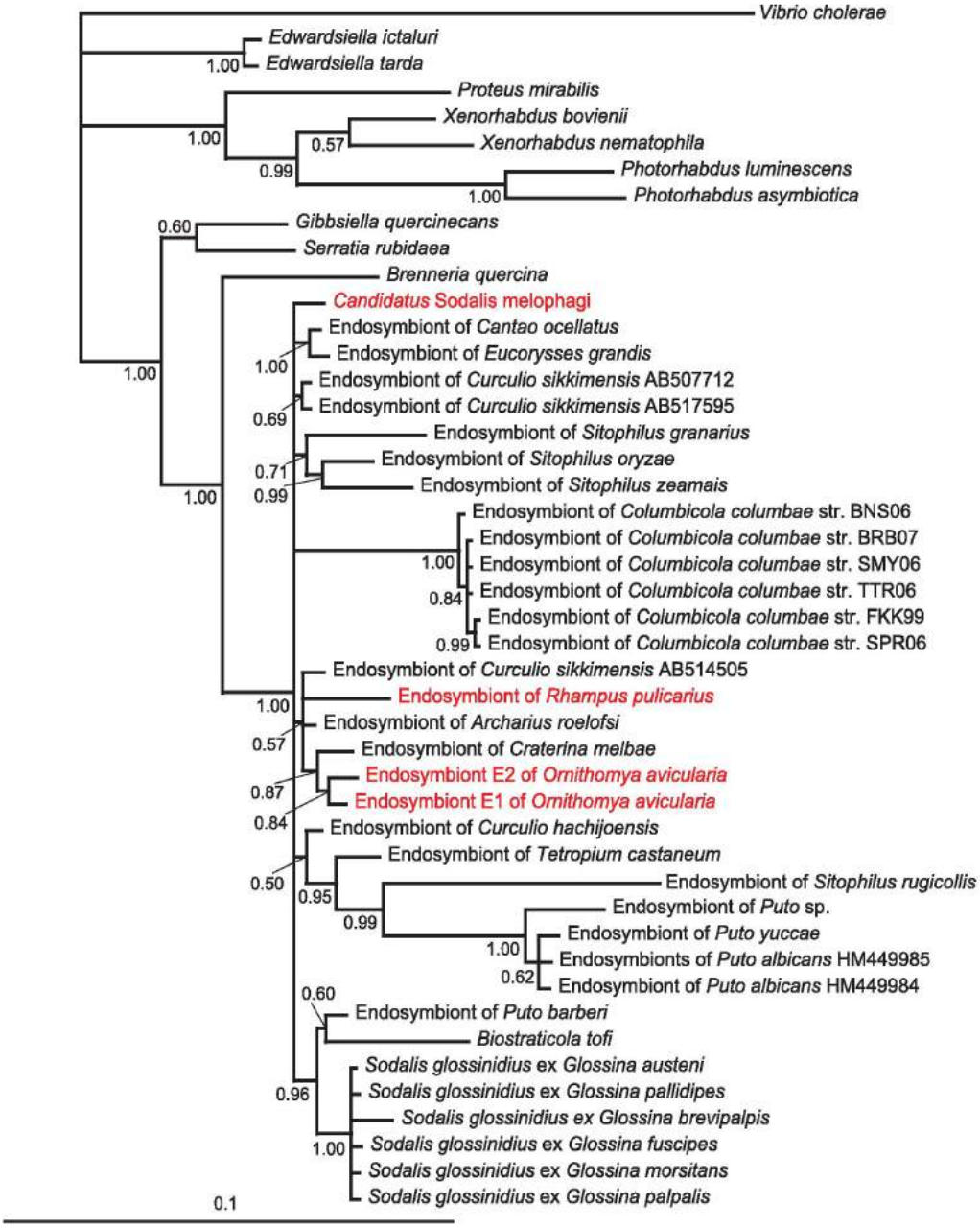


Figure 2. 16 S rDNA tree derived by BI analysis in MrBayes. Posterior probabilities are indicated by the numbers at the nodes. New *Sodalis* lineages added in this study are printed in red.
doi:10.1371/journal.pone.0040354.g002

composition and structure of the TTSS. Significance of this system in symbiosis evolution has previously been suggested [39] and investigated in several symbiotic lineages [21,40–43]. These genes are also among the few sequences that are currently available and can be compared between different *Sodalis* taxa. In *S. glossinidius*, experimental work indicated that TTSS of SPI1 type from *Salmonella* (later designated as SSR-2) is essential for entering the host cell. Presence and apparent functionality of this system was subsequently confirmed in another *Sodalis*, the primary endosymbiont of *Sitophilus zeamais* (SZPE) [42]. However, further studies

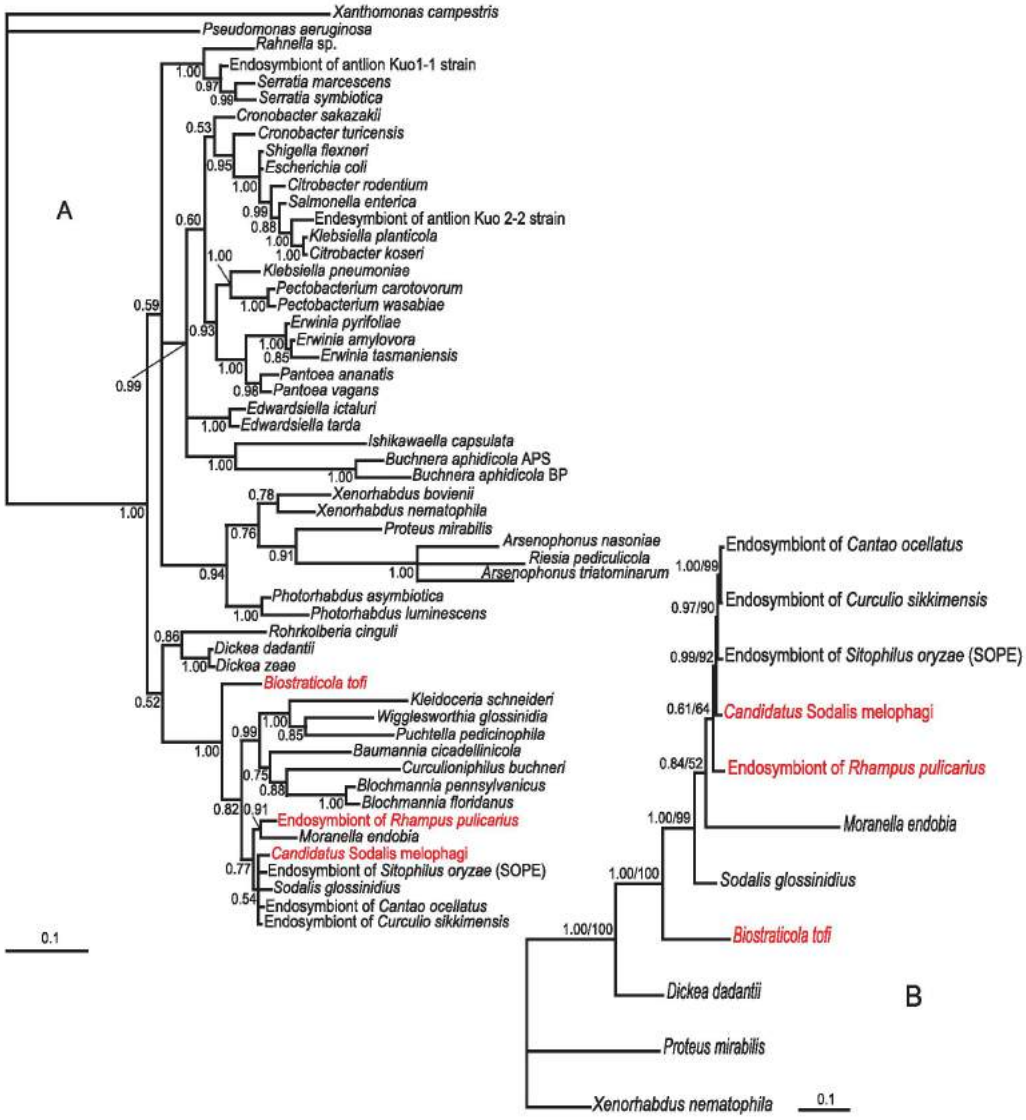


Figure 3. Phylogenetic trees derived from *groEL* matrices by BI in MrBayes. New *Sodalis* lineages added in this study are printed in red. **A**: The tree inferred from aminoacid matrix. Posterior probabilities are indicated by the numbers at the nodes. **B**: The tree inferred from nucleotide matrix restricted taxonomically to the *Sodalis* branch. The numbers at the nodes show the posterior probabilities and bootstrap values from the identical topology obtained by ML in PhyML.
doi:10.1371/journal.pone.0040354.g003

revealed another two copies of TTSS in *S. glossinidius* genome, SSR-1 closely related to the Ysa system of *Yersinia*, and SSR-3 related to *Salmonella* SPI2 [26,27,40]. The secretion systems in *Candidatus* Sodalis melophagi show clear similarity to *S. glossinidius* genes but the whole machinery is much less complex; SSR-1 is completely missing and SSR-2 is highly eroded. However, the form corresponding to the SSR-3 is complete and possibly functional. Its structure and gene content is highly similar to that in *S. glossinidius*, except for a pseudogenized state of *sscA*, encoding a putative chaperone of secreted protein SseC [44].

The differences between TTSS in *S. glossinidius* and *Candidatus* Sodalis melophagi pose an interesting question about the origin, role and significance of TTSS and its different copies in the genus *Sodalis*. As the two compared lineages, *Candidatus* Sodalis melophagi and *S. glossinidius* are not closely related and the whole *Sodalis* tree is currently undersampled, it is impossible to hypothesize whether SSR-1 was lost in the former one or acquired by the latter one after diversification of their ancestor. The significance of TTSS in the two lineages is even more difficult to assess. In *Candidatus* Sodalis melophagi, the SSR-3 alone or

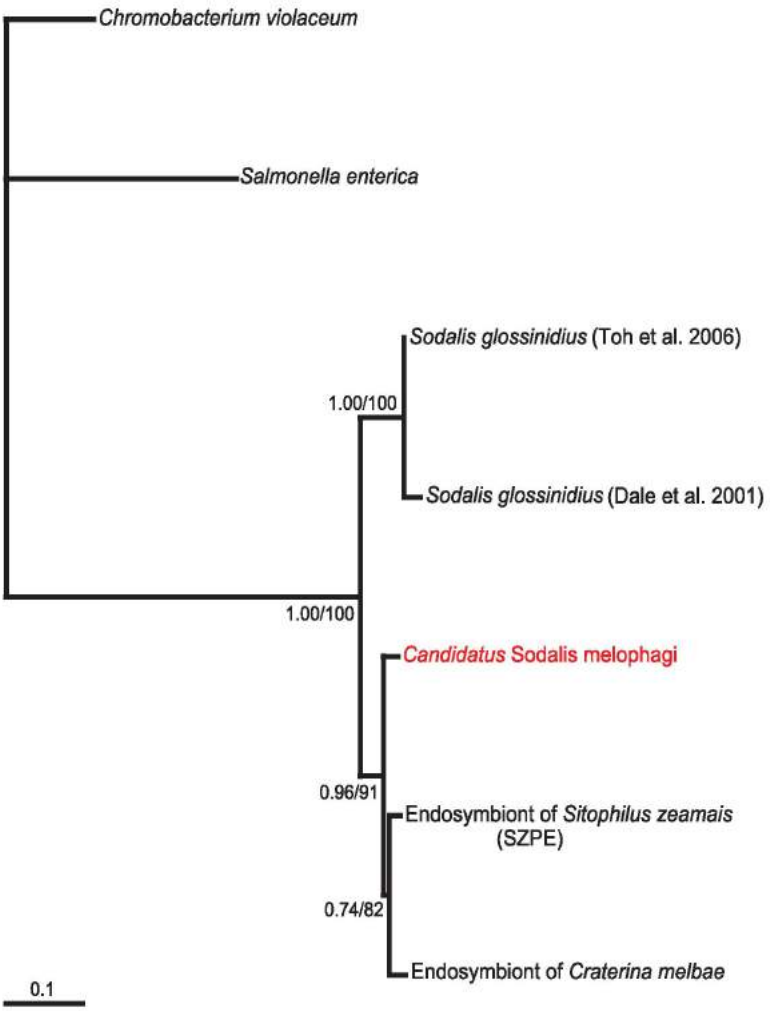


Figure 4. Phylogenetic tree derived from *spaPQR* region by BI in MrBayes. The numbers at the nodes show the posterior probabilites and bootstrap values from the identical topology obtained by ML in PhyML. New *Sodalis* lineages added in this study are printed in red.
doi:10.1371/journal.pone.0040354.g004

possibly in synergy with the retained functional genes of SSR-2, seems to be sufficient for maintenance of the intracellular lifestyle. This view, however, is based solely on the observed presence/absence of the genes and their comparison to the experimental results from *S. glossinidius* [21]. It should be taken into account that other related systems, such as flagellar export apparatus, can participate in the host cell invasion. Despite the knockout experiments, the situation is similarly unclear in *S. glossinidius* where the presence of SSR-3 has not been known at the time of the experimental work and its significance could not be investigated. From the genomic point of view, *S. glossinidius* has so far been the only member of the genus for which a detailed characterization of the genome and some metabolic capabilities is available. The quality of *Candidatus* Sodalis melophagi paired-end data retrieved from the bacteriome sequencing and additional mate pair data recently obtained from the pure culture suggest that a draft genome of this symbiont could be established and used for further comparisons. Such analysis, comparing reduction, structure, and possible adaptive changes of independent but closely related bacteria from two hosts with similar but unique biology will provide important insight into the symbiogenetic processes. In respect to the future work, it may help to discriminate between the random and symbiosis-associated modifications and indicate candidate genes for a more detailed investigation.

