

DR. TANJA PYHÄJÄRVI (Orcid ID : 0000-0001-6958-5172)

MS. TIINA MARIA MATTILA (Orcid ID : 0000-0002-1298-7370)

Article type : Perspective

Title: New model species for Arctic-alpine plant molecular ecology

Running title: Arctic-alpine model species

Tanja Pyhäjärvi<sup>1,\*</sup> & Tiina M. Mattila<sup>2</sup>

1 Department of Forest Sciences, University of Helsinki, FIN-000 14 University of Helsinki, Finland

2 Department of Organismal Biology, Uppsala University, SE-752 36 Uppsala, Sweden

\*Correspondence: Tanja Pyhäjärvi Department of Forest Sciences, University of Helsinki, FIN-00014 University of Helsinki, Finland, [tanja.pyhajarvi@helsinki.fi](mailto:tanja.pyhajarvi@helsinki.fi)

Keywords: *Draba nivalis*, genome, adaptation, Arctic-alpine, linkage map, Brassicaceae

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/1755-0998.13335](https://doi.org/10.1111/1755-0998.13335)

This article is protected by copyright. All rights reserved

Arctic and alpine, high latitude and high elevation environments are one of the most stressful environments for species to inhabit. This harshness manifests itself in lower species richness in comparison to more southern vegetation zones (Francis & Currie, 2003). Furthermore, the climatic oscillations - past and predicted - have the most dramatic effect on these ecosystems. For example, in regions of continental ice sheets - the northernmost part of Western Europe and North America - the Arctic species assemblages are no older than a few thousands of years, which is a relatively short period from an evolutionary perspective. Although similar environments may have existed further south during the Ice Age, allowing some preadaptation for the Arctic species, the current habitat is a unique combination of environmental factors such as the climate, soil, bedrock, and photoperiod. Hence, understanding the evolutionary forces shaping Arctic-alpine species will be important for predicting these vulnerable environments' population viability and adaptive potential in the future. In this issue of *Molecular Ecology Resources*, Nowak et al. (202#) present extensive genome-wide resources for an Arctic-alpine plant *Draba nivalis*. This adds a valuable new member into the cabbage family models for evolutionary genetics and adaptation studies, to accompany e.g., *Arabidopsis* (Hu et al., 2011; The *Arabidopsis* Genome Initiative, 2000), *Arabis* (Willing et al., 2015) and *Capsella* (Slotte et al., 2013). A whole new avenue will open up for molecular ecological studies not only for *D. nivalis*, but the whole large *Draba* genus with its diverse ecological and evolutionary characteristics.

*D. nivalis* is a self-compatible perennial plant with a Holarctic distribution (Figure 1). It belongs to a large genus of *Draba*, of whose 390 species typically inhabit Arctic-alpine, marginal, low-competition environments (Jordon-Thaden, Al-Shehbaz, & Koch, 2013). For evolutionary genetic studies, the genus is biologically diverse with its large number of species containing variation in e.g., ploidy, mating system, and extent of distribution (Jordon-Thaden et al., 2013). *D. nivalis* itself is known to contain cryptic species (Grundt, Kjølner, Borgen, Rieseberg, & Brochmann, 2006), thus facilitating speciation studies. Multiple evolutionary lineages allow joint analysis of adaptation, colonization and speciation processes and the new genomic resources will help to pinpoint the underlying molecular mechanisms.

The genomic resources published by Nowak et al. 202# include genome assembly (302 Mb assembled into eight chromosomes), gene and transposable element annotation, linkage map, and information on polymorphic sites in the genome, to mention the focal resources. It is worth noting that a chromosome level assembly is still not simple to achieve, but greatly improves

understanding the evolutionary forces shaping diversity that include linked selection and unequal distribution of genomic elements across chromosomes. In Nowak et al. 202# a combination of Illumina short-reads, Nanopore long-reads, proximity ligation data, and linkage map from a cross between Norwegian and Alaskan strains resulted in a genome assembly of which contiguity and completeness is comparable with many Brassicaceae model species assemblies (Table 1). Particularly important steps for improving the assembly contiguity were the proximity ligation step (increased the scaffold N50 from 30 Kb to 2.9 Mb and decreased the scaffold L50 from 2,663 to 30) and anchoring the scaffolds based on the linkage map (increased the scaffold N50 from 4.4 Mb to 31 Mb).

