

Research Article

Heavy Metal Pollution, Selection, and Genome Size: The Species of the Žerjav Study Revisited with Flow Cytometry

Eva M. Temsch,¹ Wilhelm Temsch,² Luise Ehrendorfer-Schratt,³ and Johann Greilhuber¹

¹ Department of Systematic and Evolutionary Botany, University of Vienna, Rennweg 14, 1030 Vienna, Austria

² Core Unit for Medical Statistics and Informatics, Medical University of Vienna, Währinger Gürtel 18-20, 1090 Vienna, Austria

³ Department of Biogeography and Botanical Garden, University of Vienna, Rennweg 14, 1030 Vienna, Austria

Correspondence should be addressed to Eva M. Temsch, eva.temsch@univie.ac.at

Received 2 March 2010; Accepted 29 April 2010

Academic Editor: Jaroslav Doležal

Copyright © 2010 Eva M. Temsch et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The Death Valley at Žerjav in northern Slovenia exhibits a gradient of heavy metal pollution in the soil with severe consequences for species richness and composition along this gradient. Recently, a progressive loss of large-genome species in parallel with increasing concentrations of heavy metals has been shown. Here, we have measured the genome size of a near-complete sample of these species with flow cytometry and analysed the correlation of heavy metal pollution with the C- and Cx-values assigned to the test plots. The method of probability analysis was a hypergeometric distribution method. We confirm, on a different methodological basis than previously, that along the pollution gradient, species with high C- and Cx-values are increasingly underrepresented. This lends support to the “large genome constraint hypothesis”, predicting that plants with large genomes are at a disadvantage under all aspects of evolution, ecology, and phenotype, because junk DNA imposes a load to the organism.

1. Introduction

The molecular mechanisms that lead to the more than 2000-fold variation in the amount of DNA in higher plants' genomes (Zonneveld [1]), are no longer as mysterious as they were when the term “C-value paradox” was first used (Thomas [2]). The biological significance of this variation, however, is still debated. Several hypotheses are available, of which the Nucleotype Hypothesis has been the most fruitful since its formulation by Bennett [3, 4]. It emphasizes the importance of the physicommechanical properties of the cell nucleus (the nucleotype), in addition to the genotype and the environment, for niche occupation, adaptation, and competitiveness of an organism, in both plants and animals (Gregory [5]). Up to the present, the only well-studied nucleotypic parameters are the C-value and the Cx-value. C-value or holoploid genome size (Greilhuber et al. [6]), that is, the DNA amount contained in the chromosome complement of an organism, is directly and positively correlated with cell size (Bennett [4], Knight and Beaulieu [7]). Cx-value or monoploid genome size, the DNA amount contained in the single basic chromosome set with chromosome number x ,

has a strong positive influence on cell cycle duration (Bennett [3, 4], Francis et al. [8]). This is clearly a simplification inasmuch as cell and nuclear size vary according to cell type and physiological status (Bennett and Rees [9]). Cell cycle duration is influenced by genetic controls (Francis [10]) and environment, notably temperature (Verma and Lin [11], Verma [12]). Moreover, the distinction between C- and Cx-value is blurred in the case of (palaeo)polyploids with diploidized genomes. Nevertheless, the correlation C-value/cell size/nuclear size holds closely within certain cell types, for example, meristematic and embryonal cells (Greilhuber [13]) and stomata (Knight and Beaulieu [7]), while the correlation Cx-value/cell cycle time has been proven repeatedly (Bennett [3], Verma and Lin [11]). It is noteworthy that (neo)polyploids do not show longer cell cycle duration compared to diploids (Verma and Lin [11]), and sometimes it is even slightly shorter (Bennett [3]). Since the Cx-value of an organism together with the ploidy level determines the C-value, cell cycle duration is also correlated with C-values in global analyses (Francis et al. [8]). Several lines of evidence lend support to the opinion that large genomes are a burden to organisms and restrict

their adaptability. This concept was recently forwarded as the “Large Genome Constraint Hypothesis” (Knight et al. [14]). It would thus be predicted that under stressful environments species with too large a genome increasingly go extinct, either locally or totally. However, the species and genome size spectrum along stress gradients in natural environments has rarely been investigated.