Materials and Methods

Ethic Statement

All field studies did not involve protected or endangered organisms. They were not performed on privately-owned or protected locations and were performed according to the law of the Czech Republic.

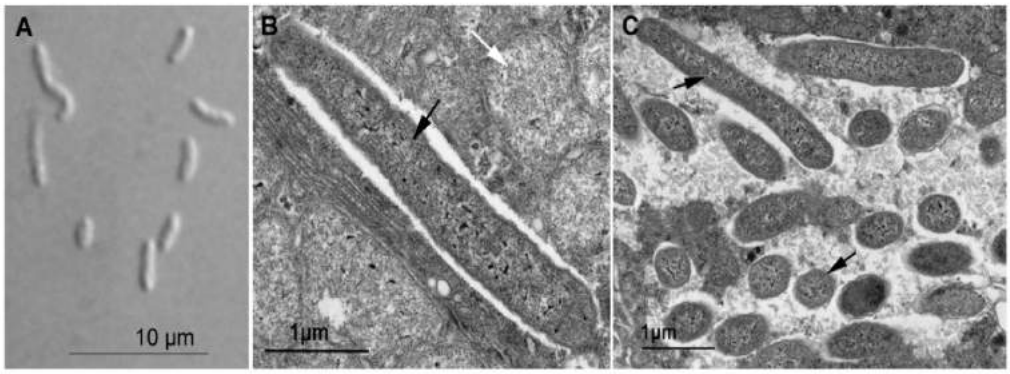


Figure 5. Morphology and ultrastructure of *Candidatus* Sodalis melophagi. A: *In vitro* cell culture in Nomarski contrast. **B**, **C**: Cells of *Candidatus* Sodalis melophagi in bacteriome. Black arrows – cells of *Candidatus* Sodalis melophagi, white arrows – cells of the primary endosymbiont of the genus *Arsenophonus*.
doi:10.1371/journal.pone.0040354.g005

Insect Samples, DNA Extraction, Sequencing and Data Assembly

Adults of *Melophagus ovinus* were obtained commercially from a licensed family sheep farm at Krásetín, Czech Republic. A midgut region with the bacteriome was dissected into the phosphate buffered saline (PBS) and total genomic DNA was extracted from each single adult by QIAamp DNA Micro Kit (Qiagen). DNA concentration was determined using NanoDrop2000 (Thermo Scientific) and its quality was assessed in 2% agarose gel using standard electrophoresis. PCR was carried out as described previously [18] with bacterial primers for 16 S rDNA and *groEL* (Table S3). The same procedure was followed for three additional samples. Two of them were *Sodalis* bacteria from other insect hosts, *Ornithomya avicularia* (Diptera, Hippoboscidae) and *Rhampus pulicarius* (Coleoptera, Curculionidae), and the third was *Biostraticola tofi* strain BF36^T, a free-living bacterium supposedly closely related to *Sodalis* [45]. Laboratory culture of *B. tofi* was obtained from the DSMZ microorganism collection (Germany), *R. pulicarius* was collected near České Budějovice (Czech Republic); according to the law of the Czech Republic, no permits are required for collection of this organism. Samples of *O. avicularia* were provided by Department of Zoology (University of South Bohemia); the collections were done during the ornithological research performed in accordance with the law of the Czech Republic.

We used 3 µg of genomic DNA isolated from *M. ovinus* bacteriome as a template for an Illumina paired-end library with an insert size of 300 bp. Library construction and sequencing on one lane in a 100 bp run was carried out at Keck Microarray Resource, West Haven CT, USA. Reads obtained underwent adaptor and quality trimming, and were further processed using two different approaches. First, the reads with significant BLASTN [46] hit to *Sodalis glossinidius* genome sequence (NC_007712) were filtered and the retrieved subset was assembled using CLC Genomic Workbench (CLC bio A/S) with parameters set to the following values: similarity 0.9, length fraction 0.9, costs for deletion/insertion/mismatch 3. Second, de-novo assembly of all the processed reads into contigs was done on CLC Genomic Workbench (CLC bio A/S) under the same parameter setting. Contigs from this assembly were binned based on their average coverage and BlastX [46] hits against available bacterial genomes. Since some of the sequences were used in this study for phylogenetic reconstruction, we checked the accuracy of the Illumina-derived sequences by an independent Sanger sequencing of the *groEL* gene.

Type Three Secretion System (TTSS) Annotation

Contigs spanning corresponding genes to TTSS islands from *Sodalis glossinidius* were retrieved based on the Blast X [46] results from both assemblies. In order to obtain a single contig for each of the TTSS islands, gaps were closed using targeted Sanger sequencing (Table S3). Mapped sequences were checked for presence of single nucleotide polymorphisms (SNPs). ORF prediction was done using CLC Genomic Workbench (CLC bio A/S) with the minimum length set to 30 AA. Gene annotation based on similarity confirmed by BLAST searches was performed manually in the same software. All genes that contained frame shift mutation or stop mutations were tentatively classified as pseudogenes. The *hilA* previously classified as pseudogene [26,27] was annotated as functional based on recent experimental work [47].

Phylogenetic Analyses

We used three different regions for the phylogenetic reconstruction, 16 S rDNA, *groEL* and the TTSS region consisting of the *spaP-spaQ-spaR* genes (Tables S4, S5, S6). The *groEL* amino acid dataset was aligned using ClustalW algorithm in BioEdit with default parameters [48] and all ambiguously aligned sites were removed from the further analyses. To gain more precise phylogenetic resolution within the *Sodalis* branch, we narrowed the taxon sampling and prepared an additional matrix using nucleotide sequences for the *groEL* gene. The matrix was aligned as described above and edited manually. The 16 S rDNA and the *spaPQR* dataset were aligned in the Mafft program [49], using the E-INS-i strategy with default parameters and manually edited in BioEdit [48]. For 16 S rDNA, the ambiguously aligned positions were eliminated in Gblocks [50].

Two approaches were used to infer phylogenetic trees, maximum likelihood (ML) and Bayesian inference (BI). For ML we used PhyML v3.0 [51,52] with the SPR search algorithm. BI was performed in MrBayes 3.1.2 [53,54] with five million generations and tree sampling every 100 generations. AWTY [55] was used to check the MCMC convergence and determine burn-in. Evolutionary substitution models for proteins and DNA were selected by ProtTest 2.4 [56] and for DNA by jModelTest 0.1.1 [57], respectively. For DNA sequences, the General Time

Reversible (GTR) model was used with an estimated proportion of invariable sites (I) and heterogeneity of evolutionary rates modeled by the eight substitution rate categories of the gamma (Γ) distribution and the gamma shape parameter (alpha) estimated from the data. LG+I+Γ was determined as the best fitting model for the amino acid *groEL* dataset, and it was used for the ML analyses. Since this model is not implemented in MrBayes, we replaced it with WAG+I+Γ for the BI analysis. JTT+I+Γ model was used for the *spaPQR* concatenate in both BI and ML analyses.

Cultivation of *Candidatus* Sodalis Melophagi sp. nov

The insects were surface sterilized using 96% ethanol and hemolymph was collected into a 1.5 ml eppendorf tube containing 500 µl of Mitsuhashi-Maramorosch (MM) medium [58] with 20% heat inactivated foetal bovine serum (FBS). The tube was incubated overnight at 27°C without shaking, a slightly modified protocol of Matthew et al. [25] was followed. Liquid culture was plated onto 10% sheep blood MMI plates solidified by 1% agar. The medium was supplemented with 100 µg/ml Polymyxin B and 10 mg/ml Amphotericin B to prevent contamination by Gram-negative non-symbiotic bacteria [59] and fungi. The plates were incubated at 27°C in a microaerobic atmosphere generated by the Campygen pack system (Oxoid) producing 5% of oxygen balanced with carbon dioxide. Bacteria were further inoculated into flasks containing C6/36 mosquito cells [60] (LGC Standards, Czech Republic) in MM medium with 20% heat inactivated FBS.

Microscopy

A three day old liquid culture of C6/36 cells infected with *Candidatus* Sodalis melophagi was used for microscopic examination. Cells were harvested and fixed in 4% formaldehyde in PBS and observed under the BX53 microscope (Olympus) using Nomarski contrast. For electron microscopy, the midgut region containing bacteriome was dissected directly into a 2.5% glutaraldehyde in 0.1 M phosphate buffer and prefixed at 4°C overnight. The tissue was then postfixed at 4°C for 60 min with 2% osmium tetroxide in phosphate buffer. After dehydration through ethanol series, the samples were embedded in Spurr resin. Ultrathin sections were stained with uranyl acetate and lead citrate and examined in transmission electron microscope JEOL JEM-1010.

References

1. Snyder AK, McMillen CM, Wallenhorst P, Rio RV (2011) The phylogeny of *Sodalis*-like symbionts as reconstructed using surface-encoding loci. FEMS Microbiol Lett 317: 143–151.
2. Heddi A, Charles H, Khatchadourian C, Bonnot G, Nardon P (1998) Molecular characterization of the principal symbiotic bacteria of the weevil *Sitophilus oryzae*: a peculiar G + C content of an endocytobiotic DNA. J Mol Evol 47: 52–61.
3. Dale C, Maudlin I (1999) *Sodalis* gen. nov. and *Sodalis glossinidius* sp. nov., a microaerophilic secondary endosymbiont of the tsetse fly *Glossina morsitans morsitans*. Int J Syst Bacteriol 49 Pt 1: 267–275.
4. Toju H, Hosokawa T, Koga R, Nikoh N, Meng XY, et al. (2010) "*Candidatus* Curculioniphilus buchneri," a novel clade of bacterial endocellular symbionts from weevils of the genus *Curculio*. Appl Environ Microbiol 76: 275–282.
5. Toju H, Fukatsu T (2011) Diversity and infection prevalence of endosymbionts in natural populations of the chestnut weevil: relevance of local climate and host plants. Mol Ecol 20: 853–868.
6. Conord C, Despres L, Vallier A, Balmand S, Miquel C, et al. (2008) Long-term evolutionary stability of bacterial endosymbiosis in curculionoidea: additional evidence of symbiont replacement in the dryophthoridae family. Mol Biol Evol 25: 859–868.
7. Campbell BC, Bragg TS, Turner CE (1992) Phylogeny of symbiotic bacteria of four weevilspecies (coleoptera: curculionidae) based on analysis of 16 S ribosomal DNA. Insect Biochem Mol Biol 22: 415–421.
8. Fukatsu T, Koga R, Smith WA, Tanaka K, Nikoh N, et al. (2007) Bacterial endosymbiont of the slender pigeon louse, *Columbicola columbae*, allied to endosymbionts of grain weevils and tsetse flies. Appl Environ Microbiol 73: 6660–6668.
9. Nováková E, Hypša V (2007) A new *Sodalis* lineage from bloodsucking fly *Craterina melbae* (Diptera, Hippoboscoidea) originated independently of the tsetse flies symbiont *Sodalis glossinidius*. FEMS Microbiol Lett 269: 131–135.
10. Sameshima S, Hasegawa E, Kitade O, Minaka N, Matsumoto T (1999) Phylogenetic comparison of endosymbionts with their host ants based on molecular evidence. Zoolog Sci 16: 993–1000.
11. Gruwell ME, Hardy NB, Gullan PJ, Dittmar K (2010) Evolutionary relationships among primary endosymbionts of the mealybug subfamily phenacoccinae (hemiptera: Coccoidea: Pseudococcidae). Appl Environ Microbiol 76: 7521–7525.
12. Burke GR, Normark BB, Favret C, Moran NA (2009) Evolution and diversity of facultative symbionts from the aphid subfamily Lachninae. Appl Environ Microbiol 75: 5328–5335.
13. Kaiwa N, Hosokawa T, Kikuchi Y, Nikoh N, Meng XY, et al. (2010) Primary gut symbiont and secondary, *Sodalis*-allied symbiont of the Scutellerid stinkbug *Cantao ocellatus*. Appl Environ Microbiol 76: 3486–3494.
14. Kaiwa N, Hosokawa T, Kikuchi Y, Nikoh N, Meng XY, et al. (2011) Bacterial symbionts of the giant jewel stinkbug *Eucorysses grandis* (Hemiptera: Scutelleridae). Zoolog Sci 28: 169–174.
15. Grünwald S, Pilhofer M, Höll W (2010) Microbial associations in gut systems of wood- and bark-inhabiting longhorned beetles (Coleoptera: Cerambycidae). Syst Appl Microbiol 33: 25–34.

