



Prof. Dr. Hanno Schaefer
TU Munich, WZW
Dept. Ecology & Ecosystem
Management
Plant Biodiversity Research &
Herbarium TUM
Emil-Ramann-Str. 2
D-85354 Freising
GERMANY

phone: 0049-(0)8161-715884
email: hanno.schaefer@tum.de

To
Prof. A. Lebeda
Faculty of Science
Palack University Olomouc
Chair of habilitation committee for
Ing. Petr Smýkal, Ph.D.

Evaluation of the habilitation thesis
"Genetic diversity of genus *Pisum*, its exploitation and pea domestication"

Freising, Nov. 25, 2016

The thesis of Dr. Smýkal consists of four chapters summarizing 22 selected publications (eight of them with Dr. Smýkal as first author, four with Dr. Smýkal as senior author): the first chapter is a general introduction to the legume family, and more specifically the tribe Fabeae. A discussion of the use, taxonomy, and phylogeny of the five genera in that tribe is followed by an introduction to the pea genome and pea genomics studies. Chapter two focusses on genetic diversity of the genus *Pisum* with focus on *Pisum sativum*, both the wild populations as well as the material stored in germplasm collections worldwide. In the third chapter, the domestication history of pea is discussed, while the fourth chapter gives an overview of modern pea breeding. The thesis ends with a short conclusion and outlook section.

Chapter 1 is in general a good introduction and overview but there are some shortcomings regarding the discussion of the generic circumscriptions and the interpretation of the current literature. Contrary to what is stated on p. 15 and p. 19, *Lathyrus* in its current circumscription is not monophyletic. *Vicia* as it is understood today, also is not monophyletic. Instead, a monophyletic *Pisum* and monophyletic *Vavilovia* are both nested in *Lathyrus*, and this clade together with a monophyletic *Lens* is nested in *Vicia*. The tree shown in Fig. 2 (p. 13) is outdated. The statement on p. 15 that the relationship between *Pisum* and *Lathyrus* "is not completely clear" is also outdated. It would have been better to start the chapter with a more general discussion of the generic concepts and the suggestions how to revise them to obtain a natural classification, instead of following the outdated genus

concepts. Why discuss (in chapter 1.2.1.) a paraphyletic *Lathyrus*? Why discuss (in chapter 1.2.2) a paraphyletic *Vicia*? Why maintain a distinct genus *Vavilovia*, when it is deeply nested in *Lathyrus*? Discussion of morphological character evolution is not meaningful for paraphyletic groups. In my opinion, it would have been better to follow up the suggestion of Schaefer et al. 2012 (recently followed in the "Monographie des Leguminosae de France", P. Coulot & P. Rabaute, 2016), who suggest a recircumscription of the genera in the tribe based on natural groups: 1) *Ervum*, 2) *Ervilia*, 3) *Lathyrus* (including *Pisum* & *Vavilovia*, 4) *Vicia* s.str. (including *Lens*). This means, unfortunately, that we have to say goodbye to the nice name *Pisum sativum* and accept that the better name for the pea is *Lathyrus oleraceus* Lamarck (as mentioned on p. 24). Otherwise, I have only some minor remarks for the first chapter: on page 12, the "monotypic genus *Vavilovia formosa*" is mentioned, while of course this is the species. On page 16, I do not understand the difference between a wild ancestor of a crop (here *Vicia faba*) going extinct and being "domesticated entirely". In both cases, no wild population survives, so the taxon goes extinct.

For Chapter 2, the focus was shifted to diversity analyses of pea and a manuscript submitted to Molecular Ecology with Dr. Smýkal as first author. With a very impressive amount of work, he and his coauthors analysed almost 500 samples of wild peas and pea landraces looking at trnSG and ITS diversity but also at whole genomes. The latter was done with 150 samples using the DArTseq approach, which gave thousands of informative SNPs. The analyses gave the perhaps surprising result (at least to me), that the samples can be divided in seven distinct clusters with *P. abyssinicum* and *P. fulvum* clearly distinct but *Vavilovia* mixed in the *P. sativum*/*P. elatius* group. Maybe I misunderstand these results but if not, this is rather counter-intuitive and contradicts the morphological and ecological results as well as the earlier phylogenetic studies, which all see *Vavilovia* as a distinct lineage (which is, however, deeply nested in the *Lathyrus* clade). It would have been nice to discuss this issue a bit further. Is it a saturation problem or due to introgression? Or is it some sort of artefact caused by the Structure software? Maybe I overlooked it but I did not see any discussion trying to explain these results in the light of the very nice *Vavilovia* studies of Smýkal et al., which all point out the distinctness of that lineage.