Here, we return to a recent study, in which a number of species and their genome sizes were studied with regard to a gradient of heavy metal pollution along which they grew (Vidic et al. [15]). The locality studied is the “Death Valley”, Dolina Smrti, at Žerjav in northern Slovenia. The heavy metal pollution of the soil stems from a former lead smelter and reaches 33.3 g/kg soil at a plot on which plants are still growing. Heavy metals and plant community composition had been analyzed in five test plots (A–E) at distances of 330 m (E), 420 m (D), 520 m (C), 670 m (B), and ~2000 m (A) from the smelter chimney, the last site being the control plot. Diploid herbaceous species had been recorded and collected, and their genome size measured with Feulgen DNA image densitometry. The number of species occurring at the test plots was shown to decrease with increasing pollution, as was the number of species with large genomes (arbitrarily defined as those of the upper quartile of the C-values in the total sample of 70 species). Using a simulation method it was shown that the probability of finding such low frequencies of large genomes by chance decreases with increasing lead pollution. The results were interpreted as being supportive of the Large Genome Constraint Hypothesis (Knight et al. [14]).

The intention behind undertaking a new study on this system is to corroborate the conclusions from the Žerjav study using a technical approach now in much wider use than Feulgen densitometry, that is, flow cytometry using propidium iodide as the DNA stain. Flow cytometry is not only faster, but also more precise thanks to the high number of nuclei measured and other characteristics of the technique (Greilhuber et al. [16]). There was, furthermore, one genome size of *Anemone nemorosa* that was unsupported by literature data [17] and that had to be clarified and possibly corrected. The species measured were collected mostly not in the Death Valley, but this is unproblematic for the present purpose, because genome sizes are fairly stable within a narrow taxon (species at a ploidy level).

To calculate of the chance probability of finding at the test plots the given frequency of large-genome species, we use instead of a lottery simulation method the hypergeometric distribution method. In the Žerjav study by Vidic et al. [15], the probability of finding the observed frequency of large genomes solely by statistical fluctuation was determined by a lottery simulation. But there is a formula to calculate this probability. Distribution of these probabilities is the hypergeometric distribution. Calculation is always advantageous over a lottery simulation, since the result is a priori infinitely precise. A lottery simulation is always an approximation to the calculated value; a repeat will most likely not give the same result. To avoid numerical problems, established statistics software should be used to calculate the hypergeometric distribution.

2. Materials and Methods

2.1. Plant Material. Living plant species as listed in Vidic et al. [15] were collected in Austria, Croatia, and Slovenia (Table 1). Whole plants were wrapped with wet tissue and stored in a plastic bag at the collection sites, and usually kept in the refrigerator for up to one week until investigation with flow cytometry. For *Euphorbia amygdaloides* whole dry fruits were collected. Samples exhibiting potentially problematic compounds (e.g., mucous polysaccharides) were incubated for up to one month for a starving period in the refrigerator, after which time young etiolated leaves or shoots were used for flow cytometry with better success. Herbarium specimens are deposited in WU. Identification of the taxa is based on Fischer et al. [18]. These accessions cover 60 species from 51 genera and 22 families. The C-value of *Orobanchе alba* is cited from Weiss-Schneeweiss et al. [19] and those of *Knautia drymeia* and *Scabiosa columbaria* from Temsch and Greilhuber [20] (Table 2). Altogether, 63 species from 54 genera and 24 families are included in this paper.