Supporting Information

Figure S1 16 S rDNA tree derived by ML method in PhyML. Bootstrap values are indicated by the numbers at the nodes. New *Sodalis* lineages added in this study are printed in red. Values under 50 are not shown.
(EPS)

Figure S2 Phylogenetic trees derived from *groEL* aminoacid matrix under ML in PhyML. Bootstrap values are indicated by the numbers at the nodes. New *Sodalis* lineages added in this study are printed in red. Values under 50 are not shown.
(EPS)

Table S1 List of sequences acquired in this study.
(DOC)

Table S2 Characteristics of particular datasets.
(DOC)

Table S3 List of primers used in this study.
(DOC)

Table S4 List of 16 S rDNA sequences used for phylogenetic inference.
(DOC)

Table S5 List of *groEL* sequences used for phylogenetic inference.
(DOC)

Table S6 List of *spaPQR* sequences used for phylogenetic inference.
(DOC)

Acknowledgments

We are grateful to Nancy A. Moran and Keck Microarray Resource for technical assistance with the genome sequencing.

Author Contributions

Wrote the paper: TC FH EN VH. Designed the study: EN VH. Established in vitro culture: TC FH. Generated and assembled the genomic data: EN. Performed the light and electron microscopy characterization: TC FH VH. Performed phylogenetic analyses: TC FH. Annotated and analyzed the TTSS sequences: TC.

16. Thao ML, Clark MA, Baumann L, Brennan EB, Moran NA, et al. (2000) Secondary endosymbionts of psyllids have been acquired multiple times. Curr Microbiol 41: 300–304.
17. Thao ML, Gullan PJ, Baumann P (2002) Secondary (γ-Proteobacteria) endosymbionts infect the primary (β-Proteobacteria) endosymbionts of mealybugs multiple times and coevolve with their hosts. Appl Environ Microbiol 68: 3190–3197.
18. Nováková E, Hypša V, Moran NA (2009) *Arsenophonus*, an emerging clade of intracellular symbionts with a broad host distribution. BMC Microbiol 9: 143.
19. Dale C, Welburn SC (2001) The endosymbionts of tsetse flies: manipulating host-parasite interactions. Int J Parasitol 31: 628–631.
20. Pontes MH, Dale C (2011) Lambda red-mediated genetic modification of the insect endosymbiont *Sodalis glossinidius*. Appl Environ Microbiol 77: 1918–1920.
21. Dale C, Young SA, Haydon DT, Welburn SC (2001) The insect endosymbiont *Sodalis glossinidius* utilizes a type III secretion system for cell invasion. Proc Natl Acad Sci U S A 98: 1883–1888.
22. Cheng Q, Aksoy S (1999) Tissue tropism, transmission and expression of foreign genes in vivo in midgut symbionts of tsetse flies. Insect Mol Biol 8: 125–132.
23. Runyen-Janecky LJ, Brown AN, Ott B, Tujuba HG, Rio RV (2010) Regulation of high-affinity iron acquisition homologues in the tsetse fly symbiont *Sodalis glossinidius*. J Bacteriol 192: 3780–3787.
24. Pontes MH, Babst M, Lochhead R, Oakeson K, Smith K, et al. (2008) Quorum sensing primes the oxidative stress response in the insect endosymbiont, *Sodalis glossinidius*. PLoS One 3: e3541.
25. Matthew CZ, Darby AC, Young SA, Hume LH, Welburn SC (2005) The rapid isolation and growth dynamics of the tsetse symbiont *Sodalis glossinidius*. FEMS Microbiol Lett 248: 69–74.
26. Toh H, Weiss BL, Perkin SA, Yamashita A, Oshima K, et al. (2006) Massive genome erosion and functional adaptations provide insights into the symbiotic lifestyle of *Sodalis glossinidius* in the tsetse host. Genome Res 16: 149–156.
27. Belda E, Moya A, Bentley S, Silva FJ (2010) Mobile genetic element proliferation and gene inactivation impact over the genome structure and metabolic capabilities of *Sodalis glossinidius*, the secondary endosymbiont of tsetse flies. BMC Genomics 11: 449.
28. Aksoy S (1995) *Wigglesworthia* gen. nov. and *Wigglesworthia glossinidia* sp. nov., taxa consisting of the mycetocyte-associated, primary endosymbionts of tsetse flies. Int J Syst Bacteriol 45: 848–851.
29. Cheng Q, Ruel TD, Zhou W, Moloo SK, Majiwa P, et al. (2000) Tissue distribution and prevalence of *Wolbachia* infections in tsetse flies, *Glossina* spp. Med Vet Entomol 14: 44–55.
30. O'Neill SL, Gooding RH, Aksoy S (1993) Phylogenetically distant symbiotic microorganisms reside in *Glossina* midgut and ovary tissues. Med Vet Entomol 7: 377–383.
31. Lindh JM, Lehane MJ (2011) The tsetse fly *Glossina fuscipes fuscipes* (Diptera: Glossina) harbours a surprising diversity of bacteria other than symbionts. Antonie Van Leeuwenhoek 99: 711–720.
32. Wu D, Daugherty SC, Van Aken SE, Pai GH, Watkins KL, et al. (2006) Metabolic complementarity and genomics of the dual bacterial symbiosis of sharpshooters. PLoS Biol 4: e188.
33. McCutcheon JP, Moran NA (2007) Parallel genomic evolution and metabolic interdependence in an ancient symbiosis. Proc Natl Acad Sci U S A 104: 19392–19397.
34. Akman L, Yamashita A, Watanabe H, Oshima K, Shiba T, et al. (2002) Genome sequence of the endocellular obligate symbiont of tsetse flies, *Wigglesworthia glossinidia*. Nat Genet 32: 402–407.
35. Snyder AK, Deberry JW, Runyen-Janecky L, Rio RV (2010) Nutrient provisioning facilitates homeostasis between tsetse fly (Diptera: Glossinidae) symbionts. Proc Biol Sci 277: 2389–2397.
36. Rio RV, Symula RE, Wang J, Lohs C, Wu YN, et al. (2012) Insight into the transmission biology and species-specific functional capabilities of tsetse (Diptera: glossinidae) obligate symbiont *Wigglesworthia*. MBio 3.
37. Buchner P (1965) Endosymbiosis of animals with plant microorganisms. New York: Interscience Publishers.
38. Krieg NR, Hoffman PS (1986) Microaerophily and oxygen toxicity. Annu Rev Microbiol 40: 107–130.
39. Coombes BK (2009) Type III secretion systems in symbiotic adaptation of pathogenic and non-pathogenic bacteria. Trends Microbiol 17: 89–94.
40. Dale C, Jones T, Pontes M (2005) Degenerative evolution and functional diversification of type-III secretion systems in the insect endosymbiont *Sodalis glossinidius*. Mol Biol Evol 22: 758–766.
41. Hueck CJ (1998) Type III protein secretion systems in bacterial pathogens of animals and plants. Microbiol Mol Biol Rev 62: 379–433.
42. Dale C, Plague GR, Wang B, Ochman H, Moran NA (2002) Type III secretion systems and the evolution of mutualistic endosymbiosis. Proc Natl Acad Sci U S A 99: 12397–12402.
43. Moran NA, Degnan PH, Santos SR, Dunbar HE, Ochman H (2005) The players in a mutualistic symbiosis: insects, bacteria, viruses, and virulence genes. Proc Natl Acad Sci U S A 102: 16919–16926.
44. Cirillo DM, Valdivia RH, Monack DM, Falkow S (1998) Macrophage-dependent induction of the *Salmonella* pathogenicity island 2 type III secretion system and its role in intracellular survival. Mol Microbiol 30: 175–188.
45. Verbarg S, Fruhling A, Cousin S, Brambilla E, Gronow S, et al. (2008) *Biostraticola tofi* gen. nov., spec. nov., a novel member of the family Enterobacteriaceae. Curr Microbiol 56: 603–608.
46. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215: 403–410.
47. Pontes MH, Smith KL, De Vooght L, Van Den Abbeele J, Dale C (2011) Attenuation of the sensing capabilities of PhoQ in transition to obligate insect-bacterial association. PLoS Genet 7: e1002349.
48. Hall T (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser: 95–98.
49. Katoh K, Kuma K, Toh H, Miyata T (2005) MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res 33: 511–518.
50. Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol 17: 540–552.
51. Guindon S, Delsuc F, Dufayard JF, Gascuel O (2009) Estimating maximum likelihood phylogenies with PhyML. Methods Mol Biol 537: 113–137.
52. Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52: 696–704.
53. Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17: 754–755.
54. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572–1574.
55. Nylander JA, Wilgenbusch JC, Warren DL, Swofford DL (2008) AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. Bioinformatics 24: 581–583.
56. Abascal F, Zardoya R, Posada D (2005) ProtTest: selection of best-fit models of protein evolution. Bioinformatics 21: 2104–2105.
57. Posada D (2008) jModelTest: phylogenetic model averaging. Mol Biol Evol 25: 1253–1256.
58. Mitsuhashi JM, Maramorosch K (1964) Leafhopper tissue culture: embryonic, nymphal and imaginal tissue from aseptic insects. Contribution from Boyce Thompson Institute: 435–460.
59. Dale C, Beeton M, Harbison C, Jones T, Pontes M (2006) Isolation, pure culture, and characterization of "*Candidatus* Arsenophonus arthropodicus," an intracellular secondary endosymbiont from the hippoboscid louse fly *Pseudolynchia canariensis*. Appl Environ Microbiol 72: 2997–3004.
60. Igarashi A (1978) Isolation of a Singh's *Aedes albopictus* cell clone sensitive to Dengue and Chikungunya viruses. J Gen Virol 40: 531–544.

4.3 Chapter 3

Víchová B., Majláthová V., Nováková M., Stanko M., Hviščová I., Pangrácová L., Chrudimský T., Čurlík J., Peťko B., 2013. Anaplasma infections in ticks and reservoir hosts from Slovakia, Central Europe. Infection, Genetics and Evolution.

Abstract

Anaplasma phagocytophilum is a worldwide distributed bacterium with a significant medical and veterinary importance. It grows within the phagosome of infected neutrophils and is responsible for human granulocytic anaplasmosis (HGA), tick-borne fever (TBF) of small ruminants and cattle, canine and equine granulocytic anaplasmosis, but infects also a great variety of wildlife species. Wild ungulates and rodents are considered reservoirs of infection in natural foci. The objective of this study was to determine the spectrum of animal species involved in the circulation of A. phagocytophilum in Slovakia and to analyze the variability of obtained nucleotide sequences, in order to determine whether genotypes from Slovakia cluster according to host-species or geographical location. Several animal species and vector ticks were screened for the presence of members of the family Anaplasmataceae using PCR based methods. Additional data on the molecular evidence of Anaplasma ovis and Candidatus Neoehrlichia mikurensis are presented. These pathogens were detected in tested sheep flocks and rodents with the mean infection rates of 8.16% and 10.75%, respectively. A. phagocytophilum was genotyped by 16S rRNA and groEL gene sequencing. Bacterial DNA was confirmed in questing ixodid ticks, in domesticated canine, wild rodents and several species of wild ungulates. In European isolates, 16S rRNA gene does not seem to be an appropriate locus for the analyses of heterogeneity as it is too conservative. Similarly, 16S rRNA isolates from our study did not reveal any polymorphisms. All isolates were identical in overlapped region and showed identity with sequences from ticks, horses or ruminants previously isolated elsewhere in the world. On the other hand, the groESL heat shock operon is widely used for determination of diversity and the analyses have already revealed considerable degree of heterogeneity. Tested ungulates were infected with A. phagocytophilum to a considerable extent. High proportions of red and roe deer tested positive and the rates of infection reached over 60.0%. GroEL sequences from canine, wild ungulates and ticks from Slovakia clustered within a clade

together with isolates from horses, humans, wild ungulates and ticks from Slovakia or elsewhere in the world. Sequences from rodents clustered apart from those obtained from wild ungulates, ticks and humans. These results suggest that European rodents do not harbour A. phagocytophilum strains with strong zoonotic potential such as those from United States.

