



UNIVERSITÄT  
BAYREUTH

Universität Bayreuth, TÖK I, •95440 Bayreuth, GERMANY

School of Doctoral Studies in Biological  
Sciences  
Faculty of Science  
University of South Bohemia  
Ceske Budejovice  
Cech Republic

LEHRSTUHL TIERÖKOLOGIE I

Prof. Dr. Heike Feldhaar

Postanschrift:  
Universität Bayreuth  
95440 Bayreuth

Gebäude: NW I  
Universitätsstraße 30  
95447 Bayreuth  
Telefon: +49 / 0921 / 55-2645  
+49 / 0921 / 55-2651 (Skr.)  
Telefax: +49 / 0921 / 55-2784

Internet: <http://www.bayceer.uni-bayreuth.de/toek1/index.php?lang=de>

e-mail: [feldhaar@uni-bayreuth.de](mailto:feldhaar@uni-bayreuth.de)

9.6.2012

## **Review of PhD-thesis of Mrs Eva Novakova entitled “Molecular phylogeny and genome evolution of insect symbiotic bacteria“**

The main focus of the PhD-thesis presented by Mrs Eva Novakova is to integrate phylogenetic and genomic studies of endosymbiotic bacteria in order to understand evolutionary patterns involved in the transition from free-living to endosymbiotic life-style. One focus of her work is on the comparison of genomic alterations following integration into insect hosts with different ecological niches. Differences in host-niche such as trophic level of the host (e.g. herbivore vs. blood-feeding) are presumed to cascade down to the symbiotic bacteria and result in different patterns of genome degradation of endosymbionts.

The other focus of her work is on methodological problems involved in performing phylogenetic analyses of endosymbiotic bacteria. Mrs Novakova stresses that most of the phylogenetic analysis tools may fail on endosymbiotic bacteria. E.g. the most widely used models of DNA evolution such as the general time reversible-model (GTR) should fail when dealing with endosymbiotic bacteria as the substitution process seems to be directed, violating the assumption of time-reversibility.

The PhD-thesis presented by Mrs Eva Novakova comprises a general introduction and seven manuscripts, of which four have already been published. Of the remaining three one is already submitted to PlosOne and two others are presented as manuscripts that can be submitted shortly. The four published chapters are two invited book chapters as well as a BMC Microbiology publication and one that appeared in MBE. Both journals are high ranking journals listed in ISI. Mrs Novakova is first author on both of these publications as well as on

one book chapter. She will be first author on two of the three remaining yet unpublished manuscripts.

The first publication is an invited book chapter, which appeared in the prestigious and widely read 3<sup>rd</sup> Volume of "Insect Symbioses" edited by K. Bourtzis and T. Miller. The book chapter gives an excellent overview of the problems of the widely used molecular phylogenetic approaches when applied to endosymbionts.

The second publication appeared in BMC Microbiology. In this publication a phylogenetic analysis of the endosymbiont *Arsenophonus* sampled from a broad variety of hosts is presented. This publication highlights on the one hand the diverse relationships of the endosymbiont with its diverse hosts and on the other hand discusses potential problem in the phylogenetic analysis. The variety of the *Arsenophonus* sequences included in this study enabled Mrs Novakova and her coauthors to uncover various types of symbiosis and evolutionary patterns. Thus, whereas some *Arsenophonus* strains show signs of parallel host-symbiont evolution whereas in other cases closely related strains were isolated from distantly related hosts. The discussion of this publication is excellent with respect to potential sources of artefacts in phylogenetic analyses – this part is often too short or non-existent in comparable publications. This publication is an important step forward in order to understand the evolution of endosymbiosis as *Arsenophonus* seems to be a bacterium with a large diversity of host-interactions ranging from obligate to facultative and from mutualistic to parasitic. This is the first study including a broad spectrum of strains that were either published in GenBank or obtained by the authors themselves.

In the fourth publication (an invited book chapter) Mrs Novakova took part in giving a broad overview over the bacterial genus *Arsenophonus* and its diverse relations with insect hosts. In this publication, the genome sequences of the genus known to date are compared.

In the seventh – still unpublished – manuscript, Mrs Novakova and coauthors present data on comparative genomics of *Arsenophonus*. Two draft genome sequences are presented: the genome of the secondary endosymbiont of a triatomine bug *Arsenophonus triatominarum* as well as that of the blood-feeding dipteran *Melophagus ovinus*. This work is important since *Arsenophonus* shows a broad spectrum of stages of genome degeneration. *Arsenophonus* – similar to *Sodalis* – shows a tendency towards higher genome degeneration in more mutualistic associations with its hosts.