In addition to the genomic resources, Nowak et al. 202# utilized three different approaches to investigate adaptive evolution signatures from the genome sequence of *D. nivalis*: 1) comparison of Pfam abundances across species, 2) gene family composition analysis looking for significant expansions and contractions in Draba lineage, and 3) identification of sites that have been subjects of positive selection in the branch leading to *D. nivalis*. These approaches were followed by gene ontology enrichment of the putatively selected loci to reveal processes, functions and cellular components specifically affected by positive selection. Ontology enrichment revealed multiple gene attributes that had either expanded in Draba lineage or experienced codon level selection pressure or both and deserve further empirical studies. For example, the enriched oxidation-reduction homeostasis and photoperiod related genes have indication of being especially important in Arctic-alpine environments. Interestingly, also meiosis related genes were identified as positively selected and significantly expanded in the Draba lineage, adding evidence that meiotic processes, despite being part of the essential machinery of the sexual reproduction, are sensitive to environmental conditions such as temperature (Bomblies, Higgins, & Yant, 2015).

All three methods of identifying adaptation signature applied to the *D. nivalis* genome are based on comparison to other, non-Draba species and thus covers evolutionary time of the entire Draba branch. This means likely over ten million years of evolutionary history, much older than the existence of *D. nivalis* as a species and its history as an Arctic species. Investigating the species colonization and population split history in more detail will refine the existing framework of its adaptation. Population re-sequencing of different groups covering different parts of its distribution range will be technically straight-forward with the help of reference genome published by Nowak et al. 202#. Re-sequencing data will allow identification of loci whose variants were favored during the Arctic colonization, whether they are shared across the species distribution (putative preadaptation) or repeatedly adapted to similar environments independently. The utility of the *D.*

*nivalis* reference genome expands to other *Draba* species and due to their distribution in multiple continents and different Arctic-alpine regions, allows studies of further patterns of convergent adaptation.

Another pivotal resource for selection studies is the information on recombination rate variation along the genome now available for *D. nivalis*. For example, the extent of linked selection depends on recombination rate variation which is important to take into account in population genetics selection inference (Charlesworth, Morgan, & Charlesworth, 1993). Furthermore, variance in measures of differentiation, such as  $F_{ST}$ , was recently found to be correlated with recombination rate even in the absence of selection (Booker, Yeaman, & Whitlock, 2020). Hence, incorporating recombination rate as a covariate in selection scans will decrease the false positive and false negative detections in regions of low and high recombination rate, respectively. The genomic resources for *D. nivalis* will facilitate comprehensive, less-biased investigations of adaptation genomics in this genus.

Unfortunately, Arctic-alpine vascular plants, including *D. nivalis* are predicted to experience serious range-contractions and even extinctions (e.g. Niskanen, Niittynen, Aalto, Väre, & Luoto, 2019). Let's make sure that *Draba* spp. will not become model species of extinction genomics.

### Figure Legends

Figure 1. A. Species distribution map of *Draba nivalis*. Redrawn based on the distribution map from Den virtuella floran (<http://linnaeus.nrm.se/flora/>, accessed November 6, 2020). The basemap was obtained from Wikimedia Commons ([https://commons.wikimedia.org/wiki/File:Blankmap-ao-090N-north\\_pole.png](https://commons.wikimedia.org/wiki/File:Blankmap-ao-090N-north_pole.png), accessed November 10, 2020) B. *Draba nivalis* growing on a rocky slope in Northern Sweden (Photo: A. Lovisa S. Gustafsson).

Bomblies, K., Higgins, J. D., & Yant, L. (2015). Meiosis evolves: adaptation to external and internal environments. *The New Phytologist*, 208(2), 306–323.

Booker, T. R., Yeaman, S., & Whitlock, M. C. (2020). Variation in recombination rate affects detection of outliers in genome scans under neutrality. *Molecular Ecology*, 29(22), 4274–4279.

