Supplementary Data

Environmental differences are correlated with the distribution pattern of cytotypes in *Veronica* subsection *Pentasepalae* at a broad scale

Blanca M. Rojas-Andrés[1,2,3,*,†], Nélida Padilla-García[1,2,*,†], Manuel de Pedro[1,4], Noemí López-González[1,2], Luis Delgado[1], Dirk C. Albach[5], Mariana Castro[6], Sílvia Castro[6], João Loureiro[6], M. Montserrat Martínez-Ortega[1,2]

[1]*Departamento de Botánica, University of Salamanca, E-37007 Salamanca, Spain,* [2]*Biobanco de ADN Vegetal, University of Salamanca, Edificio Multiusos I+D+i, Calle Espejo s/n, 37007 Salamanca, Spain,* [3]*Department of Molecular Evolution and Plant Systematics & Herbarium (LZ), Institute of Biology, Leipzig University, Leipzig, Germany,* [4]*CREAF, E-08193 Bellaterra (Cerdanyola del Vallès), Spain,* [5]*Institute of Biology and Environmental Sciences (IBU), AG Plant Biodiversity and Evolution, Carl von Ossietzky Universität Oldenburg, Oldenburg 26111, Germany,* [6]*CFE, Centre for Functional Ecology, Department of Life Sciences, University of Coimbra, Calçada Martim de Freitas s/n, 3000-456 Coimbra, Portugal.*
†*These authors contributed equally to this work.*
For correspondence. E-mail rojas_andres@uni-leipzig.de or nelidapg@gmail.com

The following Supporting Information is available for this article:

Fig. S1. Observed and expected genome size values in *Veronica* subsect. *Pentasepalae*.

Fig. S2. Box plots of the 12 quantitative environmental variables that are significantly associated with the distribution of cytotypes.

Fig. S3. Relative frequencies of cytotypes in relation to the three most common vegetation types and the three most common biogeographical regions.

Fig. S4. Distribution of cytotypes in relation to vegetation type, and biogeographical regions.

Table S1 (*provided as a separate excel file*). Details of individuals of *Veronica* subsect. *Pentasepalae* included in this study, including population code, ploidy level, genome size values (1Cx, 1C and 2C-value), sample CV, number of measurements per individual, collector number, country, region, locality, geographical coordinates, altitude, date of collection, collectors and herbarium number. Ploidy levels marked with an asterisk indicate that chromosome counts were performed for this population. Mixed-ploidy populations are indicated with a grey shaded population code.

Table S2 (*provided as a separate excel file*). Mean and standard deviation of genome sizes for each ploidy (only samples with CV < 5%).

Table S3. Environmental variables used in univariate and multivariate analyses.