The third publication appeared in the high ranking journal Molecular Biology and Evolution. Here Mrs Novakova conducted an extensive study on an ancient gene transfer from fungi to



aphids. More than 30 aphid species were screened for the presence of the genes for carotenoid biosynthesis that aphids have obtained from the fungi. The gene products are required by animals for body coloration, as pigments in vision and other cellular functions. Until recently it was assumed that horizontal gene transfer – the hallmark of genetic versatility of bacteria – does not play an important role in eukaryotic genomes. However, through the increasing availability of genomic resources of endosymbiotic bacteria as well as their insect hosts it is becoming clear that there are events of gene transfer to the eukaryotic host. The publication on the diversification of genes involved in carotenoid biosynthesis is among the first examining the evolution of such genes transferred from a different eukaryote. Here, Mrs Novakova can show that these genes have diversified greatly within aphids after acquisition early in their evolution. The closest relatives of aphids, adelgids, are shown here to also possess these genes.

In the fifth publication, which is under revision in the ISI-listed journal PlosOne, Mrs Novakova is coauthor. Here data is presented on basic morphological and molecular traits as well as on the type three secretion system of the endosymbiont *Cand. Sodalis melophagi*. This endosymbiont is a close relative of *Sodalis glossinidius* harboured by tsetse flies. This *Sodalis* strain has been transferred to insect cell culture successfully, which is described in the manuscript. As only very few endosymbiotic bacteria can be cultured outside the host to date. Future comparative work on these endosymbionts can contribute important insights into the first steps of endosymbiosis, namely how endosymbionts manage to enter cells and can become integrated into the insect host.

In the sixth publication Mrs Novakova explores the utility of genes of *Buchnera*, the P-endosymbiont of aphids, as markers for phylogenetic reconstruction of the aphid hosts. A wide range of genes of the endosymbiont is used for this study. Mrs Novakova and her coauthors show that these genes can indeed resolve the aphid phylogeny.

All publications presented in this PhD-thesis are of very high scientific quality and value. With this PhD-thesis Mrs Novakova contributes very important new results in the field of genome evolution of endosymbionts. I would like to highlight her ability to integrate the results obtained by phylogenetic approaches with the general biology of the organisms. The discussions of her publications are written extremely well. While numerous researchers present phylogenetic data on all sorts of organisms only very few are able to transport why exactly their work is important. This is the strength of Mrs Novakova. In addition, the critical assessment of the results due to potential methodological problems is written very well. She

is able to "translate" the methodological problems into a language that researchers from outside the field will understand.

I am also very pleased with the general form of the PhD-thesis. There are very few misspellings in the general introduction, but other than that this thesis is formally correct and of very high scientific quality.

In sum, I have the pleasure to recommend that Mrs Novakova has shown that she highly deserves a PhD-degree in Biological Sciences as she has presented an excellent thesis.



Prof. Dr. Heike Feldhaar  
Lehrstuhl Tierökologie I  
Universität Bayreuth  
95440 Bayreuth  
Germany

Prof. Dr. Heike Feldhaar





Florida Museum of Natural History  
Department of Natural History  
Dickinson Hall, PO Box 117800  
Gainesville, FL 32611-0287

352-273-1971  
352-846-0287 Fax  
dreed@flmnh.ufl.edu  
<http://www.flmnh.ufl.edu>

June 12, 2012

Dear Eva:

Congratulations on an excellent dissertation. It was a pleasure to read. I would like to begin with a quick evaluation, then pose some general questions for you to answer—as if I were there for your defense, and finally a few comments on the dissertation itself.

Evaluation of Dissertation:

Eva's dissertation contained an introduction and 7 complete manuscripts, four of which are already published, and three of which are ready for publication. This is an outstanding accomplishment, and she is to be commended. Her work is both innovative and synthetic, and shows the maturity of an independent scientist. I would be exceedingly proud of one of my PhD students if they were to present a dissertation of this quality, and Eva's speaks highly not only of her, but also of her advisor.

Of the four published manuscripts two were edited book chapters and two were published in peer-reviewed journals that are very well respected in our field. This alone is quite an accomplishment. However, Eva has produced three additional manuscripts that either have been, or soon will be, submitted for publication.