2.2. Flow Cytometry (FCM). Following the chopping method of Galbraith et al. [21], about 25 mg fresh leaves (dry fruits in case of *Euphorbia amygdaloides*) from each plant sample were co-chopped in Otto's buffer I (Otto et al. [22]) together with *Pisum sativum* (1C = 4.42 pg DNA; Greilhuber and Ebert [23]), *Zea mays* (1C = 2.73 pg DNA; Doležel et al. [24]), *Secale cereale* (1C = 7.79 pg DNA; Doležel et al. [24]), or *Solanum pseudocapsicum* (1C = 1.29 pg DNA; Temsch et al. [25]) as the internal standard organisms. The resulting isolated nuclei were filtered through a 30 µm mesh and subsequently incubated with RNase A for 30 minutes at 37°C for digestion of double-stranded RNA. The nuclei were then stained in Otto's buffer II (Otto et al. [22]) containing the fluorochrome propidium iodide (PI, 50 µg/ml) for 1 hour or overnight in the refrigerator. For measurement, a CyFlow ML flow cytometer or a PAII (both Partec, Muenster, Germany) was used. Light sources were for the CyFlow a green laser (532 nm, Cobolt Samba, Cobolt AB, Stockholm, Sweden) and for the PAII a mercury lamp. One preparation was made per individual and from this at least three measurement runs were performed, with 5000 measured particles per run. Usually, the coefficient of variation (CV) was less than 3%, but whenever higher CVs occurred, one or two more runs were added. For each run, the 1C-value was calculated according to the formula: mean fluorescence intensity of the sample organism's G1 nuclei population divided by mean fluorescence intensity of the standard's G1 nuclei population times the 1C-value of the standard organism. The resulting sample values are shown in Table 2.

2.3. Chromosome Counts. Whenever the ploidy levels of our samples had to be clarified, chromosome counts were done on slides made according to protocols for Feulgen densitometric analysis (Greilhuber and Temsch [26]), which was the case in *Campanula rotundifolia*, *Fragaria vesca*, *Hieracium murorum*, *Knautia drymeia*, *Lathyrus pratensis*,

TABLE 1: Locality code, country, district, province, village, longitude, latitude, and altitude (m a.s.l.) of each collection site.

