

V I L L E D E
G E N È V EProf Tomas SCHOLZ
Institute of Parasitology
University of South Bohemia
Branisovska 31370 05 České Budjovice
Czech Republicjean.mariaux@mhn.ville-ge.ch
☎ +41 22 418 63 43*This my report on the thesis of Mr. Jan Brabec, submitted for the
grade of Ph.D. at the University of South Bohemia*

Despite the apparently simple question mentioned in the title of this thesis, the problem of understanding how the basal lineages of cestodes evolved is a difficult one, that has puzzled systematians for a long time. It is thus very welcome that someone dared to look at this question and I should first congratulate Mr. Brabec for this. These congratulations are even more deserved because he opted for a state of the art, but difficult, way to reach his goal: to develop a very new set of markers that would – in principle – allow to overcome the limits of the traditionally used characters, and to offer a significant source of independent information. This was a bold and risky topic and one must acknowledge Mr. Brabec for engaging on such uncertain path. Obviously the results he obtained were not on par with his initial expectations but others, that were probably not expected at the beginning of this work were obtained and this PhD work nevertheless allowed for a large array of very interesting outcomes. What I consider the most interesting is that Mr. Brabec went beyond the traditional “simple” use of molecular markers to elucidate the basal evolution of tapeworms and instead led basic research to enhance the use of these markers. The positive results of such results are particularly obvious in his finding on paralogous sequences in Caryophyllidea.

I briefly comment on a number of more precise points below.

I liked very much the introductory parts on the evolutionary history of flatworms and molecular systematics of Cestoda. Although this only represents a synthesis exercise, it is very nicely written and the author

demonstrated an excellent aptitude at cleverly summarizing a wide corpus of information.

I have no experience in this specific field and was pleasantly surprised by the clarity of the summary on SL *trans*-splicing. This again denotes the excellent communication skills of Mr. Brabec. One might however have appreciated a couple of illustrations to explicit some of the described mechanisms in this section. On that topic – which was central to the thesis – there is however a point, maybe naive, that I would like to hear the candidate develop. It is mentioned (P18) that the origin and evolution of *trans*-splicing is a debated point with the possibility that the phenomenon is homoplastic. It is furthermore stated that, potentially, only a small percentage of mRNAs are processed this way. I am thus wondering whether the chances of obtaining a sufficient database for phylogenetic comparisons were high enough, even if the characterization of SL genes had been more successful? How was this risk evaluated at the beginning of the thesis?

The demonstration of the existence of mt-paralogs in the nuclear genome of caryophyllids became eventually the most important published result of the thesis. This is a remarkable one that translated in a very nice paper (whose supplementary data might have been included in the thesis!). I think it is a remarkable achievement for Mr. Brabec, and a very useful one for any researcher in the field, even if it developed as “by-product” of his research (in his own terms!).

Concerning the paper on the characterization of human *Diphyllobothrium*. I appreciate the result and its elegant methodological development. This is clearly a nice achievement but its link with the main focus of the work is really circumstantial, and I am therefore not convinced by the need to integrate this paper in the thesis.

The papers on *Khawia*, *Kirstenella* and *Lobulovarium* are all very good contributions to the systematics of basal tapeworms. Although Mr Brabec is stating it modestly, his rigorous analyses of the molecular data in these papers are key contributions and significantly enhance their overall importance. This type of collaboration is extremely important, and although Mr. Brabec effective contribution might be limited to the molecular phylogenetic analyses, it remains fundamental that he could work at a more integrative level together with morphological systematicians. These papers also help justify the final title of the thesis!