Manuscript #5 has been submitted to PLoS One, which is an up-and-coming journal in the PLoS family. This manuscript, in part a species description, was well written and provides a very interesting comparison of *Sodalis* taxa. It will add to the growing body of literature on this very interesting and dynamic genus of bacterium.

Manuscript #6 was quite innovative and asked the question "to what degree can we use the genes of a primary endosymbiont to reconstruct the evolutionary history of its insect host?" Eva contends in this manuscript that the bacterial genes outperform the often-used COI gene from the insect hosts in reconstructing the host's evolutionary history. I love this approach, and in fact I use it in my own research, however, I would suggest that genes from the p-endosymbionts are no better than mtDNA genes in one respect. Just as mtDNA is inherited as a single unit and therefore represents a single view of history, the endosymbiont genes are providing the same narrow window into the host's evolutionary past. Having said that, I still think this is a solid manuscript that I will no-doubt cite many times.

Manuscript #7 is a comparative genomics study of four *Arsenophonus* taxa that are somewhat closely related but differ in their life history strategies. Eva contends that the four taxa lie along a spectrum leading to obligate endosymbiont, and that by studying them together we can gain a better understanding of symbiogenesis. I agree with this premise and like this manuscript very much. This manuscript, however, is in the greatest need of editing compared to the others. But, it may likely have the greatest research impact of them all. The work that needs to be done with this manuscript is primarily in tightening up the writing, but nothing serious in my opinion. In closing I would like to say that this is THE best dissertation I have read in several years. Great job and congratulations.

*The Foundation for The Gator Nation*

Questions to Answer:

Please be brief in your responses (i.e., a short paragraph).

-You discuss in various places in your dissertation both the uniqueness of *Candidatus* *Riesia* spp. and the fact that it is nested within the well-known genus *Arsenophonus*. You seem to favor subsuming the proposed genus *Riesia* within *Arsenophonus*. Can you describe what, if anything, would convince you to reclassify the entire *Arsenophonus* clade instead of sinking *Riesia*?

-Would you answer the question above differently if we were talking about insects or vertebrates instead of bacteria?

-I really liked your work using cutting-edge methods of phylogenetic inference. What do you think are the most important changes on the horizon for phylogenetics—not the advances you are already using, but the advances that will come in 1-5 years?

-Similarly, do you think that we'll ever really overcome the problems of extreme nucleotide composition bias in phylogenetics? This seems like a particularly difficult problem for those of us that study p-endosymbionts. If so, what will the solution be, more complex CAT-like models?

-Pardon me if you addressed this in the dissertation, if so I missed it. But, would you say that the multiple *Buchnera* genes used to infer the evolutionary history of their hosts are in any way independent of one another? Or, would you contend that they are analogs to mtDNA from the aphids in that they are inherited as a single unit and therefore only represent a single history (unlike nuclear markers from the aphids themselves)? If so, is this a major or minor limitation of this technique?

-I am curious about your thoughts on the high degree of purifying selection coupled with high mutation rates found in *Ca. Riesia pediculicola*. Do you think this is specific to the genome of that p-endosymbiont, or do you think we would find a similar pattern in other (even unrelated) p-endosymbionts of sucking lice? Why or why not?



General Comments on Dissertation:

Here are some questions and comments that I have about your dissertation beyond the grammatical and other editorial suggestions that I made directly in your PDF. You do not need to respond to me *per se*, but think about these as you make your final changes to your dissertation.

-I really think that you have to wait until more is known about *Arsenophonus* before suggesting a reclassification of the group. I really appreciate it when taxonomists take a stand for or against reclassification, and I'm glad you did. However, I think in this case there is a bit more work that we could do that would help us make a better decision and that it would be prudent to wait a while longer.

-on pg 50 you say that unfortunately, calculating evolutionary rates in p-endos and insects is of little use. I'm still not sure I understand why you say that. Is it because in most systems you don't have a real calibration point that can be used, and the researchers have to pull some arbitrary rate from another taxon? You'll see three comments by me in that area of your dissertation, and I think I still don't understand the issue.