Table S4. Associations between ploidy levels and environmental variables in populations of *Veronica* subsect. *Pentasepalae* found by the univariate analyses.
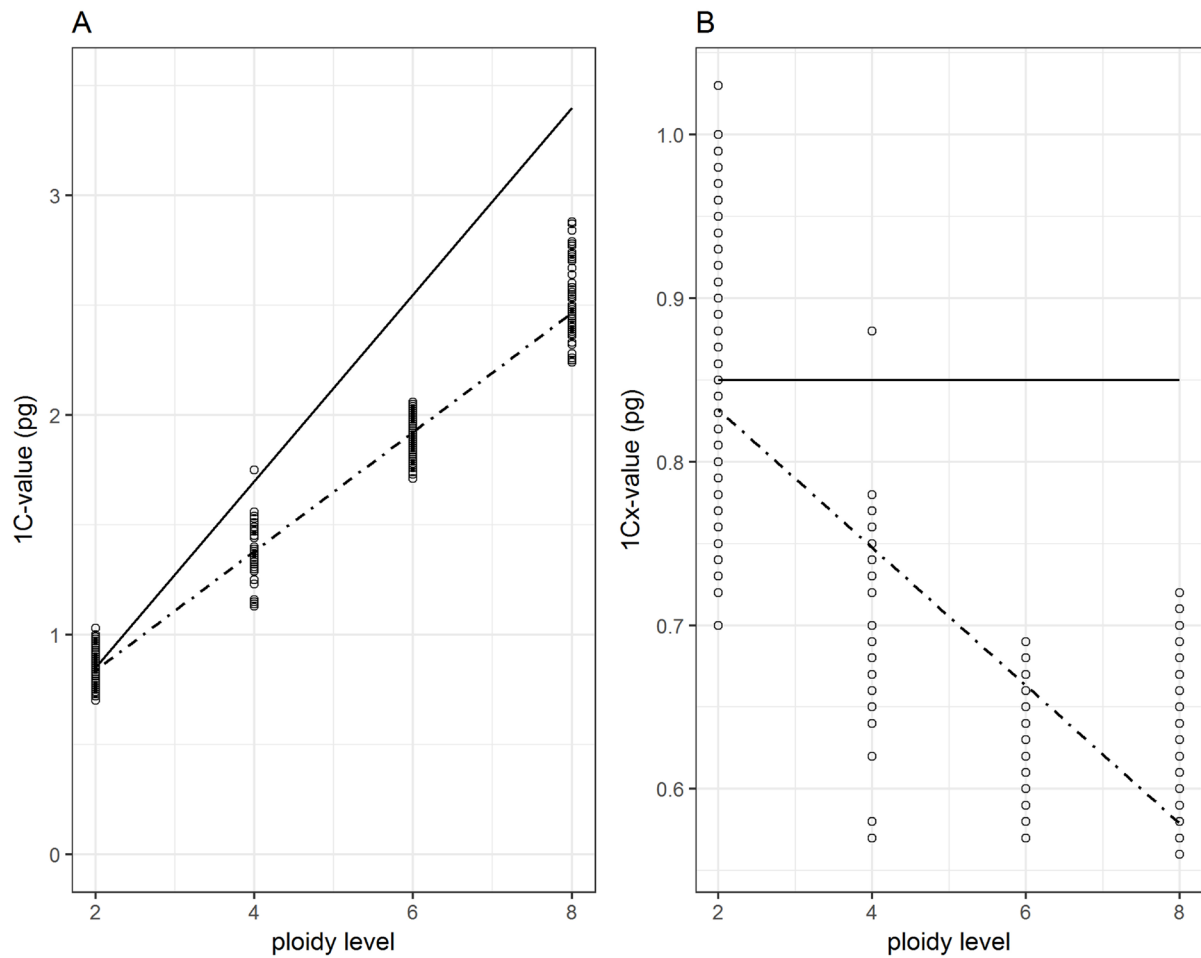
Fig. S1. Observed and expected 1C-values (A) and 1Cx-values (B) in the four groups of ploidy levels (2*x*, 4*x*, 6*x*, and 8*x*) found in *Veronica* subsect. *Pentasepalae* (only samples with CV < 5%). The solid lines indicate the correlation of mean observed values for different groups and the dot-dashed lines represent expected values based on mean observed diploid values. The upper outliers in the tetraploid group correspond to individuals of northwestern Spain (population 129; see discussion).
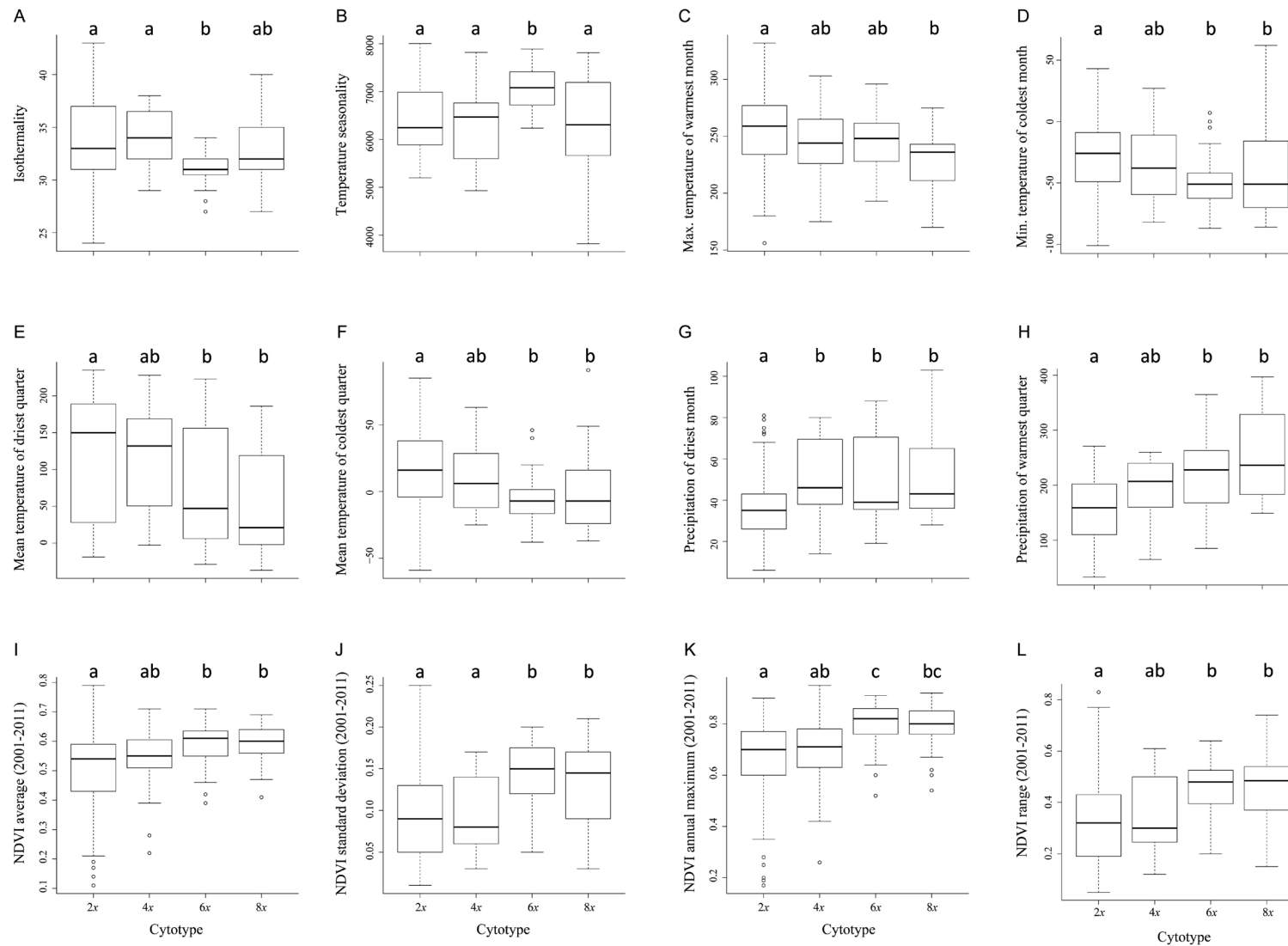
Fig. S2. Box plots of the 12 quantitative environmental variables that are significantly associated with the distribution of cytotypes. Significant differences between ploidy levels were identified by the Dunn's test ($P < 0.01$) and are indicated by different letters above the box plots.