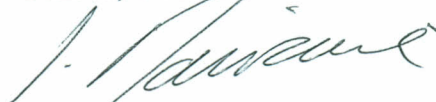
Of course what makes the strength of this thesis can also be seen as a weaker point in the sense that, globally, a certain lack of homogeneity is easy to perceive. This reaches the point where, should the original part of the research program have worked better, I wonder if the thesis might not have switched toward a more fundamental research in molecular biology instead of phylogenetics (with a study of the purpose and function of *trans*-splicing as admitted in the introduction). Although, as said above, the various contributions assembled in this work are complementary, I would

have expected a final synthesis with a personal conclusion from the author. In particular a concluding chapter on the status of the basic question of the whole study, i.e. the phylogeny of the basal cestodes lineages, and on the impact of this work, would have nicely concluded the thesis. This is basically the most important point that I feel is missing in this thesis, but it is an important one, and I would maybe encourage Mr. Brabec to think about putting his ideas on this point together in a short review paper in the future.

I have noted a few typos and minor editorial points on my version of the text and will communicate them directly to the author before the defense.

In conclusion: This is a very interesting work in which the candidate proved his ability to tackle various approaches in the field of molecular, and to some extent morphological, systematics. The strength of this work is two folds, first in the novelty and visionary approach of some of the selected molecular methods, and second in the diversity and complementarity of the methods used by Mr. Brabec. Its main weakness is the lack of a real synthesis that would better valorize the results of the thesis and help putting them in perspective. Globally however I consider that Mr. Brabec made a remarkable work and brought very significant results to the scientific community. He demonstrated his capacity to communicate efficiently, to work on difficult problems, to master and to develop various sophisticated techniques as well as to cleverly build on existing knowledge. For all these reasons, he certainly deserves a Doctor title and I recommend the acceptance of his thesis.

Geneva, 14/6/12



Prof J. Mariaux
Museum of Natural History of Geneva



To the
School of Doctoral Studies in Biological Sciences
University of South Bohemia in České Budejovice
Faculty of Science

Appraisal report of the doctoral thesis

„Molecular Systematics and Evolution of Basal Cestode Lineages“

by **Jan Brabec**

This is a cumulative PhD thesis encompassing five scientific publications in journals of significant international reputation with the applicant acting as first author in one publication (Int. J. Parasitol.), and a section on the identification and description of spliced-leader genes in basal cestode groups. The work addresses important questions of parasite flatworm phylogeny. Despite significant efforts in the past to resolve the phylogeny of tapeworms, one of the most successful and diverse groups of parasitic animals, a number of questions, particularly concerning the phylogeny of basal tapeworm groups, is still unresolved. In his PhD thesis, Jan Brabec now significantly contributes to this topic, provides explanation why phylogenetic studies on these organisms are problematic, and opens new experimental strategies for future investigations into the issue. Of special value in the context of this evaluation, due to the applicant's particularly significant contribution, are the study aspects concerning caryophyllidean phylogeny and the identification of the cestode spliced-leader genes. The identification of nuclear paralogs of mitochondrial marker genes (*numts*) in cestodes, described for the first time in Brabec et al. (2012), is indeed of particular interest for phylogeneticists and provides a clue why phylogenetic studies on these organisms (and maybe on cestodes in general, since *numts* are surely not confined to Caryophyllidea) using classical markers are so difficult. The study also stresses the need to identify a new generation of phylogenetic markers and with his contribution on spliced-leader trans-splicing in basal cestodes, Jan Brabec also provides highly valuable sequence information that most probably leads to respective crucial marker genes. Although this part of the thesis has not yet led to a

Prof. Dr. Klaus Brehm

Abt. Medizinische Parasitologie

Tel: 0931-201-46168
Sekretariat: 0931-201-46161
Fax: 0931-201-46445
E-Mail: kbrehm@hygiene.uni-wuerzburg.de
Web: www.hygiene.uni-wuerzburg.de