- Charlesworth, B., Morgan, M. T., & Charlesworth, D. (1993). The effect of deleterious mutations on neutral molecular variation. *Genetics*, 134, 1289-1303.
- Cheng, C. Y., Krishnakumar, V., Chan, A. P., Thibaud-Nissen, F., Schobel, S., & Town, C. D. (2017). Araport11: a complete reannotation of the Arabidopsis thaliana reference genome. *The Plant Journal*, 89(4), 789-804
- Francis, A. P., & Currie, D. J. (2003). A Globally Consistent Richness-Climate Relationship for Angiosperms. *The American Naturalist*, Vol. 161, pp. 523–536. doi: 10.1086/368223
- Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., ... Rokhsar, D. S. (2012). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Research*, 40(Database issue), D1178–D1186.
- Grundt, H. H., Kjølner, S., Borgen, L., Rieseberg, L. H., & Brochmann, C. (2006). High biological species diversity in the arctic flora. *Proceedings of the National Academy of Sciences of the United States of America*, 103(4), 972–975.
- Hu, T. T., Pattyn, P., Bakker, E. G., Cao, J., Cheng, J.-F., Clark, R. M., ... Guo, Y.-L. (2011). The Arabidopsis lyrata genome sequence and the basis of rapid genome size change. *Nature Genetics*, 43(5), 476–481.
- Jordon-Thaden, I. E., Al-Shehbaz, I. A., & Koch, M. A. (2013). Species richness of the globally distributed, arctic–alpine genus *Draba* L. (Brassicaceae). *Alpine Botany*, Vol. 123, pp. 97–106. doi: 10.1007/s00035-013-0120-9
- Lyons, E., & Freeling, M. (2008). How to usefully compare homologous plant genes and chromosomes as DNA sequences. *The Plant Journal*, Vol. 53, pp. 661–673. doi: 10.1111/j.1365-313x.2007.03326.x
- Nguyen, T. P., Mühlich, C., Mohammadin, S., van den Bergh, E., Platts, A. E., Haas, F. B., ... & Schranz, M. E. (2019). Genome improvement and genetic map construction for *Aethionema arabicum*, the first divergent branch in the Brassicaceae family. *G3: Genes, Genomes, Genetics*, 9(11), 3521-3530.
- Niskanen, A. K. J., Niittynen, P., Aalto, J., Väre, H., & Luoto, M. (2019). Lost at high latitudes: Arctic and endemic plants under threat as climate warms. *Diversity and Distributions*, Vol. 25, pp. 809–821. doi: 10.1111/ddi.12889
- Nowak, M. D., Birkeland, S., Mandáková, T., Choudhury, R. R., Guo, X., Gustafsson, A. L. S., ... Brochmann, C. (202x). The genome of *Draba nivalis* shows signatures of adaptation to the extreme environmental stresses of the Arctic. *Molecular Ecology Resources*. Vol. X, pp. xxx–xxx doi: 10.1111/1755-0998.13280
- Rawat, V., Abdelsamad, A., Pietzenuk, B., Seymour, D. K., Koenig, D., Weigel, D., ... & Schneeberger, K. (2015). Improving the annotation of Arabidopsis lyrata using RNA-Seq

data. PloS one, 10(9), e0137391.

Slotte, T., Hazzouri, K. M., Ågren, J. A., Koenig, D., Maumus, F., Guo, Y.-L., ... Wright, S. I. (2013). The *Capsella rubella* genome and the genomic consequences of rapid mating system evolution. *Nature Genetics*, 45(7), 831–835.

The Arabidopsis Genome Initiative. (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*, Vol. 408, pp. 796–815. doi: 10.1038/35048692

Willing, E.-M., Rawat, V., Mandáková, T., Maumus, F., James, G. V., Nordström, K. J. V., ... Schneeberger, K. (2015). Genome expansion of *Arabis alpina* linked with retrotransposition and reduced symmetric DNA methylation. *Nature Plants*, 1, 14023.

Yang, R., Jarvis, D. J., Chen, H., Beilstein, M., Grimwood, J., Jenkins, J., ... & Wang, X. (2013). The reference genome of the halophytic plant *Eutrema salsugineum*. *Frontiers in plant science*, 4, 46.