-manuscript #7 seems to be in the greatest need of editing before publication. It needs careful editing, but is otherwise in very good shape. I thought that manuscripts 5 and 6 were much better written and are ready to be submitted (I realize manuscript 5 already has). With manuscript #7 its mostly grammatical errors and writing that can be made clearer. One thing to consider is that the Results section is considerably longer than usual, but that may be because it is a comparative genomics paper. However, it should cause you to ask yourself if every result is worth discussing in a paragraph, or can some of that information be presented in a shorter format (e.g., combining paragraphs when possible, move some information to additional tables, etc.)?

Again, congratulations on an excellent dissertation. It was a pleasure to read, and I hope our paths cross again very soon.

Best regards,



David L. Reed  
Associate Curator of Mammals  
Florida Museum of Natural History

## **Review of the PhD thesis „Molecular phylogeny and genome evolution of insect symbiotic bacteria“**

PhD candidate: Eva Nováková

reviewer: Miroslav Oborník

This PhD thesis is composed of general introduction (it contains chapters called Background; Objectives and models; Research outcomes; Conclusions and future prospects; 1-12 + references 13-20), which is supplemented by set of 7 scientific papers in different stages of completion. Two of them were published as book chapters (MS#1 and MS#4), two others in the form of articles in scientific journals (MS#2 and MS#3) and the last three are presented in the form of unpublished manuscripts (MS#5-7). Eva Nováková is the first author of two published papers (MS#2 and 3) and two unpublished manuscripts (MS#6 and 7). This well demonstrates very good publication activity of the candidate, which with no doubts fulfils requirements of the University of South Bohemia for publication outcomes necessary to obtain PhD degree. As far as I can evaluate it, the thesis is written in good English.

### **Introduction:**

In the Introduction the candidate very roughly summarize the current knowledge concerning the bacterial symbionts of insects and mainly shows her contribution to this knowledge. I have several questions related to this introductory part and also to conclusions at the end of the chapter:

**Q1:** The candidate evaluates reductive processes during the evolution of genomes of symbiotic bacteria as beneficial in the view of molecular phylogeny. However, those genomes could pass through various genetic processes including gene or genome duplications before the free-living bacteria got a symbiotic character and such processes can be now hidden and impossible to be uncovered. There can be hidden paralogs, thanks to the gradual loss of paralogs in different lineages that follows gene duplication. At the same time, such unknown genome rearrangements can substantially influence phylogenetic analyses and their interpretations. Can candidate explain this aspect in details?

**Q2:** It is mentioned in the thesis that: “The short lengths of next generation sequencing results in fragmented assemblies or even misassembled rearrangements”. This is obvious problem in assembly of eukaryotic genomes, but I do not see it as much important in the case of symbiotic bacteria with a highly reduced genomes. What are lengths of reads of particular NGS method (e.g. 454, Illumina...)? Where would you like to sequence your genomes by Sanger sequencing as you proposed, when most of Sanger sequencing facilities gave up on this particular sequencing method and replaced it by NGS? How would you solve this problem?

**Q3:** How would you explain the high number of gene duplications (affecting more than 2000 gene families) in the genus *Acyrtosiphon*?

**Q5:** Please, define the boundary line, where the endosymbiotic bacteria or any similar entity can be understood as an organelle. Which process defines organelle?

### **Manuscript #1**

The MS#1 is a review chapter in the book that summarizes the current knowledge concerning the bacterial symbionts of insects. Inclusion of this book chapter into the thesis fully justifies the brief character of the thesis introduction focusing on the thesis results and conclusions instead of introducing the reader to the topic. I have two question to this manuscript:



**Q6:** When I see definitions of P and S symbionts; are bacterial symbionts of insects strictly divided to these well identifiable types or is there any continuum between these two types of symbiotic relationships?

**Q7:** What is the main obstacle in generating multigene phylogenetic analysis?

#### **Manuscript#2**

The second manuscript published in BMC Microbiology describes an analysis of extending set of bacteria of the genus *Arsenophonus*, as one of the most widespread bacteria associated to insects. Authors noted that particularly 16S rRNA gene dataset is very sensitive to various methodological artifact. At the same time they propose the necessity to use other molecular marker than 16S rDNA. My questions:

**Q8:** Do you have some gene candidates for the low level phylogeny of *Arsenophonus*?

**Q9:** On the page 3, various topologies of the tree are shown, either computed by MP with different Tv/Ts ratio. However, none of the presented trees show branch support. Why?

**Q10:** Have you tested those topologies by AU test or any similar topology tests?

#### **Manuscript#3**

This paper published in Molecular Biology and Evolution describes rare and clear case of lateral gene transfer from fungi to aphids. Although I am not very big fan of explaining everything unusual by LGT, I like this particular case, especially when aphids are known to host also eukaryotic symbionts.