**Nationales Referenzzentrum
für Meningokokken**

Tel: 0931-201-46006
Fax: 0931-201-46445
Web: www.meningococcus.de

**Konsiliarlabor
für Echinokokkose**

Tel: 0931-201-46161
Fax: 0931-201-46445
Web: www.echinococcus.de

Diagnostische Laboratorien

akkreditiert nach DIN EN ISO 15189 u.
DIN EN ISO/IEC 17025



DAC-ML-0109-00-10

publication in an international journal, Jan Brabec has already carried out the (most difficult) experimental key steps in identifying trans-spliced genes from basal lineages. From what we meanwhile know through cestode genome analyses, the original hypothesis formulated by Jan Brabec in his thesis (i.e. that trans-splicing of particular genes is conserved in cestodes), exactly gets to the point. Genome analyses on schistosomes and *Echinococcus* clearly demonstrated that both organisms contain operon structures, the expression of which involves trans-splicing, and that trans-spliced genes in both groups are usually highly homologous, rendering it likely that homologs are also trans-spliced in basal cestodes. By concentrating on this particular group of genes it should thus be possible to easily get hold on homologs in various cestode groups that fulfill criteria as phylogenetic markers. Furthermore, the analyses of Jan Brabec also provide a solid basis for further investigations into the mechanisms (and evolution) of trans-splicing in cestodes. Technically, this part of the thesis has been very well performed. The author also provides an excellent overview of trans-splicing in different metazoan lineages, presents the data in a clear and concise way and delivers a highly qualified discussion of the data. The information on operons in trematodes and cestodes is highly relevant for using trans-spliced genes as phylogenetic markers, but has only recently (or not yet) been published and is, thus, not part of the thesis. For future publication on the topic, I nevertheless recommend that they are included in the rationale and discussion. Apart from the above mentioned achievements in cestode phylogeny, I also want to point out Jan Brabec's contribution to the development of the *Diphyllobothrium* multiplex PCR which is indeed highly useful for diagnostic aspects of cestode infections and has been very well performed. The thesis is exceptionally well conceived and written. Reading it was not only highly interesting regarding scientific content but also enjoyment of literature. This should be particularly pointed out since the author did not write the thesis in his mother tongue. Even for the non-phylogeneticist (as I am), Jan Brabec provides an easily understandable and comprehensive introduction into the topic and clearly points out the achievements of cestode phylogeny as well as present shortcomings and problems. The experimental procedures are outlined in sufficient detail, the data are clearly presented, and the author's conclusions are sound.

The quality of Jan Brabec's experimental work is reflected by the fact that it already contributed to five publications in internationally known journals, among which I would particularly like to mention the first-authorship in a full-length paper in the *International Journal for Parasitology* and the co-authorship in the *Journal of Clinical Microbiology*. I am also highly confident that the identification of trans-spliced marker genes and respective

phylogenetic analyses, as experimentally started in this thesis, will lead to additional publications in international journals.

Overall, I have a very well impression of this work. By his contribution to the five publications listed in this thesis, by his introduction into the topic, his additional data on trans-splicing in cestodes, and his discussion, he has clearly demonstrated his ability to carry out independent scientific work of high quality. I thus recommend that the School of Doctoral Studies accepts Jan Brabec's thesis without changes. In the case that grading is required, I would recommend **magna cum laude (very good)**.

Würzburg, 20.06.2012



Prof. Dr. Klaus Brehm



Report on Ph.D. Thesis by Jan Brabec

Molecular Systematics and Evolution of Basal Cestode Lineages

I am satisfied that, in this thesis, the candidate has demonstrated his ability to undertake scientific research. In particular, this is demonstrated by the unpublished chapter on spliced leader *trans*-splicing and by "Paper 1".

The thesis consists of a general introduction flowing into a section on cestode systematics, which I thought was very good. However, it would have benefitted from the inclusion of a few trees. At the very least, I would have expected a summary tree or other form of diagram to highlight what is accepted and what uncertain about ordinal relationships.

Following this section was one giving the aims of the thesis, and then another on spliced leader *trans*-splicing. This occupies pages 1-35. The last section of this (spliced leader *trans*-splicing) constitutes interesting new and unpublished material.

