ISSN 0305-0270
Journal of Biogeography

Supplementary Material for

Revisiting a speciation classic: Comparative analyses support sharp but leaky phylogeographic transitions in *Bombina*

Christophe Dufresnes, Tomasz Suchan, Nazar A. Smirnov, Mathieu Denoël, Juriy M. Rosanov, Spartak N. Litvinchuk

Table of Content

Appendix I

Localities analyzed in this study

Table S1: Localities included in this study, with indications on their labelling on Fig. 3 (N°), field diagnosis (Field), latitude (N) and longitude (E), mitochondrial lineages (mtDNA) – see below – nuclear ancestry based on RAD analyses (Chernivtsi, Bochnia) or allozymes (Transcarpathian) – noted "mix" if below 0.8 to either species – and genome size mean and standard deviation (SD) per populations. MtDNA labels as follow. *B. bombina*: B.N (northern lineage), B.S (southern lineage); *B. variegata*: V.C.W / V.C.E (western / eastern Carpathian lineages), V.W (western lineage), V.B (Balkan lineage). ZISP: specimen vouchers from the Zoological Institute of Russian Academy of Sciences, St. Petersburg.

						mtDNA		nuclear		genome size			
Locality	**N°**	**Field**	**N**	**E**	**n**	**n**	**clade**	**n**	**ancestry**	**n**	**mean**	**SD**	**ZISP**
<u>Chernivtsi</u>													
Bahna	1	*variegata*	48.23	25.20	7	-	-	7	*variegata*	-	-	-	
Valia Kuzmyna	2	*variegata?*	48.15	25.99	12	12	V.C.E, V.C.W	12	*variegata*	-	-	-	
Krupianske	3	*variegata?*	48.17	26.04	10	10	V.C.E	10	*variegata*	-	-	-	
Mala Buda	4	*bombina*	48.18	26.11	1	1	B.S	1	*bombina*	-	-	-	
Zavoloka	5	*bombina*	48.25	25.89	12	12	B.S	12	*bombina*	-	-	-	
Tsetsyno	6	*variegata*	48.31	25.85	6	6	V.C.E, V.C.W	6	*variegata*	-	-	-	
Shypyntsi	7	*bombina*	48.37	25.74	15	15	B.S, B.N	15	*bombina*	-	-	-	
Sadhora Forest	8	mix	48.33	26.00	6	6	B.S, V.C.E	6	mix	-	-	-	
Chornivka	9	*bombina*	48.42	26.04	9	9	B.S, B.N	9	*bombina*	-	-	-	
Malyi Kuchuriv	10	*bombina*	48.46	25.92	2	2	B.S	2	*bombina*	-	-	-	
<u>Bochnia</u>													
Drwinia	1	*bombina*	50.10	20.43	11	11	B.N	8	*bombina*	-	-	-	
Lipówka	2	*bombina*	50.08	20.40	14	14	B.N	13	*bombina*	-	-	-	
Cikowice	3	*variegata*	49.97	20.39	9	9	V.C.W	8	*variegata*	-	-	-	
Doluszycki Forest 1	4	*variegata*	49.95	20.40	6	6	V.C.W	4	*variegata*	-	-	-	
Doluszycki Forest 2	5	*variegata*	49.94	20.42	3	3	V.C.W	3	*variegata*	-	-	-	
Doluszycki Forest 3	6	*variegata*	49.93	20.42	9	9	V.C.W	9	*variegata*	-	-	-	
Doluszycki Forest 4	7	*variegata*	49.92	20.42	8	8	V.C.W	6	*variegata*	-	-	-	
Chrostowa	8	*variegata*	49.90	20.30	5	5	V.C.W	5	*variegata*	-	-	-	

Jaroszówka	9	*variegata*	49.89	20.28	5	5	V.C.W	5	*variegata*	-	-	-	
Sobolów	10	*variegata*	49.90	20.35	11	6	V.C.W	5	*variegata*	-	-	-	
Wichraż	11	*variegata*	49.88	20.37	3	3	V.C.W	3	*variegata*	-	-	-	
Lipnica Murowana	12	*variegata*	49.85	20.53	8	5	V.C.W	5	*variegata*	-	-	-	
Iwkowa	13	*variegata*	49.82	20.53	8	8	V.C.W	7	*variegata*	-	-	-	
Bytomsko	14	*variegata*	49.81	20.46	4	4	V.C.W	4	*variegata*	-	-	-	