In the third chapter, the focus lies on domestication of pea with discussion of a very exciting paper by Smýkal and colleagues reporting an ancient DNA analysis of Early Iron Age pea seeds from an archaeological site in Serbia. Several short DNA sequences could be analysed from these samples and allowed to place them in the phylogeny as sister to all modern pea landraces plus *P. humilis* and some *P. elatius*. This suggests that they are (as expected) very early cultivated peas. A second, also very nice paper on the structure of pea seeds and seed dormancy (published in Frontiers in Plant Sciences) is here also relevant. The various analyses by Dr. Smýkal of testa structure and biochemical principle of seed dormancy in pea are also remarkable. If the thick testa of wild pea allows years and maybe decades of survival in the soil, this should have quite interesting effects on genetic diversity of populations. The often small number of aboveground individuals could then be quite misleading because all the seeds waiting in the seed bank should actually also be included in genetic diversity analysis. I guess there are still many interesting angles to explore in future studies. My only minor criticism is that some parts of the chapter are a bit redundant and could have been more focussed. For example, on p. 75, the second half of the first paragraph seems to be identical to the end of the first paragraph on p. 67. Maybe this is just an editing error?

The last chapter discusses how to use the existing diversity in pea for breeding strategies and relates to two publications by Dr. Smýkal on breeding of varieties resistant to

the pea seed-borne mosaic virus. Main focus is breeding via introgression lines. More specifically, the idea is to take advantage of resistance genes in the wild *Pisum fulvum*, without losing the domestication traits in the *Pisum sativum* cultivar. My expertise in this field is limited, so I cannot comment much on these experiments. However, I get the impression that breeding via introgression lines is a very time-consuming process. Maybe the recent development of CrisprCas technique would allow a more concerted and much quicker approach? It would have been nice to mention these recent advances and novel options in the concluding part of the thesis. With just two paragraphs, it is really short and in my opinion suffers a bit from the lack of visionary statements. The soon to be released pea genome will of course be a massive step forward but what will be the big question that can be addressed using this genome? I think it will be much more than "just" going back to Gregor Mendel's efforts trying to link phenotype and genetic information.

In summary, the research achievements and publication record of Dr. Smýkal are impressive. With more than 50 papers in peer-reviewed journals and more than 1000 citations (according to his research gate webpage; Web of Science finds 43 papers with 678 citations since the year 2000, highest impact factor 4.81), he is clearly a very productive scientist and well-established in the legume research community. I would rank him as one of the leading authorities for the *Pisum* lineage, where he is not only an expert in systematics and (population) genetics/genomics but, as demonstrated by the broad range of his papers, his expertise also includes anatomy, physiology, ecology, modelling, bioinformatics, and modern plant breeding. I also note that many of his projects were (and still are) international collaborations including famous legume researchers from Australia, USA, UK, and Turkey. Travelling in many different countries, often with local expert researchers, he has gained a lot of important connections to field botanists. So, it is clear that Dr. Smýkal is not only an excellent scientist but also a leader who can bring together very different people and motivate them to try to solve complicated questions together. This is a crucial quality in modern Science, where the complexity of each field is way too high for a single individual. In order to master all the different cutting-edge approaches, one has to collaborate with many people. This is obviously one of the reasons for D. Smýkal's success and this is also, why I am convinced that his career is only just starting and that there will be many exciting discoveries from the Smýkal lab in the coming years.

Sincerely,

A handwritten signature in blue ink that reads "Hanno Schaefer". The script is cursive and fluid.

Prof. Dr. Hanno Schaefer