The candidate stated in the "aims" section that the initial goal of the project was "to resolve the problem of basal cestode orders by obtaining bigger amount of data by targeting a specific subpopulation of cestode mRNAs—the spliced leader (SL) *trans*-spliced genes." It was not explained to my satisfaction how work on these genes was going to help resolve the phylogeny of the basal cestodes. First, the SL exon needed to be characterised for each taxon. Then, mRNAs undergoing *trans*-splicing were to be amplified using primers against the SL exon at the 5' end and the 3' poly-A tail. It seemed to me that the chance of amplifying a strictly homologous set of mRNAs for each of a large number of taxa would be low. The candidate has subsequently modified the approach to take advantage of next-generation sequencing techniques, a much more feasible approach, but the work is ongoing and not part of the thesis.

Table 3. Is the sequence of primer T7NALUR correct? I was trying to work out how the "easy gene walking" worked, and noticed that this primer has "...AGGCGCA..." in the middle, whereas the two above it in the table have "...AGGGCGA...". Generally, more information on the locations of and/or rationale for design of the primers would have been useful. I eventually worked it out, but doing so required some effort.

P26. Fig 1: one of the potential structures for Ligula has U, the other T. The secondary structures could usefully be numbered.

P33. Some references are out of alphabetical order.

The remaining five chapters each consisted of a single, published (or in-press) research paper. Since the five research papers have been peer-reviewed, they have already been "polished". The candidate's contributions to these papers were primarily in the molecular aspects, so I have concentrated on these in my comments.

Paper I

The first of the published papers ("Substitution saturation and nuclear paralogs of commonly employed phylogenetic markers in the Caryophyllidea ...") is the one in which the candidate has best been able to demonstrate his skills in molecular techniques and analysis.

A priori, should we expect mitochondrial sequence data to be of use in resolving deep branches in an ancient taxon of tapeworms?

The phrase on page 41 "while heteroplasmy is difficult to evaluate..." needs further explanation and justification. More consideration should have been given to the possibility of heteroplasmy. Might cestodes in fact have different types of mitochondria in different tissues? See, for example, Takamiya, S., K. Fukuda, T. Nakamura, T. Aoki, and H. Sugiyama. 2010. *Paragonimus westermani*

possesses aerobic and anaerobic mitochondria in different tissues, adapting to fluctuating oxygen tension in microaerobic habitats. *International Journal for Parasitology* 40:1651-1658.

Page 41, Table 2, footnote e. Substitutions relative to what? Each other? What is the difference between "mutations" and "substitutions"? And footnote c – were indels in multiples of three?

Is it worth exploring phylogenetic/systematic levels at which significant saturation in the mt dataset no longer exists? For example, what happens if the outgroup is omitted when the calculations are done? Indeed, there might be a strong case for omitting the outgroup when calculating levels of saturation. What happens if single clades on the tree are considered? Such work might usefully inform studies such as that reported in Paper III.

Fig 2. Are *Glaridacris* and *Monobothrium* distinct from each other? What does this study say about traditionally recognised families? I would have liked a bit more on morphology here.

Surely the supplementary material should also have been included in the thesis?

Can you speculate on causes for the lack of basal resolution even when genes with very different rates of evolution are used?

Paper II

The second paper ("Multiplex PCR for differential identification of broad tapeworms...") demonstrates a different approach to molecular identification studies.

Some of the species are essentially marine and some freshwater. Are there any ecological or epidemiological consequences arising from occupation of different habitats?

The abstract implies that a test like this can assist in preventing [marine] species from invading new areas. How feasible is this?

Large numbers of eggs were used as a source of DNA. What might be the smallest number of eggs that can yield sufficient DNA for PCR in tests like this?

It needs to be emphasised that appropriate positive and negative controls are required. And there are always the dangers of a) misdiagnosing something new as a known species or b) getting a negative result because of unreported intraspecific variation affecting primer sites.

Paper III

"Revision of *Khawia* ...".

Were whole specimens used for DNA extraction? Or was it possible to use just part of each worm, leaving the taxonomically important portions intact and hence retaining the worm as a voucher?

Can anything be said about the age of the genus *Khawia* or its biogeography? There is a paragraph on this in the section "Interrelationships of *Khawia* species based on molecular data". Recent human introductions seem to have been important: can any older patterns be detected?