ARTICLE IN PRESS

Infection, Genetics and Evolution xxx (2013) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Anaplasma infections in ticks and reservoir host from Slovakia

Bronislava Víchová [a,*], Viktória Majláthová [a], Mária Nováková [a], Michal Stanko [a,b], Ivana Hviščová [a], Lucia Pangrácová [a], Tomáš Chrudimský [c], Ján Čurlík [d], Branislav Peťko [a]

[a] Institute of Parasitology, Slovak Academy of Sciences, Hlinkova 3, 04001 Košice, Slovakia
[b] Institute of Zoology, Slovak Academy of Sciences, Lofflerova 10, 04001 Košice, Slovakia
[c] Faculty of Science, University of South Bohemia, Branišovská 31, 37005 České Budějovice, Czech Republic
[d] Institute for Breeding and Diseases of Animals and Fishes, University of Veterinary Medicine and Pharmacy in Košice, Komenského 73, 04181 Košice, Slovakia

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Anaplasma
Ixodes ricinus
Rodents
Wild ungulates
Slovakia

ABSTRACT

Anaplasma phagocytophilum is a worldwide distributed bacterium with a significant medical and veterinary importance. It grows within the phagosome of infected neutrophils and is responsible for human granulocytic anaplasmosis (HGA), tick-borne fever (TBF) of small ruminants and cattle, canine and equine granulocytic anaplasmosis, but infects also a great variety of wildlife species. Wild ungulates and rodents are considered reservoirs of infection in natural foci. The objective of this study was to determine the spectrum of animal species involved in the circulation of *A. phagocytophilum* in Slovakia and to analyze the variability of obtained nucleotide sequences, in order to determine whether genotypes from Slovakia cluster according to host-species or geographical location.

Several animal species and vector ticks were screened for the presence of members of the family Anaplasmataceae using PCR based methods. Additional data on the molecular evidence of *Anaplasma ovis* and *Candidatus* Neoehrlichia mikurensis are presented. These pathogens were detected in tested sheep flocks and rodents with the mean infection rates of 8.16% and 10.75%, respectively. *A. phagocytophilum* was genotyped by 16S rRNA and *groEL* gene sequencing. Bacterial DNA was confirmed in questing ixodid ticks, in domesticated canine, wild rodents and several species of wild ungulates.

In European isolates, 16S rRNA gene does not seem to be an appropriate locus for the analyses of heterogeneity as it is too conservative. Similarly, 16S rRNA isolates from our study did not reveal any polymorphisms. All isolates were identical in overlapped region and showed identity with sequences from ticks, horses or ruminants previously isolated elsewhere in the world. On the other hand, the groESL heat shock operon is widely used for determination of diversity and the analyses have already revealed considerable degree of heterogeneity.

Tested ungulates were infected with *A. phagocytophilum* to a considerable extent. High proportions of red and roe deer tested positive and the rates of infection reached over 60.0%. GroEL sequences from canine, wild ungulates and ticks from Slovakia clustered within a clade together with isolates from horses, humans, wild ungulates and ticks from Slovakia or elsewhere in the world. Sequences from rodents clustered apart from those obtained from wild ungulates, ticks and humans. These results suggest that European rodents do not harbour *A. phagocytophilum* strains with strong zoonotic potential such as those from United States.

1. Introduction

Anaplasma phagocytophilum is a Gram-negative intracellular bacterium causing febrile disease of humans (human granulocytic anaplasmosis – HGA) and animals (pasture fever, equine and canine granulocytic anaplasmosis) (Dumler et al., 2001). The principle vectors of this rickettsial pathogen are ticks from the *Ixodes ricinus* complex. Transovarial transmission in ixodid ticks has not yet been confirmed; therefore the vertebrate hosts are crucial for the maintenance and circulation of pathogen in enzootic foci. Bacteria multiply in a broad range of hosts, especially in small rodents and wild ruminants which are discussed main reservoirs (Alberdi et al., 2000; Liz et al., 2002; Petrovec et al., 2003; Hulínska et al., 2004; Polin et al., 2004; Smetanová et al., 2006; de la Fuente et al., 2008; Bown et al., 2008; Štefanidesová et al. 2008). Based on analyses of several genetic markers (16S rRNA, *groESL*, *ankA*, *msp*), the existence of intraspecific heterogeneity has been recorded within *A. phagocytophilum*. Complex of closely related strains shows differences in vector and host preference,

* Corresponding author. Tel.: +421 908 698 482.
E-mail address: vichova@saske.sk (B. Víchová).

1567-1348/$ - see front matter © 2013 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.meegid.2013.06.003

Please cite this article in press as: Víchová, B., et al. *Anaplasma* infections in ticks and reservoir host from Slovakia. Infect. Genet. Evol. (2013), http://dx.doi.org/10.1016/j.meegid.2013.06.003

geographical distribution, pathogenicity and severity of clinical manifestations (Massung et al., 2006; Stuen et al., 2002; Stuen, 2007; Carpi et al., 2009). Significant differences were detected in the epidemiology of granulocytic anaplasmosis (GA) in Europe and North America. Whereas in the USA GA belongs among the most common tick-borne diseases, it remains relatively rare in Europe (Strle, 2004).

In the United States, the circulation of two distinct variants of *A. phagocytophilum* was confirmed on the basis of revealed differences in 16S rRNA (Massung et al., 2002). Similarly, de la Fuente et al. (2005c) characterized two European monophyletic groups based on analysis of *msp4* gene. One group consisted of strains from humans, dogs and horses and second one comprised strains from wild ruminants. They did not find *A. phagocytophilum* in tested rodents. In Europe, the situation seems to be rather different compared to the United States and the ecology and circulation of distinct ecotypes are poorly understood. Analyses of 16S rRNA did not confirm the unambiguous association of two genotypes with rodents and ruminants, as it has been reported in the USA. In European ungulates both variants have been recorded. Mostly those not associated with human infections, but also strains closely related to human granulocytic anaplasmosis (HGA) derived from red deer in Slovenia (Petrovec et al., 2002). Studies from England outlined the possibility of co-existence of two distinct subpopulations of *A. phagocytophilum* circulating in separate enzootic cycles, one involving deer and *I. ricinus* ticks and the other involving field voles (*Microtus agrestis*) and nidicolous *Ixodes trianguliceps* ticks (Bown et al., 2008).

The objective of this study was to identify epizootiological situation of granulocytic anaplasmosis in Slovakia, to determine the spectrum of animal species involved in the circulation of pathogen and to analyze and characterize the variability of obtained *A. phagocytophilum* DNA sequences, in order to determine whether genotypes from Slovakia cluster according to host-species or geographical location. We also report supplementary data on the presence of *Anaplasma ovis*, an etiologic agent of pasture fever of small ruminants and the recently described tick-borne bacterium *Candidatus* Neoehrlichia mikurensis, with the potential to cause severe diseases in immunocompromised patients. *Ca.* Neoehrlichia mikurensis was discovered in wild rodents and *Ixodes ovatus* ticks in Japan (Kawahara et al., 2004). Since then, it has been identified in ixodid ticks and several rodent species, which may act as reservoir hosts (Schouls et al., 1999; Špitalská et al., 2008; Alekseev et al., 2001; Andersson and Raberg, 2011; Jahfari et al., 2012; Pangrácová et al., 2012; Shpynov, 2012; Vayssier-Taussat et al., 2012), and in blood samples of febrile patients from Sweden, Germany, Switzerland and Czech Republic with a lethal course in one case (Fehr et al., 2010; von Loewenich et al., 2010; Welinder-Olsson et al., 2010).

2. Material and methods

2.1. Collected samples

2.1.1. I. ricinus ticks

Ticks were collected at the sampling sites continually during 2006–2009. A total of 1075 questing *I. ricinus* ticks were collected by white cloth flag dragging in locations of eastern Slovakia, in a hornbeam deciduous suburban forest park in Košice city (48°43′N, 21°15′E) ($n = 213$), near the Hornád river in Košice city, in the area devastated last year by floods (48°40′N, 21°18′E) ($n = 46$), in the forest habitat of the locality Kavečany in the district of Košice city (48°46′N, 21°12′E) ($n = 108$), in the district of Michalovce city (48°45′N, 21°55′E) ($n = 73$), in the Slovak Karst National Park (48°36′N, 20°52′E) in south-eastern Slovakia ($n = 242$) and 393 ticks were collected at the recreational grounds of Teplý vrch and nearby the water basin Kurinec in the district of Rimavská Sobota city (48°28′N, 20°05′E; 48°20′N, 20°01′E). Collected ticks were preserved in 70% ethanol until DNA isolation.

2.1.2. Dog blood samples

All tested blood and tissue samples of domesticated and wild animals included in study were collected continually from 2006 to 2011.

In total, 137 blood samples of dogs suspected of having non-specific febrile disease with a tick bite history were collected by vet practitioners in the districts of Senica (48°40′N, 17°21′E), Bratislava (48°08′N, 17°06′E), Trenčín (48°53′N, 18°02′E), Liptovský Mikuláš (49°04′N, 19°37′E), Lučenec (48°19′N, 19°39′E), Rimavská Sobota (48°23′N, 20°01′E), and Košice city (48°43′N, 21°15′E). Additionally, 144 blood samples of police and military dogs, previously screened at the Institute of Parasitology for the presence of *Dirofilaria* sp. (Miterpáková et al., 2010), were included in study. Samples were collected in plastic tubes containing anticoagulant (EDTA) and stored at 4 °C until DNA isolation.

2.1.3. Wild carnivores, ungulates and rodents

Samples of tissue were collected from 248 red foxes (*Vulpes vulpes*). Major part of them originated from animals shot by hunters in eastern Slovakia subjected to the post mortem examination of the presence of *Trichinella* spp. and *Echinococcus* spp. (Dubinský et al., 2006; Hurníková and Dubinský, 2009). Samples of blood or tissue (liver, spleen or muscle) were collected from 84 hunter-killed wild boars (*Sus scrofa*), 103 red deer (*Cervus elaphus*), 13 roe deer (*Capreolus capreolus*), and 3 fallow deer (*Dama dama*) from several sites within Slovakia. Moreover, tissue samples of 57 alpine chamois (*Rupicapra r. rupicapra*) from the Slovak Paradise National Park (48°54′N, 20°20′E) were included in study. Altogether 286 samples of blood or tissue (spleen or ear) from rodents were analyzed. Rodents were trapped to live traps in the districts and surroundings of the cities Ružomberok (49°04′N, 19°18′E) ($n = 44$), Rozhanovce (48°45′N, 21°20′E) ($n = 93$), Šebastovce (48°39′N, 21°16′E) ($n = 58$) and at four sampling sites situated within the district of Lučenec city (48°19′N, 19°39′E) ($n = 91$). Trapped individuals belonged to five rodent species: *Apodemus agrarius* 77.27%, *Apodemus flavicollis* 5.24%, *Apodemus uralensis (previously Apodemus microps)* 1.74% *Myodes glareolus* (8.06%) and *Microtus arvalis* (7.69%). Blood samples were stored in plastic tubes containing anticoagulant (EDTA) at 4 °C. Similarly, all obtained tissue samples were preserved in 70% ethanol or frozen at −20 °C in plastic tubes until DNA extraction. Additionally, all samples from rodents were tested for the presence of *Ca.* Neoehrlichia mikurensis.

2.1.4. Ruminants

A total of 178 blood specimens from cattle, sampled at three farms situated within the districts of the cities Prešov (49°00′N, 21°14′E) and Košice (48°43′N, 21°15′E) were screened for the presence of *A. phagocytophilum*. Moreover, 16 blood samples were obtained from a slaughter-house in eastern Slovakia. Except of *A. phagocytophilum*, all blood samples were tested for the presence of *Anaplasma marginale*, an agent responsible for bovine anaplasmosis.

Altogether 147 blood samples from four sheep flocks in central Slovakia, Očová (48°36′N, 19°17′E), Mýto pod Ďumbierom (48°51′N, 19°37′E), Liptovský Mikuláš (49°04′N, 19°36′E) and Veterná Poruba (49°06′N, 19°40′E) were delivered to our lab in plastic tubes with anticoagulant (EDTA). Additionaly, these samples were analysed for the presence of *A. ovis*, an agent responsible for pasture fever of small ruminants. Samples were stored at 4 °C until DNA isolation.

Please cite this article in press as: Víchová, B., et al. *Anaplasma* infections in ticks and reservoir host from Slovakia. Infect. Genet. Evol. (2013), http://dx.doi.org/10.1016/j.meegid.2013.06.003

2.2. DNA extraction

Ticks were identified to species and sex according to morphological key (Siuda, 1993), and then dried on air under the white-light for approximately 15 min to evaporate the residual ethanol. Genomic DNA was then isolated by using of alkaline hydrolysis method (Guy and Stanek, 1991). Samples were stored at −20 °C until further processing. Total DNA from tissue and blood samples was extracted using NucleoSpin® Tissue and/or Blood Kit (Macherey-Nagel, Germany). Before DNA extraction all tissue samples were digested with proteinase K at 56 °C, overnight. All DNA samples were eluated in 50 µl of elution buffer and stored at −20 °C for further analyses. In order to minimize DNA contamination, DNA extractions, PCR-mix preparations and PCR reactions were conducted in separate rooms.