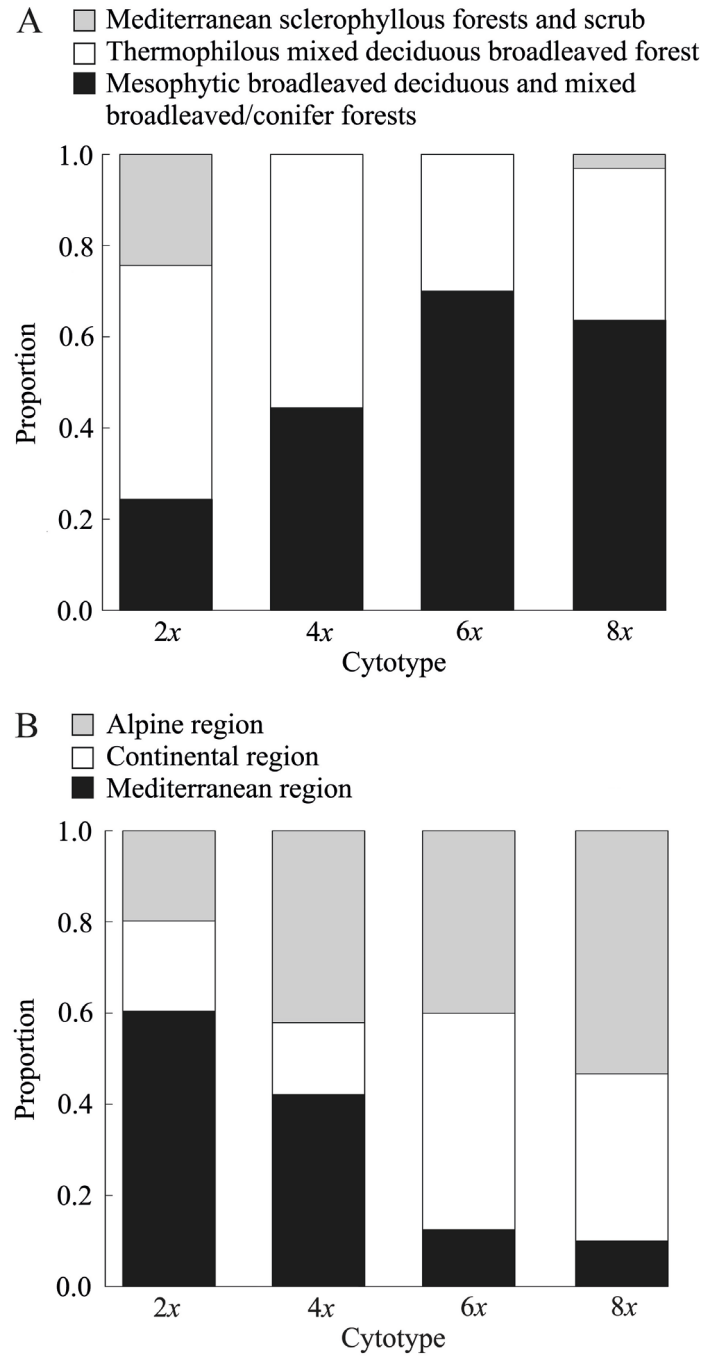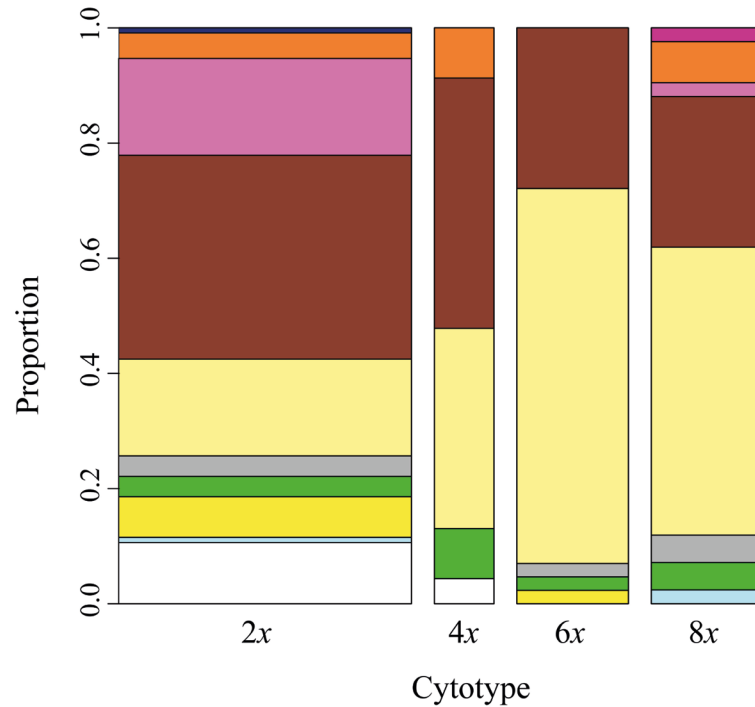
Fig. S3. Relative frequencies of cytotypes of *V.* subsect. *Pentasepalae* in relation to the three most common vegetation types (A), and the three most common biogeographical regions (B).
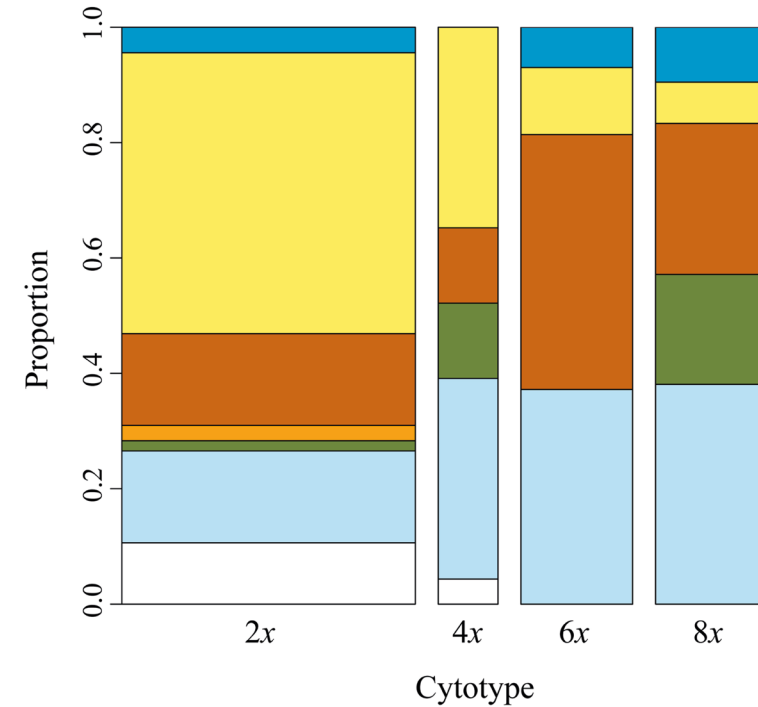
A Potential Natural Vegetation