Transcarpathian

Aeroport	-	*bombina*	48.64	22.26	18	-	-	15	*bombina*	18	21.4	0.10	5947, 6334, 7202
Andreevka	-	*variegata*	48.53	22.55	4	-	-	-	-	4	20.5	0.08	-
Batevo	-	*bombina*	48.37	22.40	34	-	-	22	*bombina*	34	21.5	0.06	5378, 5392, 5903, 6236, 7203
Belki	-	*variegata*	48.30	23.15	6	-	-	6	*variegata*	6	20.5	0.08	5904
Chernogora	-	*variegata*	47.95	24.31	8	-	-	-	-	8	20.3	0.11	-
Chinadievo	-	*variegata*	48.49	22.81	21	5	V.C.W	16	*variegata*	21	20.5	0.15	5385, 5899, 6529
Chop	-	*bombina*	48.43	22.23	36	-	-	34	*bombina*	36	21.6	0.06	5909, 7198
Chornyi Potok	-	*bombina*	48.20	22.92	1	-	-	1	*bombina*	1	21.7	-	5364
D'akovo	-	*bombina*	48.01	23.01	6	-	-	6	*bombina*	6	21.5	0.07	5377
Delovoe	-	*variegata*	47.93	24.18	13	-	-	12	*variegata*	13	20.2	0.09	5367
Dobroselie	-	*bombina*	48.16	22.77	4	-	-	4	*bombina*	4	21.5	0.06	5375
Drisina	-	*bombina*	48.33	22.65	4	3	B.S	4	*bombina*	4	21.6	0.04	5363
Glubokoe	-	*variegata*	48.55	22.42	8	-	-	8	mix	8	20.8	0.15	5370, 5854
Gudya	-	*bombina*	48.13	23.13	12	1	V.C.W	11	mix	12	21.3	0.12	5384, 5912
Kamenskoe	-	*bombina*	48.28	22.93	28	6	B.S, V.C.W	27	mix	28	21.4	0.23	5368, 5369, 5557, 6527
Kamenskoe-Seltse	-	*variegata*	48.28	22.97	3	-	-	-	-	3	21.0	0.10	5347
Karpaty	-	*variegata*	48.53	22.89	15	10	V.C.W	15	*variegata*	15	20.3	0.06	5897
Khmelnik	-	*bombina*	48.27	22.88	7	6	B.S	6	*bombina*	7	21.6	0.09	5365, 5366, 7213
Khust	-	*variegata*	48.16	23.30	4	1	V.C.W	1	*variegata*	4	20.3	0.05	5386, 5844
Klyucharka	-	*bombina*	48.42	22.65	1	1	B.S	-	-	1	21.4	-	6878
Kolchino	-	*variegata*	48.47	22.77	6	-	-	-	-	6	20.5	0.08	-
Kolochava	-	*variegata*	48.42	23.74	13	9	V.C.W	13	*variegata*	13	20.2	0.10	5898
Korolevo	-	*variegata*	48.17	23.15	8	-	-	4	mix	8	20.8	0.12	5374, 5911, 5912, 6879
Kriva	-	*variegata*	48.17	23.23	15	-	-	14	*variegata*	15	20.4	0.13	6215, 6525, 7215
Kvasovo	-	*bombina*	48.20	22.77	3	-	-	3	*bombina*	3	21.5	0.04	5376

Minai	-	*bombina*	48.58	22.28	17	-	-	16	*bombina*	17	21.6	0.09	7197
Mukachevo	-	*bombina*	48.44	22.67	36	2	B.S	34	*bombina*	36	21.6	0.08	5388, 5901, 6214, 6235, 6530
Nizhnee Solotvino	-	*bombina*	48.54	22.44	4	1	V.C.W	4	mix	4	21.5	0.26	5913, 5914
Paseka	-	*variegata*	48.53	22.92	4	-	-	4	*variegata*	4	20.3	0.05	5902
Seltse	-	*variegata*	48.29	22.98	2	-	-	-	-	2	20.3	0.13	-
Shalanka-1	-	mix	48.23	22.90	41	6	B.S, V.C.W	40	mix	41	21.1	0.45	5362, 5373, 6228, 6233, 6528
Shalanka-2	-	*bombina*	48.44	22.92	11	2	B.S	11	*bombina*	11	21.6	0.05	-
Sinevir	-	*variegata*	48.52	23.63	7	-	-	7	*variegata*	6	20.3	0.04	5379, 5806
Solotvino	-	*variegata*	47.95	23.90	42	-	-	27	*variegata*	42	20.3	0.06	5354, 5381, 5516, 5522, 5905, 7258
Svoboda	-	*bombina*	48.37	22.38	10	-	-	10	*bombina*	10	21.6	0.04	5380
Syanki	-	*variegata*	49.02	22.91	6	-	-	6	*variegata*	6	20.4	0.05	7210
Uzhgorod	-	*bombina*	48.63	22.28	6	-	-	-	-	6	21.4	0.11	6524
Velika Kopanya	-	*variegata*	48.19	23.12	3	-	-	3	mix	3	20.6	0.17	5900
Veryatsa	-	*variegata*	48.15	23.18	3	2	V.C.W	1	mix	3	20.7	0.08	5372, 5383
Vinogradov	-	*variegata*	48.13	23.08	1	-	-	-	-	1	20.9	-	-
Volosyanka	-	*variegata*	48.98	22.80	7	-	-	7	*variegata*	7	20.3	0.04	7214
Volovets	-	*variegata*	48.71	23.17	13	-	-	-	-	13	20.4	0.06	7257
Other areas													
Baranovichi	-	*bombina*	53.13	26.03	16	16	B.S, B.N	-	-	-	-	-	
Chervona Gusarivka	-	*bombina*	49.42	36.87	10	10	B.N	-	-	-	-	-	
Glinskoe	-	*bombina*	49.24	37.22	5	5	B.N	-	-	-	-	-	
Svyatogorsk	-	*bombina*	49.05	37.55	10	10	B.S, B.N	-	-	-	-	-	
Verbunovskaya dacha	-	*bombina*	49.42	36.88	6	6	B.N	-	-	-	-	-	
Baia de Arama 1	-	*variegata*	44.59	22.47	3	3	V.C.W	-	-	-	-	-	
Baia de Arama 2-3	-	*variegata*	45.00	22.47	3	3	V.C.W	-	-	-	-	-	
Svinita 1	-	*variegata?*	44.30	22.05	1	1	V.W	-	-	-	-	-	
Svinita 2	-	*variegata?*	44.30	22.06	3	3	V.B	-	-	-	-	-	