There is a statement made in the discussion: "A phylogenetic tree based on 11 morphological characters was almost completely incongruent with the tree inferred from molecular data ...". However, the morphological data seem to contain so much homoplasy that they might actually be congruent with almost any molecular tree. A test of this can be done using incongruence length difference methods. See, for example, Lee, M. S. Y. 2001. Uninformative characters and apparent conflict between molecules and morphology. *Molecular Biology and Evolution* 18:676-680.

In the section "Conflict of morphological and molecular data ..." there is discussion of the morphological dissimilarity but molecular similarity of *K. sinensis* and *K. saurogobii*. This was not

taken as far as a decision on the identities of these worms. Which is to be preferred: molecular or morphological data (since they conflict so strongly here)? If we are to use molecular data for species delimitation, these two species might be regarded as synonyms (not a very satisfactory outcome in my opinion). How much molecular difference should be accepted as intraspecific variation? At what level of divergence should we regard worms as belonging to separate species? If *K. sinensis* and *K. saurogobii* are retained as separate species, then obviously *Khawia* sp and *K. armeniaca* would have to be regarded as distinct using any molecular criteria. I appreciate there are no simple answers to these questions, but a debate on this topic is long overdue.

The word "specious" is used a few times in this paper. It should be "speciose". "Specious" has a very different meaning.

Paper IV

"Bothriocephalidean tapeworms (Cestoda) of freshwater fish in Africa..."

Figure 62 (tree). Although there is an "African clade" it includes two examples of the introduced (from Asia) *Bothriocephalus acheilognathi*, one from Africa and one from Europe. Another two species of *Bothriocephalus* were included, but were each in a very different part of the tree. I could see no explanation of this finding in the text. I would have expected paraphyly of such an important genus to be mentioned and discussed. I also noticed that the seven newly generated sequences (African plus *Senga*) form a cluster. It might be worth looking at the sequence alignment in case some anomaly has crept in.

Figures 49-52 are out of sequence relative to other figures.

Last two lines of the discussion... "Preliminary data on phylogenetic relationships of African species, inferred from partial *lsrDNA* sequences of five species representing all five African genera, showed that they form a monophyletic group (Fig. 62) ...". This is confusing, *Bothriocephalus acheilognathi*, while occurring in Africa, is a recent introduction. *Bothriocephalus claviceps* is dubiously recorded from Africa. So, the genus *Bothriocephalus* is not really an African genus and should not be included as such.

Paper V

"A new monozoic tapeworm, *Lobulovarium longiovatum* n. g., n. sp. ...".

There seem to be a lot of differences between the tree in Paper I based on *lsrDNA* + *ssrDNA*, and the one here (fig. 19), based only on *lsrDNA* sequences. I note, in particular, the different placements of *Caryophyllaeides*, but other genera are also in very different relative positions. How are these differences to be explained? Do they indicate a problem with molecular phylogeny of this group of tapeworms?

The standard of English is generally good throughout the thesis, and (as might be expected) especially good in the published papers. There were some difficulties in the first 35 pages. I will only mention a couple of them here. On page 24 the phrase "high sequence homology" appears. "Homology" is a word very frequently misused in molecular biology. "Similarity" should have been used here. Also in this section, "degenerative" and "degenerated" were used, when the word should be "degenerate".

P3 last paragraph; what does "analogously problematic" mean?

Some attention needs to be paid to verb tenses and to agreement between nouns and verbs.

Examiners are asked to "judge the scientific value of the thesis, novelty of the results and fulfillment of the objectives". I have no difficulties in stating that the thesis contains valuable and new material. I suppose that I have to say there is a problem in "fulfillment of the objectives", given that the original aim, to use SL *trans*-spliced genes as a tool to elucidate cestode evolution (page 14), could not be achieved. However, the candidate realised this in time to shift the emphasis of much of the work.

David Blair

David Blair: 8 June 2012.