**Q11:** Can you, please, mention which fungi form symbiotic relationships with aphids?

**Q12:** Is there any explanation alternative to LGT?

#### **Manuscript#4**

The fourth manuscript is again chapter from the book. It is reviewing the knowledge about the bacterial symbiotic genus *Arsenophonus*, composition of the genus, characterization of each described species and their interaction with the insect hosts; it is simply very nice and informative review. I have no questions to this MS.

#### **Manuscript#5**

This is an unpublished manuscript (submitted to PLoS ONE) describing new model for (secondary) symbiotic relationship between bacteria and insect. This particular model, *Candidatus Sodalis melophagi*, which was isolated from *Melophagus ovinus*, represents an early or intermediate stage of symbiosis usable to study various aspects influencing relationships between host and symbiotic bacteria. New bacterium is well characterized and its phylogenetic position in the frame of the genus *Sodalis* is shown. The fact that the relation of *Can. S. melophagi* and its host highly resemble a couple *Sodalis glossinidus* and tsetse fly makes this model very prospective. The independent origin of the symbiotic relationship between *Candidatus Sodalis melophagi* and *Melophagus ovinus* is remarkable. It also demonstrates the ability of horizontal transmission of bacteria of the genus *Sodalis*.

**Q13:** When looking at the trees, it is obvious that the genus *Sodalis* needs some taxonomic revision, since most of the members of the clade are not named at all. Do you plan to make such revision?

#### **Manuscript#6**

This unpublished manuscript tries to address purely resolved evolution of aphids by phylogenetic analysis of their symbiotic bacteria of the genus *Buchnera*. It appears that such analysis is more informative than the one based on aphid genes. This is an interesting



approach; however, the symbiont data tends to be influenced by phylogenetic artifacts. Since most of previous work on symbiont phylogeny was dedicated to elimination of these artifacts, I would not be happy to deal with these weird sequences again. However, any progress in problematic phylogeny as is the case of aphids helps.

**(comment)** In the figure one, also very low node support is shown in the tree. It has been generally assumed that Bayesian posterior probabilities (PP) below 95% indicate low or no support. You show all the PPs in figure 1, including branches with PP 0.32, 0.33 or 0.35, which is in my opinion not correct. Also in the figure 2, all branches supported by PP higher than 0.85 are counted as supported, which I see incorrect again. These figures should be modified before submission.

**Q14:** In figure 3 the tree lacking statistical support at all is shown. Why?

**Q15:** The danger of LBA is mentioned in the article. But only a single method how to deal with LBA is proposed there: exclusion of long branches from the dataset. In the presented trees, majority of branches are long, so their exclusion would not be very helpful. Can you propose any other methods how the LBA can be eliminated? What are the principles of this particular artifact? Which method is the most sensitive to it?

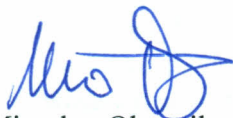
### Manuscript#7

The last and unpublished manuscript of the thesis demonstrates genomic adaptations of symbiotic bacteria of the genus *Arsenophonus* reflecting different life style of the symbiont. They show continuum from almost ordinary bacterial genomes of parasites and facultative symbionts to highly degenerated genomes of obligate mutualists. The genomic degeneration is usually reflected in losses of numerous genes and in substantial nucleotide bias. However, not all detected trends follow the universal direction in the frame of the genus *Arsenophonus*. This analysis of four genomes of endosymbiotic bacteria is really substantial piece of work.

**Q16:** In the supplementary tree S1 there are only two species of *Arsenophonus* shown and the other two are absent. Why?

### Summary:

This is an excellent PhD thesis that well demonstrates an outstanding effort of the candidate to understand the origin and evolution of the genome structure of bacterial symbionts in insects. It covers various aspects of the bacterial symbiosis starting from a simple phylogeny and ending in the genomic based approach. I highly appreciate the international mode of this PhD and I very much recommend keeping these contacts with the top laboratories alive. I cordially recommend this thesis for defense.



Miroslav Obornik  
Department of Molecular Biology  
Faculty of Science  
University of South Bohemia &  
Institute of Parasitology  
Biology Centre ASCR  
Branišovská 31  
37005 České Budějovice

In České Budějovice, June 13<sup>th</sup> 2012