Appendix II

Ecological niche modelling and species distributions

Extended Methods

To reconstruct the ecological niche of each species, we gathered 3,570 and 769 localities with known occurrence of *B. bombina* and *B. variegata*, respectively, comprising our own records, museum collections and previously published data. This initial dataset was filtered to avoid spatial autocorrelation and duplication using NicheToolBox (Osorio-Olvera, Barve, Barve, Soberón, & Falconi, 2018), and retain localities at least 10 km (0.093°) apart (see details in Brown, 2014). The final dataset comprised 1461 and 508 records for *B. bombina* and *B. variegata*, respectively (Fig. S1), available in Geofile S1.

Models were built as follow. Altitude and 19 bioclimatic layers representative of the climatic data over ~1950–2000 were extracted from the WorldClim 1.4 database (http://www.worldclim.org). Ten additional layers were considered: the aridity index (Global Aridity and Potential Evapo-Transpiration; http://www.cgiar-csi.org/data/global-aridity-and-pet-database), the global percent of tree coverage (https://github.com/globalmaps/gm_ve_v1), and eight land cover variables (spatial homogeneity of global habitat, broadleaf forests, needleleaf forests, mixed forests, shrubs, barren, herbaceous and cultivated vegetation; https://www.earthenv.org/). To consider topography in the model, four landscape layers were calculated with QGIS: aspect, exposition, slope, and terrain roughness index (each with 30 arc seconds spatial resolution). All analyses were conducted under the WGS 84 projection with species-specific masks covering to the area of occurrence of these species: 36°N–60°N, 5°E–67°E for *B. bombina*; 35°N–63°N, 13°W–35°E for *B. variegata*.

To eliminate predictor collinearity before generating the model, we calculated Pearsons's correlation coefficients for all pairs of bioclimatic variables using ENMTools (Warren, Glor, & Turelli, 2010). For correlated pairs ($|r| > 0.75$), we excluded the variable that appeared the least biologically important for our target species. The resulting dataset contained eight bioclimatic variables for *B. bombina*: Bio 1 (annual mean temperature; °C×10), Bio 2 (mean diurnal range; °C×10), Bio 3 (isothermality; Bio2/Bio7×100), Bio 4 (temperature seasonality; CV ×100), Bio 8 (mean temperature of wettest quarter; °C×10), Bio 14 (precipitation of driest month; mm), Bio 15 (precipitation seasonality; CV), and Bio 19 (precipitation of coldest quarter; mm); and ten variables for *B. variegata*: Bio 2, Bio 3, Bio 4, Bio 5 (maximum temperature of warmest month; °C×10), Bio 6 (minimum temperature of coldest month; °C×10), Bio 8, Bio 12 (annual precipitation; mm), Bio 15, Bio 18 (precipitation of warmest quarter; mm), and Bio 19. We then applied a jackknife analysis to estimate the relative contributions of variables to the MaxEnt model.

We ran the MaxEnt program for ten replicates with 25% random test percentage testing. Model calibration consisted in the evaluation of models created with distinct regularization multipliers (0.5 to 6 at intervals of 0.5), feature classes (resulted from all combinations of linear, quadratic, product, threshold, and hinge response types), and from two different sets of layers. The first set consisted of all layers. The second was restricted to the four most valuable layers (Bio 1, Bio 3, Bio 8, and Bio 14 for *B. bombina*; and Bio 4, Bio 12, Bio 18, and altitude for *B. variegata*). The best parameter settings were selected considering statistical significance (partial ROC), predictive power (omission rates E = 5%), and complexity level (AICc) obtained using the R package kuenm (Cobos, Peterson, Barve, & Osorio-

Olvera, 2019). The ClogLog output format (ranging 0–1) was chosen for processing resulting maps (Phillips, Anderson, & Schapire, 2017).