2.3. PCR

In order to verify whether the genomic DNA from each *I. ricinus* tick had been isolated successfully, a fragment of the mitochondrial *cytochrome b* gene with a length of 620 bp was amplified (Black and Roehrdanz, 1998). In case of genomic DNA from blood and tissue samples, PCR amplification of 145-bp long orthologous fragment of the vertebrate mitochondrial 12S rRNA gene was performed as published previously by Humair et al. (2007). Negative samples were excluded from the file.

Partial PCR amplifications of 16S rRNA and *groESL* genes were used for detection of *A. phagocytophilum* and subsequent sequencing. For initial screening of all ticks, tissue and blood samples, nested PCR amplification of 16S rRNA gene was realized. Primers ge3a, ge10r were used for detection of all members of the family Anaplasmataceae, in the first round of amplification. One µl of PCR product from the first round was used as the template for second, *A. phagocytophilum*-specific round with the set of primers ge9f and ge2 (Massung et al., 1998). Randomly selected 16S rRNA products were additionally tested with a nested PCR targeting more variable *groESL* gene. In the first round, pair of primers HS1 and HS6 was used. These primers amplify the end of *groES* gene, intergenic spacer and approximately 2/3 of the *groEL* gene. In the second round HSVF and HSVR primers that span a 395-bp region of the *groEL* gene were used (Lotric-Furlan et al., 1998; Sumner et al., 2003).

Primers and cycling conditions for *msp4* PCR detection of *A. marginale* and *A. ovis* were used as published previously by de la Fuente et al. (2005a,b). Detection of 16S rRNA gene of *Ca.* Neoehrlichia mikurensis was performed according to Kawahara et al. (2004). All PCR amplifications were performed in a total of 25 µl of reaction mixture of 2× DyNAzyme II Master Mix (Finnzymes, Espoo, Finland) containing 7.6 µl of deionized sterile water, 12.5 µl of each primer and 2.5 µl of DNA template. As the negative control, nuclease free water was added instead of tested DNA. Similarly, positive controls (previously sequenced DNA of *A. phagocytophilum*) were used in each assay. PCR reactions were performed in personal thermal cycler (MyCycler, Bio-Rad) and PCR products were analyzed electrophoreticaly in 1.5% agarose gel stained with GelRed stain (Roche Diagnostics) and further visualized under the UV light.

2.4. Purification and sequencing

Randomly selected 16S rRNA and *groEL* amplicons of *A. phagocytophilum*, *msp4* amplicons of *A. ovis* and 16S rRNA amplicons of *Ca.* Neoehrlichia mikurensis were purified by using of a QIAquick PCR purification kit (Qiagen). Sequence analyses were performed by using of internal primers for both genes of *A. phagocytophilum*, and primers used for PCR detection of *msp4* of *A. ovis* and 16S rRNA of *Ca.* Neoehrlichia mikurensis at the Laboratory of Biomedical Microbiology and Immunology, (University of Veterinary Medicine and Pharmacy, Košice, Slovakia). The complementary strands of sequenced products were manually assembled. Sequences were compared with the GenBank entries by Blast N 2.2.13. Identities of obtained nucleotide sequences were computed using EMBOOS pairwise alignment (needle method) after manual removing of gaps.

The *groEL* nucleotide dataset was translated into amino acids and aligned using ClustalW algorithm in BioEdit 7.0.5.3 with default parameters (Hall, 1999). All ambiguously aligned sites were removed from further analyses. Edited dataset was used in nucleotides to infer phylogeny and contained 491 parsimony-informative sites. We used two approaches to reconstruct phylogenetic trees, Bayesian Inference (BI) and Maximum Parsimony (MP). BI tree was computed using MrBayes 3.1.2 (Huelsenbeck and Ronquist, 2005) with five million generations under GTR+I+Γ model with gamma distribution in four categories. AWTY (Nylander et al., 2008) was used to check the MCMC convergence and determine burn-in. MP was performed in PAUP* 4.0b10 (Wilgenbusch and Swofford, 2003) using TBR algorithm with random sequence addition. Bootstrap analysis with 100 replicates was done to statistically support tree topology.

2.5. Nucleotide sequences obtained in this study

16S rRNA and *groEL* sequences of *A. phagocytophilum*: 16S rRNA: red deer: GQ450278, GQ122212; roe deer: GQ450277, FJ984534; fallow deer: GQ450276, GQ122211; red fox: GQ162213, GQ162214; *I. ricinus* ticks: GU724963, GU724965, GU724967; dog: GU724971; bank vole: GU724970; *groEL*: European bison: JN935924 (from previous study/data not shown); dog: JN935925; red deer: JN935926, JN935927; roe deer: JN935929; *I. ricinus*: JN935928; goat: JN935930 (from previous study/data not shown); bank vole: JN935931; 16S rRNA of *Ca.* Neoehrlichia mikurensis: stripped field mouse: GU724969.

3. Results

3.1. A. phagocytophilum in I. ricinus ticks

Altogether 1075 questing *I. ricinus* ticks were screened. The overall prevalence of infection ranged from 1.4% to 5.5%. Prevalence of 2.8% was found in a hornbeam deciduous suburban forest park of Košice city; 4.36% in ticks from the basin nearby the Hornád river in area devastated by flood; 5.5% in forest in the locality Kavečany (district of Košice city). At the sampling site in the district of Michalovce city only one nymph tested positive and the prevalence of infection was 1.4%. The prevalence of 3.3% was recorded in Zádiel/Slovak Karst and the overall prevalence of *A. phagocytophilum* in ticks from the locations in the district of Rimavská Sobota city reached 1.5%.

Randomly chosen 16S and *groEL* positive amplicons were sequenced. All partial 16S rRNA sequences of *A. phagocytophilum* from *I. ricinus* ticks (GU724963, GU724965, GU724966 and GU724967) were identical in overlapped region and 100% homologous with the 16S obtained from Slovak red deer (GQ450278) or fallow deer (GQ450276).

Few samples were sequenced for *groEL* gene. Alignment of these sequences confirmed 100% identity in overlapped region. One representative was added to GenBank database (JN935928). *GroEL* sequence from Slovak *I. ricinus* clustered together with red deer and roe deer sequences from our study (Fig. 1).

Please cite this article in press as: Víchová, B., et al. *Anaplasma* infections in ticks and reservoir host from Slovakia. Infect. Genet. Evol. (2013), http://dx.doi.org/10.1016/j.meegid.2013.06.003

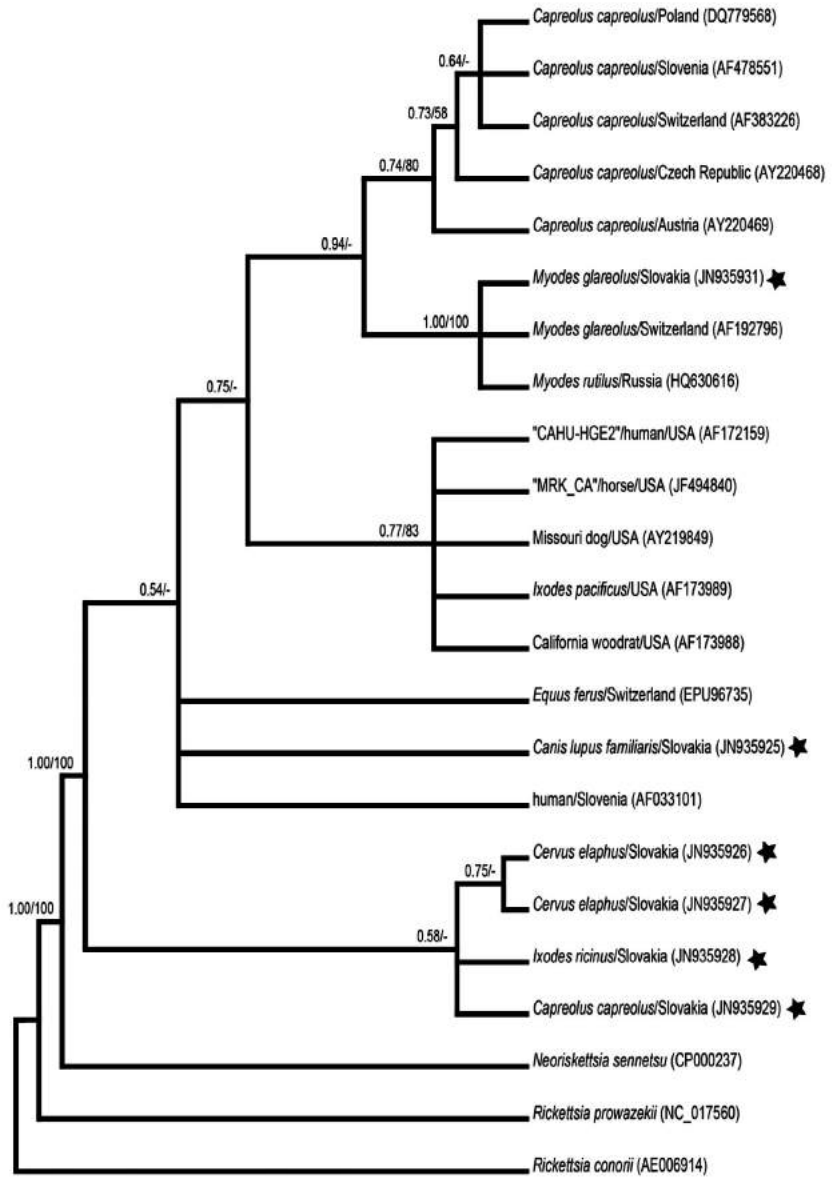


Fig. 1. Phylogenetic relationship of *A. phagocytophilum* groEL gene partial sequences downloaded from GenBank and those obtained in this study. Sequenced samples are asterisked. Phylogenetic tree represents the reconstruction of relationship produced by Bayesian inference. Numbers above branches indicate statistical support from Bayesian inference (posterior probability) and maximum parsimony (bootstrap values) respectively.

3.2. A. phagocytophilum in dogs

Only one dog blood sample tested positive for *A. phagocytophilum*. In blood smear stained with Hemacolor® Stain Set (Merck, Germany) morulae were not detected in peripheral blood granulocytes. Amplified portion of *groEL* gene from this sample was sequenced (GU724971) and included in phylogenetic analysis. This sequence clustered together with the sequence from Swiss horse, red deer and human patient from Slovenia (Petrovec et al., 2002). In four tested canine blood samples *Babesia canis canis* (EU165369) was confirmed (data not shown). In the group of 144 examined blood samples of police and military dogs we did not detect the presence of bacterial or protozoan DNA.

3.3. A. phagocytophilum in wild carnivores, ungulates and rodents

None of 248 examined red foxes from Slovakia tested positive for *A. phagocytophilum*. The prevalence of infection in wild boars was 16.7%. The overall prevalence of infection in red deer reached 17.5%. Eight tissue samples of roe deer and two chamois samples

Please cite this article in press as: Víchová, B., et al. *Anaplasma* infections in ticks and reservoir host from Slovakia. Infect. Genet. Evol. (2013). http://dx.doi.org/10.1016/j.meegid.2013.06.003

tested positive, what represents the overall prevalences of 61.5% and 3.5%, respectively.

Two of 3 tissue samples from fallow deer tested positive. Few 16S rRNA amplicons were randomly chosen and sequenced. Nucleotide sequences from red deer (GQ122212), roe deer (FJ984534) and fallow deer (GQ122211) were identical in 526 bp long overlapped region with each other and with 16S rRNA of *A. phagocytophilum* obtained previously from deer ked removed from red deer in Slovakia (GQ175174), Slovak roe deer (GQ450277) or red fox from Poland (GQ162214). Similarly, *GroEL* sequences from roe and red deer clustered together with *I. ricinus* sequence obtained in this study (Fig. 1).

The presence of *A. phagocytophilum* was confirmed in two DNA extracts from rodents. One yellow-necked mouse (*A. flavicollis*) and one bank vole (*M. glareolus*) trapped at the sampling site within the district of Ružomberok city tested positive. The overall prevalence of infection at this locality was 4.54%, and the average prevalence reached 0.69%. Both sequences from rodents were identical in overlapped regions of tested 16S rRNA and *groEL* genes. *A. phagocytophilum* sequences from our study showed 100% similarity within *groEL* gene with sequence obtained from the Swiss bank vole (AF192796) and clustered together with sequences from northern red-backed vole (*Myodes rutilus*) (HQ630616) and common shrew (*Sorex araneus*) (HQ630617) from Russia. In studied portion of 16S rRNA, our sequence showed 100% similarity with sequence from *A. agrarius* from eastern Slovakia (EF121955).