Table 1. Genome assembly summary statistics for *Draba nivalis* and selected Brassicaceae model species for comparison. The statistics were calculated with assembly-stats (<https://github.com/sanger-pathogens/assembly-stats>). The number of gene models is the count of “gene” annotations on the gff annotation files. The reference genomes

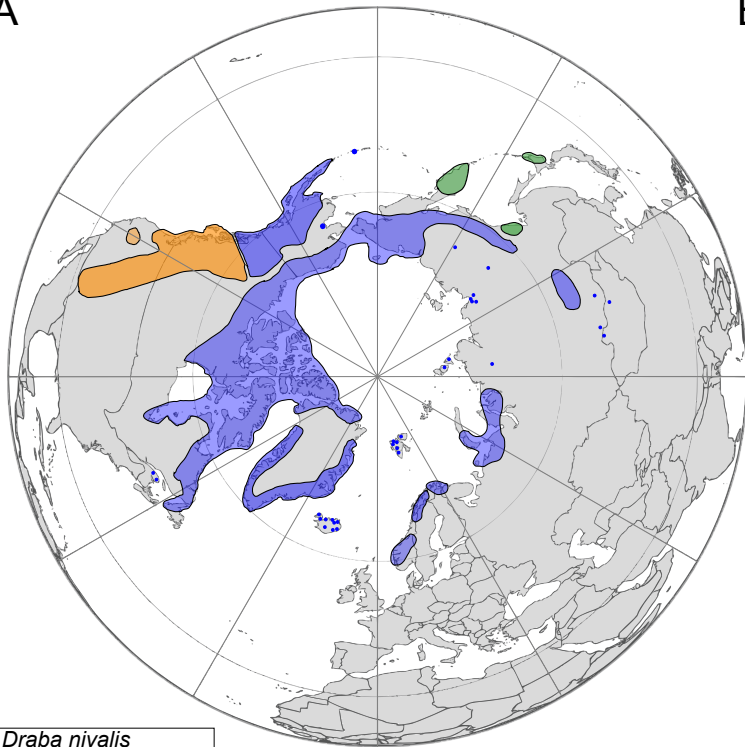
Species	Reference genome source & version	Special properties	Assembly size (Mb)	Number of scaffolds†	Unknown bases (Mb)	Number of gaps	N50 (Mb)	L50	Gene models	Original reference
<i>Draba nivalis</i>	Nowak et al. 202#	Arctic-alpine distribution	302	6648	20.67	21877	34.19	5	33537	Nowak et al. 202#
<i>Aethionema arabicum</i>	GoGe, v5_map_achored, genome id: 34234	Representative of the basal Brassicaceae lineage	196	2990‡	25.79	11261	10.14	9	23333	Nguyen et al. 2019
<i>Arabidopsis lyrata</i>	Phytozome V13, v2.1	Local adaptation, hybrid incompatibility	207	695	22.89	2953	24.46	4	31073	Hu et al. 2011; Rawat et al. 2015
<i>Arabidopsis thaliana</i>	Phytozome V13, TAIR10 & Araport11	Molecular genetics model species	120	7§	0.19	165	23.46	3	27655	The Arabidopsis Genome Initiative 2000; Cheng et al. 2016
<i>Arabis alpina</i>	<a href="http://www.arabis-alpina.org/V5">http://www.arabis-alpina.org/V5</a>	Model for perenniality	337	926	11.28	1856	36.6	4	34220	Willing et al. 2015
<i>Capsella grandiflora</i>	Phytozome V13, v1.1	Self-incompatible relative of <i>C. rubella</i> , large effective population size	105	4997	11.59	12289	0.11	223	24805	Slotte et al. 2013
<i>Capsella rubella</i>	Phytozome V12_unrestricted, 183_v1.0	Recent shift to selfing	135	853	4.76	8822	15.06	4	26521	Slotte et al. 2013
<i>Eutrema salsugineum</i>	Phytozome V13, v.1.0	Salt tolerance	243	639	4.65	2872	13.44	8	26351	Yang et al. 2013

†The assemblies did not include scaffolds shorter than 1000 bp except for *A. arabicum*

‡Shortest scaffold 886 bp

§Only chromosomal scaffolds

A



B



*Draba nivalis*  
subsp. *nivalis*  
subsp. *lonchocarpa*  
*D. kamtschatica*