To project the current ecological niches of these species on climate conditions during the Last Glacial Maximum (LGM), we applied two widely-used general atmospheric circulation models with species-specific mask and 2.5 arc minutes spatial resolutions: the Community Climate System Model (CCSM; http://www2.cesm.ucar.edu/) and the Model for Interdisciplinary Research on Climate (MIROC; Watanabe et al., 2011). Finally, to interpret the distributions in the light of landscapes during the LGM, territories covered by glaciers and seas were cropped from the resulting LGM layers (Becker, Verheul, Zickel, & Willmes, 2015; Zickel, 2016).

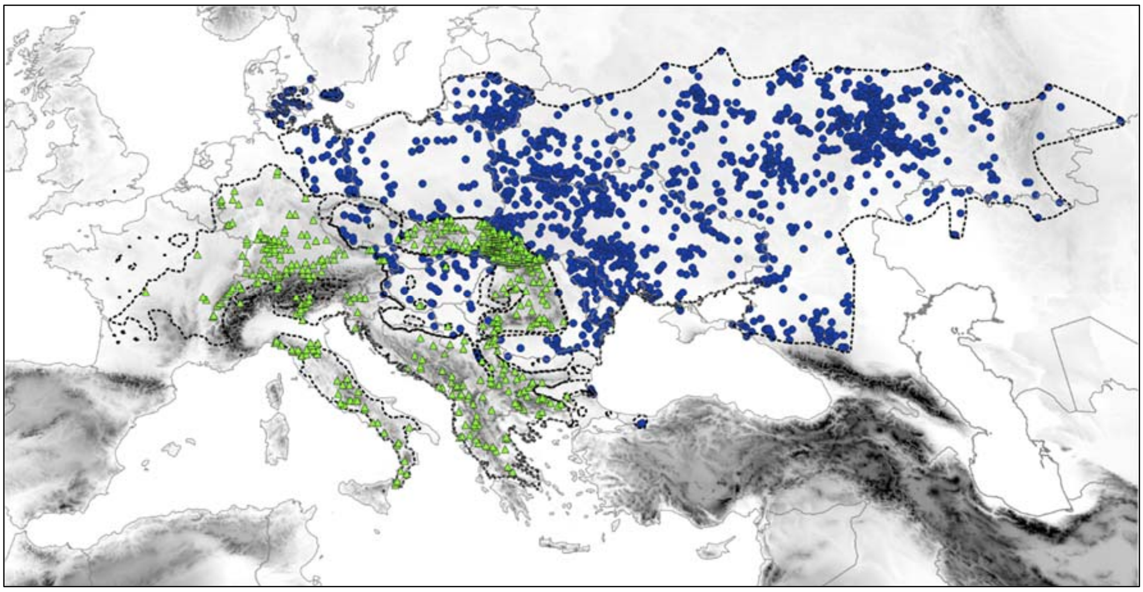

Fig. S1: Occurrence records used to build the ecological niche of *B. bombina* (blue circles) and *B. variegata* (green triangles), available in Geofile S1. Dashed lines show species distributions.

Extended Results

To calibrate the models, we assessed 696 replicates for each species. For *B. bombina*, all models were statistically significant when compared with a null model of random prediction, 144 (21%) met the omission criterion of 5%, and only a single met the AICc criteria and was statistically significant among models meeting both omission rate and AICc criteria. For *B. variegata*, 687 (99%) models were statistically significant when compared with a null model of random prediction, 68 (10%) significant models met the omission criterion of 5%, and only a single met the AICc criteria and was statistically significant among models meeting both omission rate and AICc criteria. Performance metrics for parameter settings used for creating these two final models for both species are given in Table S2. Both models were created with the first set of environmental layers. Regularization multipliers were 5.5 for *B. bombina* and 1.0 for *B. variegata*. Response types of feature classes were quadratic and hinge for *B. bombina* and linear, quadratic and product for *B. variegata*.

Table S2: Performance metrics used for selecting the final models and relative contribution (%) of variables in the niche of the two *Bombina* species.

Variable	***bombina***	***variegata***
Partial ROC	0	0
Omission rate 5%	0.048	0.042
AICc	61703	20711
Delta AICc	0	0
Annual mean temperature (Bio1)	6.5	–
Mean diurnal range (Bio 2)	0	0.9
Isothermality (Bio 3)	4.7	0.9
Temperature seasonality (Bio 4)	7.9	6.6
Maximum temperature of warmest month (Bio 5)	–	1.9
Minimum temperature of coldest month (Bio 6)	–	6.3
Mean temperature of wettest quarter (Bio 8)	12	0.5
Annual precipitation (Bio 12)	–	0.3
Precipitation of driest month (Bio 14)	50.8	–
Precipitation seasonality (Bio 15)	0.6	7.9
Precipitation of warmest quarter (Bio 18)	–	10.3
Precipitation of coldest quarter (Bio 19)	0.4	1.1
Broadleaf forest	0.1	12.3
Needleleaf forest	0.1	5.2
Mixed forest	0.1	5.2
Shrubs	0.1	4.7
Barren	1.6	3.9
Herbaceous vegetation	0	2.7
Cultivated vegetation	5.2	8
Altitude	4.4	2.2
Aridity index	0.4	1.1
Aspect	0.1	0.2
Exposition	0	0.1
Habitat homogeneity	0.8	0.1
Slope	0	2.8
Terrain roughness index	0	7.7
Tree coverage percent	4.2	7.1