For the illustration of the relationship between *A. phagocytophilum* sequences obtained from Slovak hosts and ticks we constructed the phylogenetic tree by using of Bayesian inference (Fig. 1). All partial nucleotide sequences of 16S rRNA gene from our study were identical and showed 99–100% similarity with sequences from vector ticks or various hosts previously obtained elsewhere in the world. Nucleotide sequence similarities between obtained *groEL* gene portions ranged from 92.8% to 99.7%. Despite the shortness of obtained nucleotide sequences (JN935924-31), we found altogether 31 polymorphic sites within analyzed gene portion. Multiple alignments showed that 22 out of all 31 nucleotide substitutions occurred exclusively in sequence derived from bank vole and most (64.0%) of revealed substitutions had a character of transitions. Positions of mutations were calculated with respect to the complete *groEL* gene of *A. phagocytophilum* "HZ strain" with a length of 1653 bp (NC007797) (Table 1). Obtained sequences with a length of 388 bp were translated to the corresponding amino acid sequences with a start codon GTT and compared with referential sequence of complete *groEL* gene. Two changes, caused by the A ↔ G substitution at the second base of the codon (AAT–AGT) occurred. Asparagin (N) was substituted by serine (S) (Table 2). Both changes were neutral and likely do not dramatically alter secondary structure and the function of coded protein.

Additionally, at three out of five sampling sites, the presence of *Ca.* Neoehrlichia mikurensis was detected in *Appodemus* spp. and *M. glareolus* species. At the locality Rozhanovce, the prevalence of *Ca.* Neoehrlichia mikurensis infection reached 13.97%. Five (8.62%) out of 58 tested DNA extracts from the locality Šebastovce were infected, and 25 of 91 examined samples from the district of Lučenec city tested positive, what represents the prevalence of 27.47%. The mean prevalence of *Ca.* Neoehrlichia mikurensis in rodents was 10.75%. Obtained sequences were identical with 16S rRNA obtained previously from *I. ricinus* ticks from Slovakia (GU724964). The Blast analysis demonstrated the 99.3% similarity with nucleotide sequence obtained from the long-tailed brown-toothed shrew (*Soriculus leucops*) from China (GU227699).

3.4. Anaplasma in ruminants

All cattle samples tested negative for the presence of *A. phagocytophilum* and *A. marginale*. Similarly, *A. phagocytophilum* did not occur in examined sheep flocks. On the other hand, presence of

Table 1
Polymorphic sites within the studied portion of *groEL* operon in obtained sequences (order of sequences in first row: referential sequence NC007797; European bison JN935924; dog JN935925; red deer JN935926; red deer JN935927; *I. ricinus* JN935928; roe deer JN935929; goat JN935930; bank vole JN935931).

Position in referential sequence/position in obtained sequence 5-3	Referential sequence	Bison	Dog	Red deer	Red deer	*I. ricinus*	Roe deer	Goat	Bank vole
831/20	G	G	G	G	G	G	G	G	A
840/29	C	C	C	T	T	T	T	C	C
846/35	C	C	C	C	C	C	C	T	T
849/38	A	G	G	G	G	G	G	A	A
850/39	A	A	A	A	A	C	A	A	A
855/44	A	A	A	A	A	A	A	A	G
864/53	T	T	T	T	T	T	T	T	A
885/74	A	A	A	A	A	A	A	A	G
909/98	C	C	C	C	C	C	C	C	T
933/122	C	C	C	T	T	T	T	T	T
949/138	C	C	C	C	C	C	C	C	T
951/140	T	T	T	T	T	T	T	T	G
969/158	G	A	A	A	A	A	A	A	T
975164	C	C	C	C	C	C	C	C	T
987/176	A	A	A	A	A	A	A	A	G
996/185	T	C	C	C	C	C	C	C	T
999/188	A	A	A	A	A	A	A	A	C
1102/191	T	C	T	T	T	T	T	T	T
1020/209	T	T	T	T	T	T	T	T	C
1043/232	A	A	A	A	A	A	A	A	G
1059/248	G	G	G	G	G	G	G	G	A
1067/256	A	A	A	A	A	A	A	A	G
1101/290	A	A	A	A	A	A	A	A	G
1110/299	A	A	A	A	A	A	A	A	G
1113/302	G	G	A	A	G	G	A	G	T
1128/317	C	C	T	C	C	C	T	C	T
1137/326	A	A	A	A	A	A	A	A	G
1170/359	G	G	G	G	G	G	G	G	T
1176/365	A	A	A	A	A	G	A	A	G
1179/368	C	C	C	C	C	C	C	C	T
1197/386	T	T	T	T	T	T	T	T	C

Please cite this article in press as: Víchová, B., et al. *Anaplasma* infections in ticks and reservoir host from Slovakia. Infect. Genet. Evol. (2013), http://dx.doi.org/10.1016/j.meegid.2013.06.003

Table 2
Mutations observed in *groEL* region and their consequences in coded protein (calculated from GTT start codon).

Position in DNA	Codon	Mutation type
18	GCG → GCA	Transition
27	TTC → TTT	Transition
33	GAC → GAT	Transition
36	AGA → AGG	Transition
37	AGA → CGA	Transversion
42	AAA → AAG	Transition
51	CTT → CTA	Transversion
72	GTA → GTG	Transition
96	GAC → GAT	Transition
120	GAC → GAT	Transition
136/138	CTT → TTG	Transition/transversion
156	GTG → GTA → GTT	Transition/transversion
159	CGA → CGC	Transversion
162	ATC → ATT	Transition
174	GCA → GCG	Transition
183	ATT → ATC	Transition
186	ATA → ATC	Transversion
189	GGC → GGT	Transition
207	TCT → TCC	Transition
230	AAT → AGT	Transition N → S
246	CAG → CAA	Transition
254	AAT → AGT	Transition N → S
288	AGA → AGG	Transition
297	TTA → TTG	Transition
300	GCG → GCA → GCT	Transition/transversion
315	GGC → GGT	Transition
324	GTA → GTG	Transition
357	GTG → GTT	Transversion
363	GAA → GAG	Transition
366	CGC → CGT	Transition
384	GAT → GAC	Transition

A. ovis was confirmed in tested sheep samples. Rates of *A. ovis* infection reached considerable degrees. The overall prevalence of infection ranged from 6.2% to 50.0%.

4. Discussion

Several species of wild and domesticated animals as well as competent vector ticks were screened for the presence of members of the family Anaplasmataceae by using of PCR based molecular methods. In the present study we confirmed *A. phagocytophilum* in questing ixodid ticks from all sampling sites, in domesticated canine, in wild ungulates including wild boars, roe and red deer, fallow deer, alpine chamois and wild rodents.

In Europe, the rate of *A. phagocytophilum* infection in questing *I. ricinus* varies across locations from zero or very low up to 30% and it is usually higher in adults than nymphs (Blanco and Oteo, 2002; Strle, 2004; Rosef et al., 2009). In Russia, *bacteria* were detected in *I. ricinus* ticks collected in the Baltic region (Alekseev et al., 2001) and *Ixodes persulcatus* ticks collected in Vologda Province (Eremeeva et al., 2006), in Western Siberia (Rar et al., 2005), and in the studied areas from the Urals to the Far East of Russia (Shpynov et al., 2006; Shpynov, 2012). In Slovakia, the prevalence of infection in *I. ricinus* ticks is low (under 5.0%). The infection rate in tested questing ticks ranged from 1.4% to 5.5%, what is in accordance with previously published data (Špitalská and Kocianová, 2003; Derdáková et al., 2003; Smetanová et al., 2006).

Lately, it has been shown that *A. phagocytophilum* is maintained in enzootic foci even in the absence of *I. ricinus* ticks, with *I. trianguliceps* ticks as the main vectors (Ogden et al., 1998; Bown et al., 2003, 2008). Ogden with co-workers (1998) did not find *A. phagocytophilum* in rodents trapped in areas where *I. ricinus* ticks were abundant, but *I. trianguliceps* ticks were not present. In Slovakia, *A. phagocytophilum* was not revealed in rodent samples from areas where *I. trianguliceps* did not occur and all feeding *I. ricinus* ticks removed from infected rodents were free of infection (Pangrácová et al., 2012). Studies from Europe support the hypothesis that *A. phagocytophilum* ecotypes, associated with rodents are more likely maintained in foci by *I. trianguliceps* not by *I. ricinus* ticks.

Only one dog blood sample from our study tested positive for *A. phagocytophilum*. The tick infestation was the common feature of all tested dogs. The vast majority of blood specimens originated from clinically healthy individuals or the most common signs included increased temperature, apathy, lameness, anorexia, vomiting and local inflammations at the tick-bite site. *Anaplasma* – positive dog had no clinical symptoms leading to the suspicion for anaplasmosis or other tick-borne disease. None of police and military dogs tested positive. However, these results were not surprising as all dogs were under intensive veterinary control.

The presence of *A. phagocytophilum* in red foxes from Slovakia has so far not been considered. Similarly, we did not find *A. phagocytophilum* in tested animals. Only serological studies from Switzerland have suggested that red foxes might be infected with Anaplasma (Pusterla et al., 1999). First molecular evidence of bacteria in red foxes was recently recorded in Poland (Karbowiak et al., 2009). These results demonstrate the potential ability of foxes to serve as a reservoir of infection for dogs and other wild canine, but further studies are needed.

Tested cattle and sheep were free of *A. phagocytophilum* and *A. marginale* infection. In Europe, bovine anaplasmosis caused by *A. phagocytophilum* is relatively widespread and clinical manifestations, caused by the presence of bacteria have been recorded (Woldehiwet and Scott, 1993). However, cattle from our study were kept in tie-in stalls or housed within fenced grazing areas. The likelihood of coming into contact with wild ungulates infested with potentially infected *I. ricinus* ticks was significantly minimized. Unlike deer, small mammals are not affected by fencing and are able to ensure the circulation of Anaplasma even within fenced areas (Bown et al., 2008). Anyway, we did not record *A. phagocytophilum* in cattle, but this could also be caused by the differences in virulence of *A. phagocytophilum* strains. To our knowledge, in Slovakia the clinical case of infection in cattle has not been published so far.

However, in sampled sheep flocks we confirmed the presence of *A. ovis*, an agent responsible for tick-borne fever of small ruminants. Infection was confirmed in all tested flocks from middle and northern Slovakia with the prevalence of infection from 6.2% to 50.0%. High level of *A. ovis* infection has previously been confirmed on farms located in southeastern and northern Slovakia. The authors revealed also the presence of mixed infections of *A. phagocytophilum* and *A. ovis*, what suggests that Slovakia is a country with endemic occurrence of pasture fever, gradually spreading further north (Derdáková et al., 2011).

Tested wild ungulates were infected with *A. phagocytophilum* to a considerable extent. High proportions of red deer and roe deer tested positive and the rates of infection reached 17.5% and 61.5%, respectively. Additionally, two out of three examined samples from fallow deer tested positive what suggest this species as another potential wildlife reservoir.

Totally, 57 samples of alpine chamois were examined. The prevalence of infection was 3.5%. This is the first evidence of *A. phagocytophilum* in alpine chamois from Slovakia. In opposite to strictly protect *Rupicapra rupicapra* subsp. tatrica, which occurs exclusively in the Tatra National Park, an alpine chamois from the National Park Slovak Paradise are considered as an over – reproduced and are legally hunted as a fair game. Chamois live at moderately high altitudes and are adapted to living in open rocky terrains and alpine meadows. They can be found at elevations of at least 3,600 m and during the winter go to lower forests of around 800 m (Aulagnier et al., 2008). The chamois can carry infected ticks

Please cite this article in press as: Víchová, B., et al. *Anaplasma* infections in ticks and reservoir host from Slovakia. Infect. Genet. Evol. (2013), http://dx.doi.org/10.1016/j.meegid.2013.06.003

while overcoming various altitudes and affect the spreading of ticks and tick-borne pathogens further in higher altitudes.

Our study confirms the presence of *A. phagocytophilum* in 16.7% of tested wild boars. Previous studies suggested that wild boars may carry human pathogenic ecotype of *A. phagocytophilum* (Petrovec et al., 2002; Strasek Smrdel et al., 2009; Michalik et al., 2012). As far as boars are most abundant and widespread game in Europe, the identification and further study of zoonotic potential of *A. phagocytophilum* strains could support the surveillance for this pathogen in people from the same regions. Nevertheless, all our 16S rRNA sequences from boars were identical to sequences obtained from the other tested ungulates. Unfortunately, from wild boars, we were not able to obtain clearly legible *groEL* sequences. High rates of infection in examined wild animals suggest that ungulates in Slovakia are involved in the enzootic circulation of pathogen and serve as competent reservoirs for *A. phagocytophilum* as it has already been suggested by Štefanidesová et al. (2008).

Altogether 286 samples of rodents were screened for the presence of *A. phagocytophilum*. Only two rodents carried *A. phagocytophilum* and the average prevalence of infection was 0.69%. Previously, the presence of *A. phagocytophilum* was confirmed in *A. agrarius* and *A. uralensis* from eastern Slovakia (Štefančíková et al., 2008). Sequences obtained in this study from bank vole and yellow – necked mouse were identical in overlapped regions of 16S rRNA and *groEL* genes. These sequences were different than those derived from humans what suggest, that European rodents do not harbour *A. phagocytophilum* strains with strong zoonotic potential such as those from the United States.