B Biogeographic Regions

Legend A:
- No information (12 african populations)
- Arctic tundras and alpine vegetation
- Subarctic, boreal and nemoral-montane open woodlands as well as subalpine and oromediterranean vegetation
- Mesophytic and hygromesophytic coniferous and mixed broadleaved-coniferous forests
- Mesophytic broadleaved deciduous and mixed broadleaved/conifer forests
- Thermophilous mixed deciduous broadleaved forests
- Mediterranean sclerophyllous forests and scrub
- Xerophytic coniferous forests, coniferous woodland and scrub
- Vegetation of floodplains, estuaries, fresh-water polders and other moist or wet sites
- Coastal vegetation and inland halophytic vegetation
- Forest steppes (Meadow steppes alternating with nemoral deciduous forests) and dry grasslands alternating with dry scrub

Legend B:
- No information (12 african populations)
- Alpine Region
- Atlantic Region
- Black Sea Region
- Continental Region
- Mediterranean Region
- Pannonian Region

Fig. S4. Distribution of cytotypes of *V.* subsect. *Pentasepalae* in relation to vegetation types (A), and biogeographical regions (B).

Table S3. Summary of the 33 environmental variables used in univariate and multivariate analyses, its definition, spatial resolution, and source. Variables in bold represent the 17 low correlated variables used in the multivariate analysis.

Variable	Definition	Spatial Resolution	Data source
Bio 1	Annual Mean Temperature (ºC *10)		
Bio 2	Mean Diurnal Range (Mean of monthly (max temp - min temp)) (°C * 10)		
Bio 3	**Isothermality (BIO2/BIO7) (* 100) (%)**		
Bio 4	**Temperature Seasonality (standard deviation * 100) (°C * 10)**		
Bio 5	Max. Temperature of Warmest Month (°C * 10)		
Bio 6	Min. Temperature of Coldest Month (°C * 10)		
Bio 7	Temperature Annual Range (BIO5-BIO6) (°C * 10)		
Bio 8	Mean Temperature of Wettest Quarter (°C * 10)		WorldClim
Bio 9	**Mean Temperature of Driest Quarter (°C * 10)**	30 arc seconds (ca. 1 km)	(1960-1990)
Bio 10	Mean Temperature of Warmest Quarter (°C * 10)		(Hijmans *et al.*, 2005)
Bio 11	**Mean Temperature of Coldest Quarter (°C * 10)**		
Bio 12	Annual Precipitation (mm)		
Bio 13	Precipitation of Wettest Month (mm)		
Bio 14	**Precipitation of Driest Month (mm)**		
Bio 15	Precipitation Seasonality (Coefficient of Variation) (%)		
Bio 16	Precipitation of Wettest Quarter (mm)		
Bio 17	Precipitation of Driest Quarter (mm)		
Bio 18	**Precipitation of Warmest Quarter (mm)**		
Bio 19	Precipitation of Coldest Quarter (mm)		
NDVI av	**NDVI* annual average**	250 m	NASA EOSDIS LP DAAC

NDVI sd	NDVI* annual standard deviation		(Vermote, 2015)
NDVI min	NDVI* annual minimum		
NDVI max	NDVI* annual maximum		
NDVI ran	**NDVI* annual range**		
Tree cover	**Percent Tree Cover from 2010**	250 m	Global Land Cover Facility (VCF) (NASA) (DiMiceli *et al.*, 2011)
Slope	**Slope map in percent from 2002 to 2010**	2 km	
Sunr av	**Solar radiation average from 2011 (× 365/8 kWh / m2)**	700 m	WorldGrids (Reuter and Hengl, 2012)
Sunr sd	**Solar radiation standard deviation from 2011 (× 365/8 kWh/m2)**	700 m	
HFP	**Global Human Footprint from 1995 to 2004**	1 km	WCS and CIESIN (2005)
Soil	**Soil type**	1 km	World Soil information (Hengl *et al.*, 2014)
N Formation	**Potential natural vegetation**	Scale 1:2,500,000	Map of the Natural Vegetation of Europe (Bohn *et al.*, 2003)
Land cover	**Land cover from 2008 to 2012**	300 m	European Space Agency-CCI (Land Cover Map, 2014)
Biogeo	**Biogeographical region from 2008 to 2011**	**	European Environment Agency (2011)

*NDVI: Normalized Difference Vegetation Index averaged over the years 2001-2011.

** The scale varies between 1:1 000 000 and 1:10 000 000 depending on regions.

References cited in Table S3

Bohn U, Neuhäusl R, Gollub G, *et al.* 2003. *Karte der natürlichen Vegetation Europas/Map of the natural vegetation of Europe. Maßstab/Scale 1: 2 500 000.* Münster: Bundesamt für Naturschutz.

DiMiceli C, Carroll ML, Sohlberg RA, Huang C, Hansen MC, Townshend JRG. 2011. *Annual Global Automated MODIS Vegetation Continuous Fields (MOD44B) at 250 m Spatial Resolution for Data Years Beginning Day 65, 2000 - 2010, Collection 5 Percent Tree Cover, University of Maryland, College Park, MD, USA.* http://www.glcf.umd.edu/data/vcf/index.shtml [Accessed 13 October 2015].

