POSUDEK OPONENTA HABILITAČNÍ PRÁCE

Masarykova univerzita

Uchazeč	Mgr. Natália Martínková, Ph.D.
Habilitační práce	Modelling in phylogenetic framework
Oponent	Prof. RND. Ivan Horáček, CSc.
Pracoviště oponenta, instituce	Univerzita Karlova v Praze, Katedra zoologie

viz Příloha 11_1a

Dotazy oponenta k obhajobě habilitační práce (počet dotazů dle zvážení oponenta)

1. Některé fylogenetické stromy v článku 2.1.2 naznačují pozici rodu *Arvicola* uvnitř rodu *Microtus* jakožto sesterské skupiny kladu *Chionomys*, což odporuje detailnímu fosilnímu záznamu ukazujícího víceméně pozvolný přechod rodu *Arvicola* z *Mimomys savini* téměř million let před FAD *Chionomys*. Jaké je pro to vysvětlení?
2. Do jaké miry může být mutační rychlost ovlivněna abundancí, demografií a kolonizační historií lokální populace?

Závěr

Habilitační práce Natálie Martínkové „Modelling in phylogenetic framework" ***splňuje*** požadavky standardně kladené na habilitační práce v oboru Zoologie.

Brno dne

....
podpis

A review of habilitation thesis

Natalie Martínková: Modelling in phylogenetic framework

Being invited to review the Natalie'a habilitation thesis I was quite curious which topic she choose to design it. I expected either the molecular phylogeny and phylogeography of rodents, adaptive rearrangements in island populations and dynamic of genotype changes in marginal populations of a species range or the geographic and contextual variation of population effects under a mass epidemic influence as has been the case with white nose syndrome of Holoarctic bats. The habilitant worked in all these branches, published a lot on each of these topics and obviously would be able to compile an acceptable habilitation thesis on whatever of them.

Yet, Natalie decided for an other design. The present version of her habilitation thesis is centered with topics of advance mathematical methods promising to improve the mining of phylogenetic and functional information from diverse empirical data, particularly the extensive genomic datasets combined with a broad taxonomic screening of diverse traits and contextual factors.

The covered methods are briefly surveyed in 50 pages of the opening text with numerous cross references to a selection of habilitant's reserach papers in supplement (18 items, on 247 pp. forming the major bulk of the volume), which examplify productive application of particular methods.

The major attention is paid to the factors influencing phylogenetic inferences from genetic data, particularly in context of the multilocus aligment in very large datesets with incomplete sequences and non-random distribution of mising data, i.e. under the conditons which drastically limit a possiblity of consistent phylogenetic inference but at the same time are quite common in research practice. Using simulation models the habilitant tested the effects of diverse prior inputs upon posterior probabilities of branch length with aid of maximum likelihood and Bayesian estimates of the tree spaces generated by Markov chain Monte Carlo sampling under different parametres of exponential distribution. She demonstrated the effects of missing data upon increase of branch length indicating unrealistically long phylogenies in Bayesian inferrences. She further discusses effects of differences in mutation rates at different loci and claims for supplementing phylogenetic procedure with analyses of the aspect called terrace, a cloud of trees differing in topology but showing the same likelihood. Extent of terraces and common features of included trees can thus provide a global information on reliability of suggested output tree topology. Actual application of these improvements are demonstrated on a series of molecular phylogenetic analyses of diverse mammalian and non-mammalian taxa (papers 2.1.1 - 2.1.8), particularly the papers devoted to enormously intricated topics of phylogenetic relations within extremely diversified genus *Microtus*. Here, the incomplete sequence data in most OTUs and obviously different substitution rates between individual mitochondrial and nuclear genes present the factor which makes the inferences of the multilocus alignment a task enormously complicated. Nevertheless a detailed terrace analysis (paper 1.1.2) provided to the inferred tree a robust support. Yet, similarly as indicated by some previous analyses (e.g. Galewski et al. 2006) still certain discrepances remain. It concerns either topology of branching and different signal indicated by some nuclear and the mitochodrial markers which probably masked the effects of former ones in a concatenated aligment. My question concerning such a possibility address particularly the inner position of the genus *Arvicola* within *Microtus* as a sister clade of *Chionomys*. Such a possibility seems to be in a strong contradiction with a detailed fossil record suggesting more or less gradual transition of the former genus from *Mimomys savini*, almost one milion year later than FAD of *Chionomys*. This point, not discussed in he respective paper, is undoubtedly worth of explanation.