In tested genomic DNA from blood and tissue samples of studied rodents (*Appodemus* spp., *M. glareolus*) the presence of *Ca.* Neoehrlichia mikurensis was confirmed. Prevalence of infection reached over 10.0%, what lies within the range found in rodents from other European countries (Andersson and Raberg, 2011; Jahfari et al., 2012; Vayssier-Taussat et al., 2012). Of interest in present study was also the diversity of *A. phagocytophilum* from wide spectrum of animal species. Obtained 16S rRNA and *groEL* sequences were analyzed in order to track intraspecific heterogeneity of *A. phagocytophilum* and to assess the possibility of co-circulation of different ecotypes in Slovakia. In Europe, 16S rRNA gene does not seem to be an appropriate locus for the analyses of heterogeneity as it is too conservative. Similarly, 16S rRNA sequences from our study did not reveal any SNPs. On the other hand, the groESL heat shock operon is widely used for determination of the diversity and the analyses of different *groESL* sequences already revealed considerable degree of heterogeneity (Rymaszewska, 2008).

5. Conclusions

The results of the study reported herein suggest that *A. phagocytophilum* infections in various regions of Slovakia are maintained in wild ungulates especially, with *I. ricinus* ticks serving as biological vectors. Knowledge of the presence of pathogens, the rates of infection and identity of circulating genotypes in vector ticks and hosts in different regions are crucial for the assessment of the risk ratio for human health and may help to establish the surveillance and control programs for vector-transmitted pathogens.

We have detected the presence of *A. phagocytophilum* in a wide spectrum of tested animals and confirmed the undoubted importance of wild ungulates for its circulation in natural foci of Slovakia. There appears to be greater strain diversity among *A. phagocytophilum* from vectors and hosts in Europe in opposite to US strains. We have obtained different nucleotide sequences from ticks and ungulates in comparison to those from rodents. The results suggest that European rodents most likely do not harbour *A. phagocytophilum* strains with strong zoonotic potential such as isolates from the United States.

Acknowledgements

This study has been realized thanks to the financial support of the Research & Development Operational Programme funded by the ERDF: Environmental protection against parasitozoonoses under the influence of global climate and social changes (code ITMS: 26220220116) (0.5), Projects of Slovak Agency for Research and Development APVV 0267-10, LPP 0341-06 and Projects of the Scientific Grant Agency of the Ministry of Education SR and Slovak Academy of Sciences VEGA 2/0113/12, VEGA 2/0055/11, VEGA 2/0137/10 and VEGA 1/07911/11. The authors would like to thank to Dr. Astéria Štefančíková, Dr. Eva Bullová, Dr. Martina Miterpáková and Dr. Zuzana Hurníková, from the Institute of Parasitology (Slovak Academy of Sciences) and to Prof. Ladislav Paule from the Technical University in Zvolen, who kindly provided samples for this study.

References

Alberdi, M.P., Walker, A.R., Urquhart, K.A., 2000. Field evidence that roe deer (*Capreolus capreolus*) are a natural host for *Ehrlichia phagocytophila*. Epidemiol. Infect. 124, 315–323.

Alekseev, A.N., Dubinina, H.V., Van De Pol, I., Schouls, L.M., 2001. Identification of *Ehrlichia* spp. and *Borrelia burgdorferi* in *Ixodes* ticks in the Baltic regions of Russia. J. Clin. Microbiol. 39, 2237–2242.

Andersson, M., Raberg, L., 2011. Wild rodents and novel human pathogen *Candidatus* Neoehrlichia mikurensis, Southern Sweden. Emerg. Infect. Dis. 17, 1716–1718.

Aulagnier, S., Giannatos, G., Herrero, J., 2008. Rupicapra rupicapra. IUCN 2012. IUCN Red List of Threatened Species. Version 2012.2. <www.iucnredlist.org>.

Black, W.C., Roehrdanz, R.L., 1998. Mitochondrial gene order is not conserved in arthropods: prostriate and metastriate tick mitochondrial genomes. Mol. Biol. Evol. 15 (12), 1772–1785.

Blanco, J.R., Oteo, J.A., 2002. Human granulocytic ehrlichiosis in Europe. Clin. Microbiol. Infect. 8, 763–772.

Bown, K.J., Begon, M., Bennet, M., Woldehiwet, Z., Ogden, N.H., 2003. Seasonal dynamics of *Anaplasma phagocytophila* in a rodent-tick (*Ixodes trianguliceps*) system, United Kingdom. Emerg. Infect. Dis. 9 (1), 63–70.

Bown, K.J., Lambin, X., Telford, G.R., Ogden, N.H., Telfer, S., Woldehiwet, Z., Birtles, R.J., 2008. Relative importance of *Ixodes ricinus* and *Ixodes trianguliceps* as vectors for *Anaplasma phagocytophilum* and *Babesia microti* in field vole (*Microtus agrestis*) populations. Appl. Environ. Microbiol. 74, 7118–7125.

Carpi, G., Bertolotti, L., Pecchioli, E., Cagnacci, F., Rizzoli, A., 2009. *Anaplasma phagocytophilum* groEL gene heterogeneity in *Ixodes ricinus* larvae feeding on roe deer in Northeastern Italy. Vector Borne Zoonotic Dis. 9 (2), 179–184.

de la Fuente, J., Naranjo, V., Ruiz-Fons, F., Hofle, U., de Mera Ig, F., Villanua, D., Almazán, C., Torina, A., Caracappa, S., Kocan, K.M., Gortazár, C., 2005a. Potential vertebrate reservoir hosts and invertebrate vectors of *Anaplasma marginale* and *A. phagocytophilum* in central Spain. Vector Borne Zoonotic Dis. 5 (4), 390–401.

de la Fuente, J., Torina, A., Caracappa, S., Tumino, G., Furla, R., Almazan, C., Kocan, K.M., 2005b. Serologic and molecular characterization of *Anaplasma* species infection in farm animal and ticks from Sicily. Vet. Parasitol. 133 (4), 357–362.

de la Fuente, J., Massung, R.F., Wong, S.J., Chu, F.K., Lutz, H., Meli, M., von Loewenich, F.D., Grzeszczuk, A., Torina, A., Caracappa, S., Mangold, A.J., Naranjo, V., Stuen, S., Kocan, K.M., 2005c. Sequence analysis of the msp4 gene of *Anaplasma phagocytophilum* strains. J. Clin. Microbiol. 43 (3), 1309–1317.

de la Fuente, J., Ruiz-Fons, F., Naranjo, V., Torina, A., Rodriguez, O., Gortázar, C., 2008. Evidence of *Anaplasma* infections in European roe deer (*Capreolus capreolus*) from southern Spain. Res. Vet. 84, 382–386.

Derdáková, M., Halánová, M., Stanko, M., Štefančíková, A., Čisláková, L., Peťko, B., 2003. Molecular evidence for *Anaplasma phagocytophilum* and *Borrelia burgdorferi* sensu lato in *Ixodes ricinus* ticks from eastern Slovakia. Ann. Agric. Environ. Med. 10 (2), 269–271.

Derdáková, M., Štefančíková, A., Špitalská, E., Taragel'ová, V., Košťálová, T., Hrkľová, G., Kybicová, K., Schánilec, P., Majláthová, V., Várady, M., Peťko, B., 2011. Emergence and genetic variability of *Anaplasma* species in small ruminants and ticks from Central Europe. Vet. Microbiol. 153 (3–4), 293–298.

Dubinský, P., Malczewski, A., Miterpáková, M., Gawor, J., Reiterová, K., 2006. *Echinococcus multilocularis* in the red fox *Vulpes vulpes* from the East Carpathian region of Poland and the Slovak Republic. J. Helminthol. 80 (3), 243–247.

Dumler, J.S., Barbet, A.F., Bekker, C.P., Dasch, G.A., Palmer, G.H., Ray, S.C., Rikihisa, Y., Rurangirwa, F.R., 2001. Reorganization of genera in the families Rickettsiaceae and Anaplasmataceae in the order Rickettsiales: unification of some species of *Ehrlichia* with *Anaplasma*, *Cowdria* with *Ehrlichia* and *Ehrlichia* with *Neorickettsia*, descriptions of six new species combinations and designation of

Please cite this article in press as: Víchová, B., et al. *Anaplasma* infections in ticks and reservoir host from Slovakia. Infect. Genet. Evol. (2013), http://dx.doi.org/10.1016/j.meegid.2013.06.003

Ehrlichia equi and 'HGE agent' as subjective synonyms of *Ehrlichia phagocytophila*. Int. J. Syst. Evol. Microbiol. 51 (Pt 6), 2145–2165.

Eremeeva, M.E., Oliveira, A., Robinson, J.B., Ribakova, N., Tokarevich, N.K., Dasch, G.A., 2006. Prevalence of bacterial agents in *Ixodes persulcatus* ticks from the Vologda Province of Russia. Ann. N. Y. Acad. Sci. 1078, 291–298.

Fehr, J.S., Bloemberg, G.V., Ritter, C., Hombach, M., Lüscher, T.F., Weber, R., Keller, P.M., 2010. Septicemia caused by tick-borne bacterial pathogen *Candidatus* Neoehrlichia mikurensis. Emerg. Infect. Dis. 16 (7), 1127–1129.

Guy, E.C., Stanek, G., 1991. Detection of *Borrelia burgdorferi* in patients with Lyme disease by the polymerase chain reaction. J. Clin. Pathol. 44 (7), 610–621.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.

Huelsenbeck, J.P., Ronquist, F., 2005. Bayesian analysis of molecular evolution using MrBayes. In: Nielsen, R. (Ed.), Statistical Methods in Molecular Evolution. Springer, New York.

Hulínska, D., Langrova, K., Pejčoch, M., Pavlásek, I., 2004. Detection of *Anaplasma phagocytophilum* in animals by real-time polymerase chain reaction. APMIS 112, 239–247.

Humair, P.F., Douet, V., Morán Cadenas, F., Schouls, L.M., Van de Pol, I., Gern, L., 2007. Molecular identification of bloodmeal source in *Ixodes ricinus* ticks using 12S rDNA as a genetic marker. J. Med. Entomol. 44 (5), 869–880.

Hurníková, Z., Dubinský, P., 2009. Long-term survey on *Trichinella* prevalence in wildlife of Slovakia. Vet. Parasitol. 159 (3–4), 276–280.

Jahfari, S., Fonville, M., Hengeveld, P., Reusken, C., Scholte, E.J., Takken, W., 2012. Prevalence of *Neoehrlichia mikurensis* in ticks and rodents from north-west Europe. Parasites Vectors 5, 74.

Karbowiak, G., Víchová, B., Majláthová, V., Hapunik, J., Peťko, B., 2009. *Anaplasma phagocytophilum* infection of red foxes (*Vulpes vulpes*). Ann. Agric. Environ. Med. 16, 299–300.

Kawahara, M., Rikihisa, Y., Isogai, E., Takahashi, M., Misumi, H., Suto, C., Shibata, S., Tsuji, M., 2004. Ultrastructure and phylogenetic analysis of '*Candidatus* Neoehrlichia mikurensis' in the family Anaplasmataceae, isolated from wild rats and found in *Ixodes ovatus* ticks. Int. J. Syst. Evol. Microbiol. 54 (5), 1837–1843.

Liz, J.S., Sumner, J.W., Pfister, K., Brossard, M., 2002. PCR detection and serological evidence of granulocytic ehrlichial infection in roe deer (*Capreolus capreolus*) and chamois (*Rupicapra rupicapra*). J. Clin. Microbiol. 40, 892–897.

Lotric-Furlan, S., Petrovec, M., Zupanc, T.A., Nicholson, W.L., Sumner, J.W., Childs, J.E., Strle, F., 1998. Human granulocytic ehrlichiosis in Europe: clinical and laboratory findings for four patients from Slovenia. Clin. Infect. Dis. 27 (3), 424–428.

Massung, R.F., Slater, K., Owens, J.H., Nicholson, W.L., Mather, T.N., Solberg, V.B., Olson, J.G., 1998. Nested PCR assay for detection of granulocytic ehrlichiae. J. Clin. Microbiol. 36 (4), 1090–1095.

Massung, R.F., Lee, K., Mauel, M., Gusa, A., 2002. Characterization of the rRNA genes of *Ehrlichia chaffeensis* and *Anaplasma phagocytophila*. DNA Cell Biol. 21 (8), 587–596.

Massung, R.F., Mather, T.N., Levin, M.L., 2006. Reservoir competency of goats for the Ap-variant 1 strain of *Anaplasma phagocytophilum*. Infect. Immun. 74, 1373–1375.

Michalik, J., Stańczak, J., Cieniuch, S., Racewicz, M., Sikora, B., Dabert, M., 2012. Wild boars as hosts of human-pathogenic *Anaplasma phagocytophilum* variants. Emerg. Infect. Dis. 18 (6), 998–1001.

Miterpáková, M., Antolová, D., Hurníková, Z., Dubinský, P., Pavlacka, A., Németh, J., 2010. *Dirofilaria* infections in working dogs in Slovakia. J. Helminthol. 84 (2), 173–176.