European Environment Agency (EEA). 2011. *Biogeographical regions.* http://www.eea.europa.eu/data-and-maps/data/biogeographical-regions-europe-1 [Accessed 21 April 2013].

Hengl T, de Jesus JM, MacMillan RA, *et al.* 2014. SoilGrids1km–Global Soil Information Based on Automated Mapping. *PLoS ONE* **9**: e105992.

Hijmans R, Cameron S, Parra J, Jones P, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25: 1965–1978.

Land Cover Map. 2014. *European Space Agency – Land Cover Climate Change Initiative.* http://maps.elie.ucl.ac.be/CCI/viewer/index.php (Accessed 13 October 2015).

Reuter HI, Hengl T. 2012. Worldgrids - a public repository of global soil covariates. In: Minasny B, Malone BP, McBratney AB, eds. *Digital soil assessments and beyond.* Sydney: Taylor & Francis Group, 287–292.

Vermote E. 2015. *MOD09Q1 MODIS/Terra Surface Reflectance 8-Day L3 Global 250m SIN Grid V006* [Data set]. NASA EOSDIS LP DAAC. http://doi.org/10.5067/MODIS/MOD09Q1.006 [Accessed 26 October 2015].

WCS (Wildlife Conservation Society), and CIESIN (Center for International Earth Science Information Network) Columbia University. 2005. *Last of the Wild Project, Version 2, 2005 (LWP-2): Global Human Footprint Dataset (Geographic).* Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). http://dx.doi.org/10.7927/H4M61H5F [Accessed 15 October 2015].

Table S4. Associations between ploidy levels and environmental variables in populations of *Veronica* subsect. *Pentasepalae* found by the univariate analyses. Differences were tested either by one-way ANOVA, Kruskal-Wallis test (KW) or contingency tables (CT). d.f. = degree of freedom. Variables in bold showed significant association after Bonferroni correction (*P*-values < 0.0003).

Variable	Test	d.f.	Test statistics	*P*
Bio 1	ANOVA	3	6.038	0.0006
Bio 2	KW	3	7.221	0.065
Bio 3	KW	3	24.036	**2.45e-05**
Bio 4	KW	3	29.3944	**1.85e-06**
Bio 5	ANOVA	3	9.063	**1.12e-05**
Bio 6	KW	3	22.1684	**6.02e-05**
Bio 7	KW	3	14.1403	0.0027
Bio 8	KW	3	6.3198	0.0970
Bio 9	KW	3	29.4925	**1.76e-06**
Bio 10	ANOVA	3	6.219	0.0005
Bio 11	KW	3	31.613	**6.32e-07**
Bio 12	KW	3	11.075	0.011
Bio 13	KW	3	8.5406	0.036
Bio 14	KW	3	28.4057	**2.99e-06**
Bio 15	KW	3	11.4439	0.010
Bio 16	KW	3	9.8693	0.020
Bio 17	KW	3	18.6895	0.0003
Bio 18	ANOVA	3	25.65	**3.06e-14**
Bio 19	KW	3	5.8955	0.117
NDVI av	KW	3	28.946	**2.299e-06**
NDVI sd	KW	3	44.904	**9.695e-10**
NDVI min	KW	3	0.539	0.91
NDVI max	KW	3	46.867	**3.71e-10**
NDVI ran	ANOVA	3	16.72	**8.26e-10**
Tree cover	KW	3	15.3953	0.00151
Slope	KW	3	2.329	0.507
Sunr mean	KW	3	2.6223	0.454
Sunr sd	KW	3	6.9591	0.073
HFP	KW	3	1.0817	0.782
Soil	CT	33	41.53	0.147
N Formation	CT	30	70.43	**4.25e-05**
Land cover	CT	24	37.27	0.041
Biogeo	CT	18	80.26	**7.716e-10**