Appendix III

Additional genetic analyses of the hybrid zones

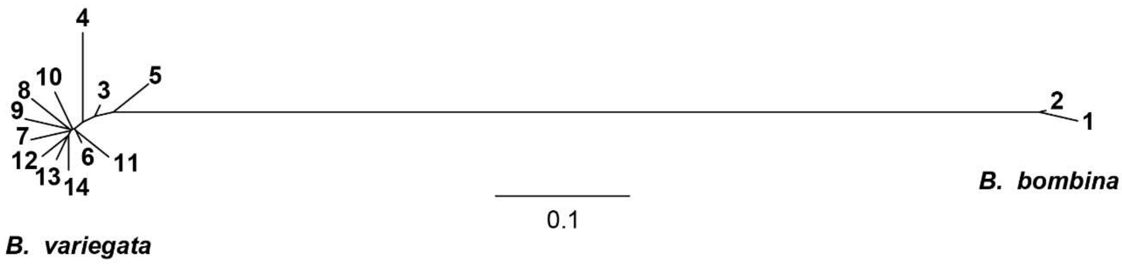

Fig. S2: Neighbor-Joining tree of pairwise genetic distance (F_{st}) on the initial Bochnia RAD-seq dataset (6,010 SNPs). Numbers indicate localities (see Fig. 3)

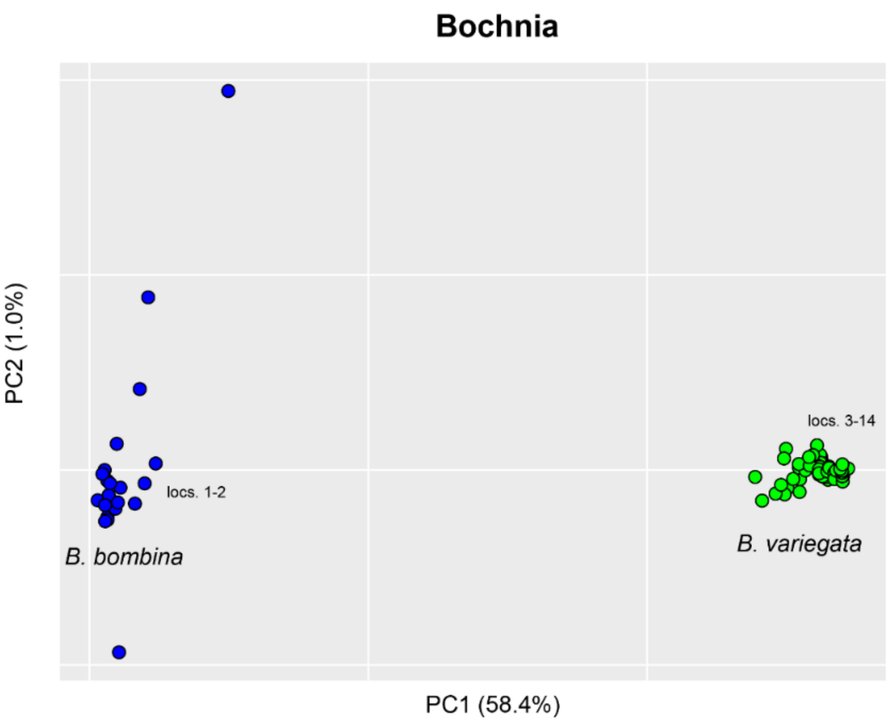

Fig. S3: PCA on the initial Bochnia RAD-seq dataset (6,010 SNPs). Loci building PC1 (species-diagnostic) were selected for the hybrid zone analyses