Nylander, J.A., Wilgenbusch, J.C., Warren, D.L., Swofford, D.L., 2008. AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. Bioinformatics 24 (4), 581–583.

Ogden, N.H., Brown, K., Horrocks, B.K., Woldehiwet, Z., Bennet, M., 1998. Granulocytic Ehrlichiae infection in ixodid ticks and mammals in woodlands and uplands of the U.K. Med. Vet. Entomol. 12 (4), 423–429.

Pangrácová, L., Stanko, M., Mošanský, L., Bona, M., Fričová, J., Víchová, B., Peťko, B. 2012. Association of *Anaplasma phagocytophilum* genotypes and Neoehrlichia mikurensis with rodents as reservoir hosts in Slovakia. In: MEEGID XI: 11the International Conference on Molecular Epidemiology and Evolutionary Genetics of Infectious Diseases, October 30-November 2nd, 2012, New Orleans, USA. Program and Abstract Book – New Orleans: MEEGID, 2012, p. 12.

Petrovec, M., Bidovec, A., Sumner, J.W., Nicholson, W.L., Childs, J.E., Avsic-Zupan, T., 2002. Infection with *Anaplasma phagocytophila* in cervids from Slovenia: evidence of two genotypic lineages. Wien. Klin. Wochenschr. 114 (13–14), 641–647.

Petrovec, M., Sixl, W., Schweiger, R., Mikulasek, S., Elke, L., Wust, G., Marth, E., Strasek, K., Stunzner, D., Avsic-Zupanc, T., 2003. Infections of wild animals with *Anaplasma phagocytophila* in Austria and the Czech Republic. Ann. N. Y. Acad. Sci. 990, 103–106.

Polin, H., Hufnagl, P., Haunschmid, R., Grubel, F., Ladurner, G., 2004. Molecular evidence of *Anaplasma phagocytophilum* in *Ixodes ricinus* ticks and wild animals in Austria. J. Clin. Microbiol. 42, 2285–2286.

Pusterla, N., Deplazes, P., Braun, U., Lutz, H., 1999. Serological evidence of infection with *Ehrlichia* spp. in red foxes (*Vulpes vulpes*) in Switzerland. J. Clin. Microbiol. 37, 1168–1169.

Rar, V.A., Fomenko, N.V., Dobrotvorsky, A.K., Livanova, N.N., Rudakova, S.A., Fedorov, E.G., Astanin, V.B., Morozova, O.V., 2005. Tick borne pathogen detection, Western Siberia, Russia. Emerg. Infect. Dis. 11, 1708–1715.

Rosef, O., Paulauskas, A., Radzijevskaja, J., 2009. Prevalence of *Borrelia burgdorferi* sensu lato and *Anaplasma phagocytophilum* in questing *Ixodes ricinus* ticks in relation to the density of wild cervids. Acta Vet. Scand. 51 (1), 47.

Rymaszewska, A., 2008. Divergence within the marker region of the groESL operon in *Anaplasma phagocytophilum*. Eur. J. Clin. Microbiol. Infect. Dis. 27 (11), 1025–1036.

Schouls, L.M., van de Pol, I., Rijpkema, S.G.T., Schot, C.S., 1999. Detection and identification of *Ehrlichia*, *Borrelia burgdorferi* sensu lato, and *Bartonella* species in Dutch *Ixodes ricinus* ticks. J. Clin. Microbiol. 37, 2215–2222.

Shpynov, S., 2012. *Ixodes persulcatus*, a major vector of Alphaproteobacteria in Russia. Ticks Tick Borne Dis. 3, 304–306.

Shpynov, S., Fournier, P.E., Rudakov, N., Tarasevich, I., Raoult, D., 2006. Detection of members of the genera *Rickettsia*, *Anaplasma*, and *Ehrlichia* in ticks collected in the Asiatic part of Russia. Ann. N. Y. Acad. Sci. 1078, 378–383.

Siuda, K., 1993. Kleszcze polski (Acari: Ixodida). Systematya i Rozmieszczenie czest II. Polsie Towaryzstvo Parazytologiczne, Warszawa, ISBN: 83-601349-0-X, 376 p.

Smetanová, K., Schwarzová, K., Kociánová, E., 2006. Detection of *Anaplasma phagocytophilum*, *Coxiella burnetii*, *Rickettsia* spp., and *Borrelia burgdorferi* s. l. in ticks, and wild-living animals in western and middle Slovakia. Ann. N. Y. Acad. Sci. 1078, 312–315.

Špitalská, E., Kocianová, E., 2003. Tick-Borne microorganisms in southwestern Slovakia. Ann. N. Y. Acad. Sci. 990, 196–200.

Špitalská, E., Boldiš, V., Košťanová, Z., Kocianová, E., Štefanidesová, K., 2008. Incidence of various tick-borne microorganisms in rodents and ticks of central Slovakia. Acta Virol. 52 (3), 175–179.

Štefančíková, A., Derdáková, M., Lenčáková, D., Ivanová, R., Stanko, M., Čisláková, L., Peťko, B., 2008. Serological and molecular detection of *Borrelia burgdorferi* sensu lato and Anaplasmataceae in rodents. Folia Microbiol. 53 (6), 493–499.

Štefanidesová, K., Kocianová, E., Boldiš, V., Košťanová, Z., Kanka, P., Nemethová, D., Špitálska, E., 2008. Evidence of *Anaplasma phagocytophilum* and *Rickettsia helvetica* infection in free-ranging ungulates in central Slovakia. Eur. J. Wildl. Res. 54, 519–524.

Strasek Smrdel, K., Bidovec, A., Malovrh, T., Petrovec, M., Duh, D., Avsic Zupanc, T., 2009. Detection of *Anaplasma phagocytophilum* in wild boar in Slovenia. Clin. Microbiol. Infect. 15 (Suppl. 2), 50–52.

Strle, F., 2004. Human granulocytic ehrlichiosis in Europe. Int. J. Med. Microb. 293 (Suppl. 37), 27–35.

Stuen, S., 2007. *Anaplasma phagocytophilum* – the most widespread tick-borne infection in animals in Europe. Vet. Res. Commun. (Suppl. 1), 79–84.

Stuen, S., Van de Pol, I., Bergstrom, K., Schouls, L.M., 2002. Identification of *Anaplasma phagocytophila* (formerly *Ehrlichia phagocytophila*) variants in blood from sheep in Norway. J. Clin. Microbiol. 40, 3192–3197.

Sumner, J.W., Yabsley, M.J., Arens, M.Q., Buenning, G., Storch, G.A., Davidson, W.R., 2003. Determination of white-tailed deer agent groESL operon sequences for phylogenetic and diagnostic applications. Ann. N. Y. Acad. Sci. 990, 699–700.

Vayssier-Taussat, M., Le Rhun, D., Buffet, J.-P., Maaoui, N., Galan, M., Guivier, Charbonnel, N., Cosson, J.-F., 2012. *Candidatus* Neoehrlichia mikurensis in bank voles, France. Emerg. Infect. Dis.. http://dx.doi.org/10.3201/eid1812.120846.

von Loewenich, F.D., Geissdörfer, W., Disqué, C., Matten, J., Schett, G., Sakka, S.G., Bogdan, C., 2010. Detection of "*Candidatus Neoehrlichia mikurensis*" in two patients with severe febrile illnesses: evidence for a European sequence variant. J. Clin. Microbiol. 48 (7), 2630–2635.

Welinder-Olsson, C., Kjellin, E., Vaht, K., Jacobsson, S., Wennerås, C., 2010. First case of human "*Candidatus Neoehrlichia mikurensis*" infection in a febrile patient with chronic lymphocytic leukemia. J. Clin. Microbiol. 48 (5), 1956–1959.

Wilgenbusch, J.C., Swofford, D., 2003. Inferring evolutionary trees with PAUP*. Curr Protoc Bioinformatics, Chapter 6:Unit 6.4.

Woldehiwet, Z., Scott, G.R., 1993. 1993. Tick-borne (pasture) fever. In: Woldehiwet, Z., Ristic, M. (Eds.), Rickettsial and chlamydial diseases of domestic animals. Pergamon Press, Oxford, pp. 233–254.

Please cite this article in press as: Víchová, B., et al. *Anaplasma* infections in ticks and reservoir host from Slovakia. Infect. Genet. Evol. (2013), http://dx.doi.org/10.1016/j.meegid.2013.06.003

5 Curriculum vitae

Tomáš Chrudimský

Faculty of Science, University of South Bohemia, Branišovská 31, České Budějovice 370 05, Czech Republic
Institute of Soil Biology, Biology Centre ASCR, v.v.i., Na Sádkách 7, České Budějovice 370 05, Czech Republic
tel.: +420-38-777-5775
e-mail: chrudimsky@upb.cas.cz

Date of birth: 13. 4. 1983
Place of birth: Chlumec nad Cidlinou
Education
2005: Bachelor degree in Biology. Faculty of Biological Sciences, University of South Bohemia

2007: Master degree in Experimental Biology. Faculty of Science, University of South Bohemia

2007-present: Doctoral student of Molecular and Cell Biolgy and Genetics. Faculty of Science, University of South Bohemia

Abroad stays and fellowships
2007 – Visiting researcher. Department of vector-borne diseases, Institute of Parasitology, Slovak Academy of Sciences, Košice, Slovakia

2010 – Visiting Researcher. Laboratory of Professor Kerry M. Oliver, The University of Georgia, Athens, GA, USA

Teaching and mentoring experience
2007-2010: Co-supervisor of bachelor student Filip Husník, Faculty of Science, University of South Bohemia.

2009-2011: Teaching assistant - Biochemistry laboratory for Biology and Environmental studies; Biochemistry laboratory 1 at Faculty of Science, University of South Bohemia.

Publications

Moore RB, Oborník M, Janouškovec J, Chrudimský T, Vancová M, Green DH, Wright SW, Davies NW, Bolch CJ, Heimann K, Šlapeta J, Hoegh-Guldberg O, Logsdon JM, Carter DA. A photosynthetic alveolate closely related to apicomplexan parasites. Nature. 2008 Feb 21;451(7181):959-63. doi: 10.1038/nature06635. Erratum in: Nature. 2008 Apr 17;452(7189):900. PubMed PMID: 18288187.

Oborník M, Janouškovec J, Chrudimský T, Lukeš J. Evolution of the apicoplast and its hosts: from heterotrophy to autotrophy and back again. Int J Parasitol. 2009 Jan;39(1):1-12. doi: 10.1016/j.ijpara.2008.07.010. Epub 2008 Sep 12. Review.PubMed PMID: 18822291.

Chrudimská T, Chrudimský T, Golovchenko M, Rudenko N, Grubhoffer L. New defensins from hard and soft ticks: similarities, differences, and phylogenetic analyses. Vet Parasitol. 2010 Feb 10;167(2-4):298-303. doi: 10.1016/j.vetpar.2009.09.032. Epub 2009 Sep 23. PubMed PMID: 19836137.

Husník F, Chrudimský T, Hypša V. Multiple origins of endosymbiosis within the Enterobacteriaceae (γ-Proteobacteria): convergence of complex phylogenetic approaches. BMC Biol. 2011 Dec 28;9:87. doi: 10.1186/1741-7007-9-87. PubMed PMID:22201529; PubMed Central PMCID: PMC3271043.

Chrudimský T, Husník F, Nováková E, Hypša V. *Candidatus* Sodalis melophagi sp. nov.: phylogenetically independent comparative model to the tsetse fly symbiont Sodalis glossinidius. PLoS One. 2012;7(7):e40354. doi: 10.1371/journal.pone.0040354. Epub 2012 Jul 17. PubMed PMID: 22815743; PubMed Central PMCID: PMC3398932.

Kilian P, Valdes JJ, Lecina-Casas D, Chrudimský T, Rů ek D. The variability of the large genomic segment of Ťahyňa orthobunyavirus and an all-atom exploration of its anti-viral drug resistance. Infect Genet Evol. 2013 Dec;20:304-11. doi: 10.1016/j.meegid.2013.09.023. Epub 2013 Sep 30. PubMed PMID: 24090866.

Víchová B, Majláthová V, Nováková M, Stanko M, Hviščová I, Pangrácová L, Chrudimský T, Curlík J, Peťko B. *Anaplasma* infections in ticks and reservoir host from Slovakia. Infect Genet Evol. 2013 Jun 11. pii: S1567-1348(13)00228-1. doi: 10.1016/j.meegid.2013.06.003. [Epub ahead of print] PubMed PMID: 23770268.

chrudimsky@upb.cas.cz

Origins of bacterial endosymbionts in arthropods
Ph.D. Thesis

University of South Bohemia in České Budějovice
Faculty of Science
Branišovská 31
CZ-37005 České Budějovice, Czech Republic

Phone: +420 387 772 244
www.prf.jcu.cz, e-mail: sekret@prf.jcu.cz