The next part of the Chapter 1 of the thesis addresses the topic of molecular dating and possible shortcomings of molecular clock method. It discusses effects of different mutation rates at different loci and long branches produced by Bayesian inferences biased by missing data. Further it provides a simulation model analysing with aid of empirical data on mean selection values of random mutation the phenomenon of mutation rate decay. It show that during first generations the mutation dynamics results to rapid fixation of beneficial mutation and extensive decay of overall mutation rate. Aspects of practical application of molecular clock are discussed with examples of calibration techniques as

applied e.g. in papers 2.2.1 - 2.2.3. My question concerning that topic is as follows: to which degree could the mutation rate be influenced by abundance, demography and colonization history of a local population? One can imagine that possibly it might abruptly increase with disappearance of contraints and drift effects e.g. after an island colonization what eventually can produce a biased signal of a deep divergences as indicated e.g. in genetic data on Orkney population of *Microtus arvalis* mentioned in paper 2.2.3.

The other part of the Chapter 1 deal with aspects of codivergence and congruence among phylogenetic pattern of distant clades as a tool enabling to text hypotheses on coevolutionary relations among them. The referred papers on relations between New World arenaviruses and their rodent hosts (1.3.1) and divergent aspect of phylogenetic history iluminated by molecular and morphological analyses of the same clade (tree squirrels in paper 1.3.2) provide an excellent illustration of impressive outcomes of that approach.

Next part introduces modelling of species traits and comparative approach to analyses of genetic diversity. It operates with quantitative measures of phylogenetic correspondence of a trait appearance within phylogenetic structure of a community and results in a proposal of the K statistics characterizing globally the directionality pattern of a trait distribution. Similarly as in the last part of the Chapter 1, focused on quantitative techniques for identification of outliers distribution, the application is demonstrated by a reference to complex analyses of diverse phenomena associated with the white nose syndrome in bats, the topics investigated in frame of a large research project coordinated by the habilitant. Despite I am not completely convinced on explanatory power of the reported analyses in the case of white nose syndrome, I entirely admit that they attributed significantly to the complexity of the project outputs and hearing of them from international research community. In any case, the described methods particularly those quantifying dynamic of a trait distribution promise, being applied in a more suitable model situations, a novel views on pattern of adaptive evolution and its contextual background.

The penultimate paper of the thesis, attributed to another approach in outliers analysis (1.5.1), introduces a novel statistical technique, SigHunt (genomic signature hunter), for discovering genomic islands and identification of genomic loci impacted by horizontal gene transfer. The technique is based on quite elegant algorithm and can be relatively easily applied without any larger computing demands. It provides a promising tool for phylogenetic studies in the era of genomic analyses.

I will not comment on quantitative tools for evaluation author contribution (Chapter 2) and discuss the suggested contribution values of the habilitant. In my eyes all are very underestimated - the contribution of a teacher, a project leader and/or simply a colleague rich in previous experinces and inovative ideas can hardly be included in such calculations or expressed in terms of a percentage. And, I am pretty sure that just this was an essential contribution of the habilitant to most of the multi-authored paper included in the thesis. By the way, all these papers appeared in top journals of the branch (2xMol.Evolution, Scientific Reports, 2xPlosOne, Mol.Phyl.Evol., J.Biogeography, Bioinformatics, Chromosome Res. etc.).
Already these facts as well as the syntesizing ambitions of the thesis and innovative exposure of all the items composing it clearly demonstrate that the habilitant is a prominent scientific personality whose research production continuously responds to a very high international standard.

In conclusion: The thesis by Natalia Martinkova exceeds excessively the level expected for successful habilitation. Her qualities respond to the requirements associated with the respective academic degree without any doubt.

3 0 -03- 2019

Prof. Dr. Ivan Horáček CSc.
Department of Zoology
Faculty of Sciences
Charles University
Prague