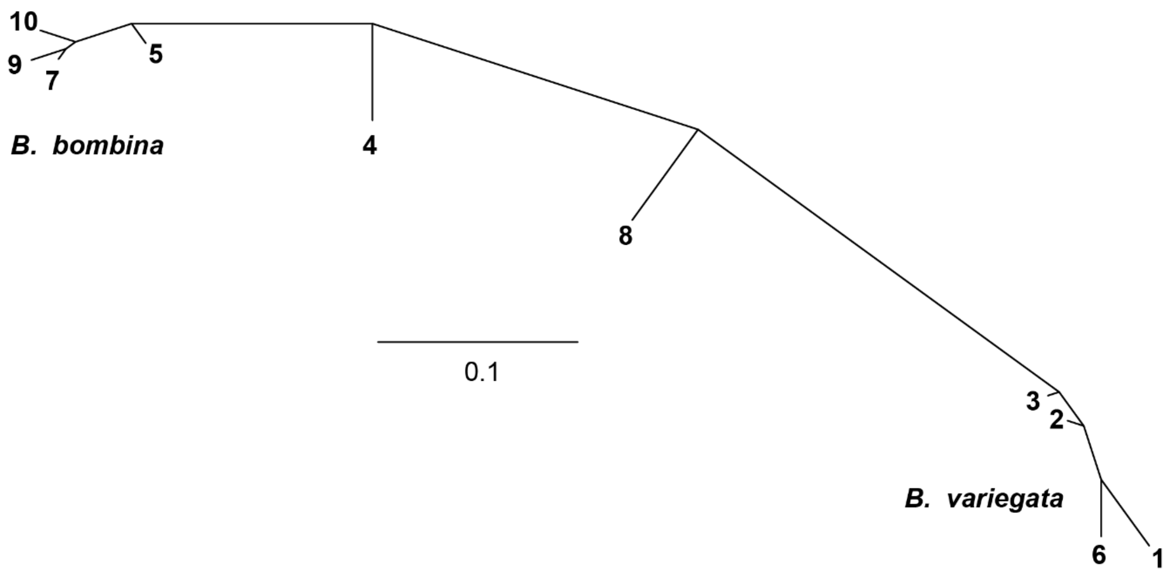

Fig. S4: Neighbor-Joining tree of pairwise genetic distance (F_{st}) on the initial Chernivtsi RAD-seq dataset (4,294 SNPs). Numbers indicate localities (see Fig. 3)

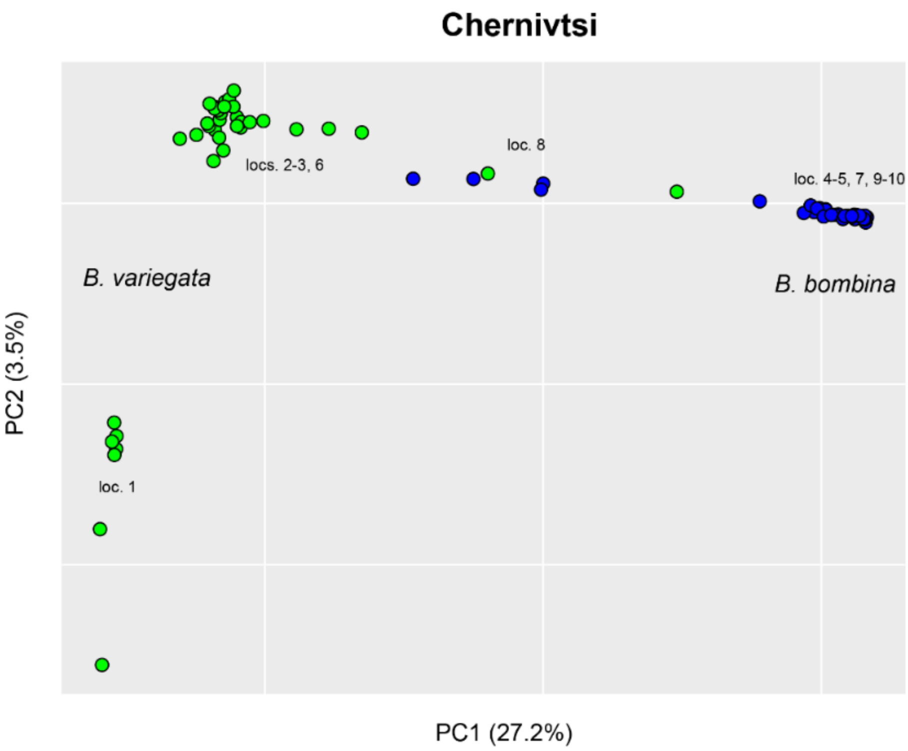

Fig. S5: PCA on the Chernivtsi RAD-seq dataset (4,294 SNPs). Loci building PC1 (species-diagnostic) were selected for the hybrid zone analyses

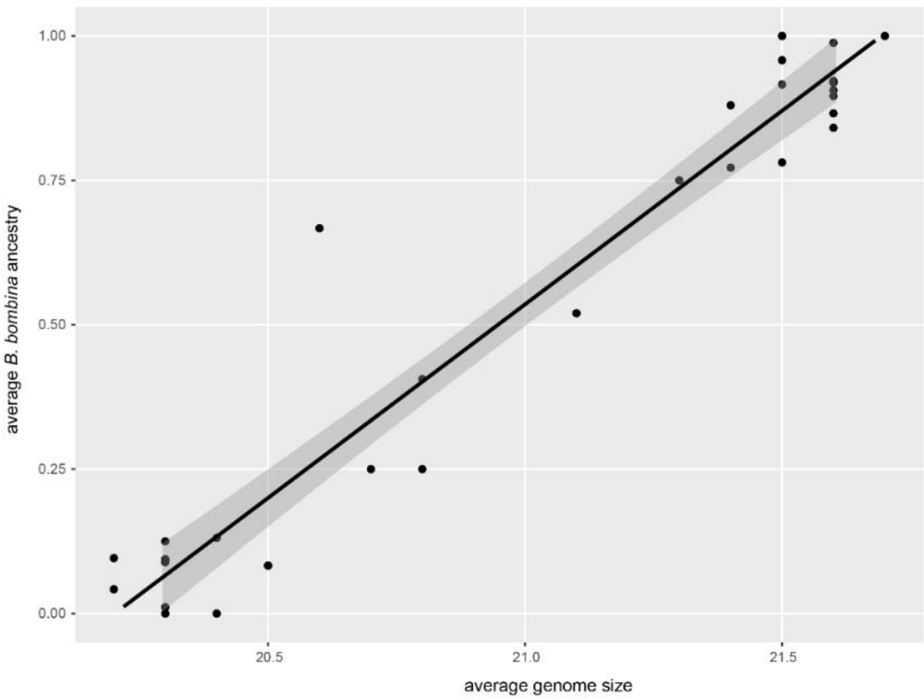

Fig. S6: Correlation between allozyme ancestry and average genome size in populations from the Transcarpathian hybrid zone.

The linear regression is shown ($R^2 = 0.93$, P <<<0.01).

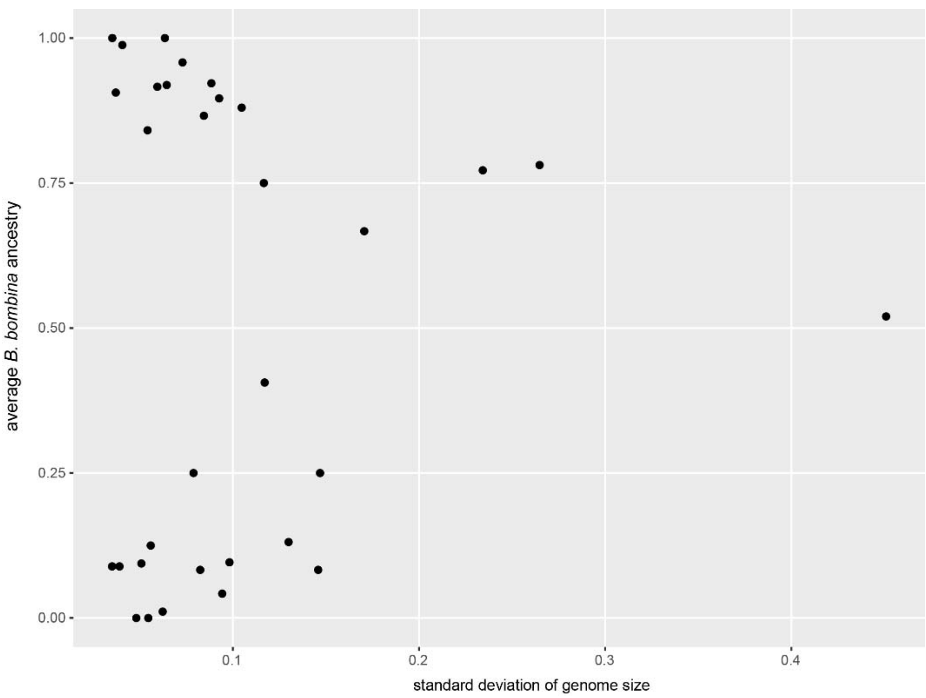

Fig. S7: Average ancestry of populations as a function of their variation in genome size in the Transcarpathian hybrid zone. Admixed populations are marked by higher variability in genome sizes.

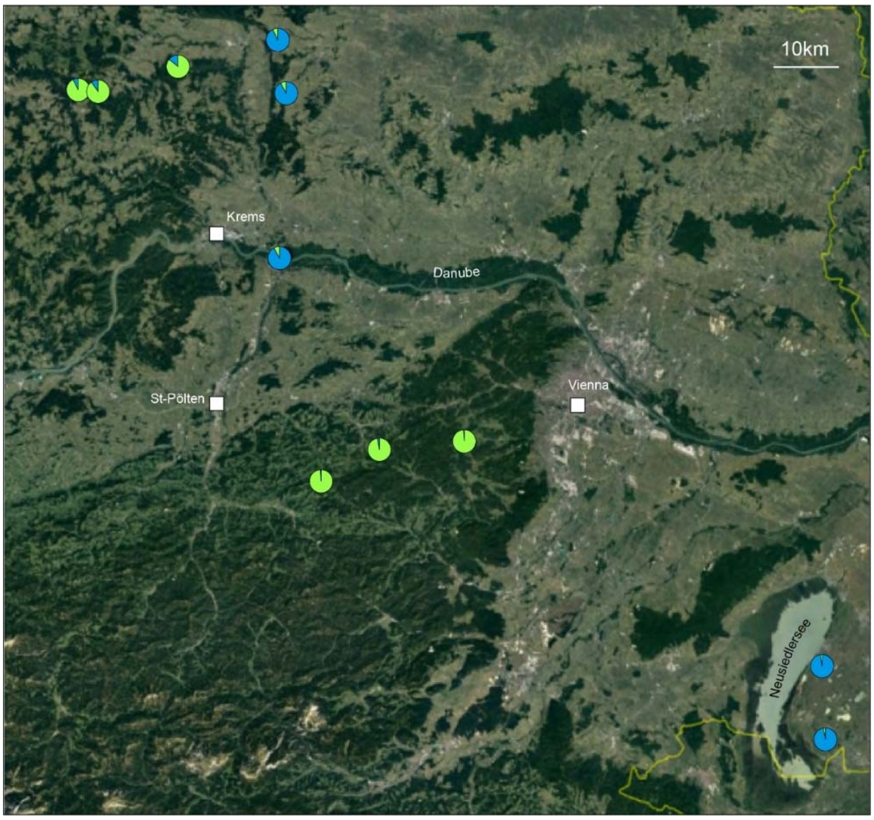

Fig. S8: Average population ancestry in northeastern Austria based on the allozyme data from Gollmann (1984)

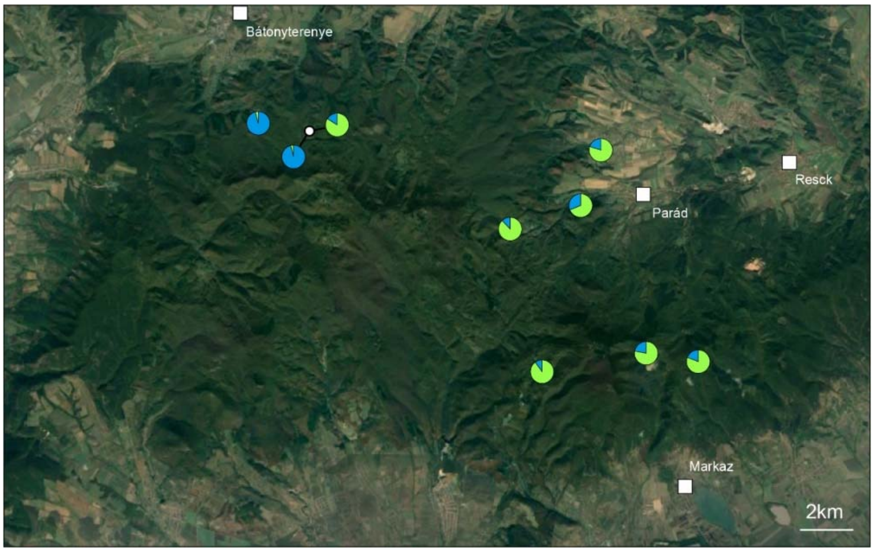

Fig. S9: Average population ancestry in the Mátra Mountains, based on the allozyme data from Gollmann (1987)

Literature cited

Becker, D., Verheul, J., Zickel, M., & Willmes, C. (2015). *LGM paleoenvironment of Europe – Map. CRC806-Database.* Available at https://crc806db.uni-koeln.de/dataset/show/lgm-paleoenvironment-of-europe--map1449850675/ (accessed 29 January 2020).

Brown, J. L. (2014). SDM toolbox: a python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. *Methods in Ecology and Evolution*, 5, 694–700. doi: 10.1111/2041-210X.12200.

Brown, J. L., Hill, D. J., Dolan, A. M., Carnaval, A. C., & Haywood, A. M. (2018). PaleoClim, high spatial resolution paleoclimate surfaces for global land areas. *Scientific Data*, 5, 180254. doi: 10.1038/sdata.2018.254.

Cobos, M. E., Peterson, A. T., Barve, N., Osorio-Olvera, L. (2019). kuenm: an R package for detailed development of ecological niche models using Maxent. *PeerJ*, 7, e6281. doi: 10.7717/peerj.6281.

Gollmann, G. (1984). Allozymic and morphological variation in the hybrid zone between *Bombina bombina* and *Bombina variegata* (Anura, Discoglossidae) in northeastern Austria. *Journal of Zoological Systematics and Evolutionary Research*, 22, 51–64. doi: 10.1111/j.1439-0469.1984.tb00562.x

Gollmann, G. (1987). *Bombina bombina* and *Bombina variegata* in the Mátra mountains (Hungary): New data on distribution and hybridization (Amphibia, Anura, Discoglossidae). *Amphibia-Reptilia*, 8, 213–224.

Osorio-Olvera, L., Barve, V., Barve, N., Soberón, J., Falconi, M. (2018). *ntbox: From getting biodiversity data to evaluating species distribution models in a friendly GUI environment.* R package version 0.2.5.4. Available at https://github.com/luismurao/ntbox

Phillips, S. J., Anderson, R. P., Schapire, R. E. (2006). Maximum entropy modelling of species geographic distributions. *Ecological Modelling*, 190, 231–259. doi: 10.1016/j.ecolmodel.2005.03.026.

Phillips, S. J., Anderson, R. P., Dudík, M., Schapire, R. E., & Blair, M. E. (2017). Opening the black box: an open-source release of Maxent. *Ecography*, 40, 887–893. doi: 10.1111/ecog.03049.

Warren, D. L., Glor, R. E., & Turelli, M. (2010) ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography*, 33, 607–611. doi: 10.1111/j.1600-0587.2009.06142.x.

Zickel, M. (2016) *Paleocoastline 130 m below mean sea level.* Available at https://crc806db.uni-koeln.de/layer/show/324/