Locality Code	District	Province ^a	Village	Longitude ^b	Latitude ^b	Altitude
<i>Austria</i>						
1	Eisenstadt-U.	B	Siegenderdorf	16° 34' 52''	47° 46' 42''	150
2	Mattersburg	B	Forchtenstein	16° 18' 30''	47° 41' 49''	730
3	Neusiedl am See	B	Jois	16° 46' 10''	47° 57' 06''	150
4	Oberwart	B	Bernstein	16° 17' 02''	47° 25' 20''	610
5	Klagenfurt Land	C	Ferlach	14° 15' 46''	46° 29' 09''	640–700 ^c
6	Villach Land	C	Villach	13° 46' 32''	46° 34' 22''	550
7	Völkermarkt	C	Völkermarkt	14° 38' 44''	46° 41' 09''	540
8	Baden	L	Altenmarkt an der Triesting	15° 56' 01''	48° 00' 43''	480
9	Baden	L	Furth an der Triesting	16° 00' 17''	47° 58' 45''	400–440 ^c
10	Baden	L	Pfaffstätten	16° 15' 22''	48° 02' 15''	320
11	Bruck/Leitha	L	Götzendorf a. d. Leitha	16° 31' 10''	48° 01' 34''	180
12	Bruck/Leitha	L	Hundsheim	16° 55' 27''	48° 07' 47''	280–300 ^c
13	Krems	L	Krems/Hollenburg	15° 41' 32''	48° 22' 15''	320
14	Lilienfeld	L	Kleinzell	15° 41' 29''	47° 58' 42''	1050
15	Lilienfeld	L	St. Ägyd am Neuwalde	15° 31' 04''	47° 48' 59''	800
16	Melk	L	Dunkelsteinerwald	15° 27' 05''	48° 15' 13''	430
17	Melk	L	Texingtal	15° 17' 18''	48° 01' 26''	720
18	Mistelbach	L	Mistelbach	16° 38' 19''	48° 33' 49''	210
19	Mödling	L	Perchtoldsdorf	16° 15' 07''	48° 07' 33''	320
20	Neunkirchen	L	Reichenau a.d. Rax	15° 45' 40''	47° 44' 37''	590
21	Neunkirchen	L	Reichenau a.d. Rax	15° 46' 01''	47° 42' 51''	1600
22	Neunkirchen	L	Reichenau a.d. Rax	15° 46' 26''	47° 43' 10''	1580
23	Neunkirchen	L	Schwarzau im Gebirge	15° 38' 41''	47° 47' 38''	790
24	St. Pölten	L	Traismauer	15° 45' 52''	48° 20' 48''	190
25	St. Pölten	L	Traismauer	15° 45' 56''	48° 22' 12''	190
26	St. Pölten	L	Traismauer	15° 47' 23''	48° 22' 00''	185
27	Tulln	L	Sieghardtskirchen	16° 03' 24''	48° 14' 07''	240
28	Tulln	L	Sieghardtskirchen	16° 04' 20''	48° 13' 47''	280
29	Wiener Neustadt	L	Gutenstein	15° 51' 00''	47° 48' 40''	990
30	Wiener Neustadt	L	Miesenbach	15° 59' 59''	47° 51' 07''	650
31	Wiener Neustadt	L	Winzendorf-Muthmannsdorf	16° 07' 21''	47° 49' 42''	500
32	Wien-Umgebung	L	Purkersdorf	16° 11' 47''	48° 11' 21''	210
33	Bruck an der Mur	S	Bruck an der Mur	15° 17' 07''	47° 23' 42''	480
34	Hartberg	S	Waldbach	15° 48' 12''	47° 27' 19''	720
35	Liezen	S	Ramsau am Dachstein	13° 37' 23''	47° 27' 27''	1800
36	Liezen	S	Weng im Gässle	14° 32' 09''	47° 36' 53''	850
37	Mürzzuschlag	S	Altenberg an der Rax	15° 37' 19''	47° 41' 48''	1200
38	Mürzzuschlag	S	Altenberg an der Rax	15° 37' 25''	47° 42' 16''	1500
39	Mürzzuschlag	S	Altenberg an der Rax	15° 38' 29''	47° 40' 53''	1500
40	Weiz	S	Ratten	15° 45' 54''	47° 30' 16''	800
41	Lienz	T	Kals am Großglockner	12° 37' 31''	47° 01' 35''	1638
42	Gmunden	U	Grünau im Almtal	13° 57' 00''	47° 45' 14''	590
43	Vienna	V	Vienna/Lainzer Tiergarten	16° 14' 40''	48° 10' 13''	295
44	Vienna	V	Vienna/Gallitzinberg	16° 16' 24''	48° 13' 43''	440
45	Vienna	V	Vienna/Lobau	16° 28' 24''	48° 12' 19''	155
46	Vienna	V	Vienna/HBV	16° 23' 02''	48° 11' 36''	170
<i>Croatia</i>						
47	Istarska	—	Prascari	13° 45' 54''	45° 16' 45''	300

TABLE 1: Continued.

Locality Code	District	Province ^a	Village	Longitude ^b	Latitude ^b	Altitude
<i>Slovenia</i>						
48	Koroska	—	Mezica	14° 52' 10''	46° 29' 05''	550
49	Notranjska	—	Ilirska Bistrica	14° 13' 26''	45° 34' 58''	400
50	Notranjska	—	Postojna	14° 00' 36''	45° 47' 34''	780
51	Notranjska	—	Postojna	14° 12' 03''	45° 47' 49''	580
52	Primorska	—	Divaca	13° 58' 37''	45° 39' 43''	450
53	Stajerska	—	Ptuj	16° 00' 00''	46° 22' 12''	200

^aAbbreviations of Austrian provinces. B = Burgenland, C = Carinthia, L = Lower Austria, S = Styria, T = Tyrol, U = Upper Austria, V = Vienna, ^bWorld Geodetic System 1984, 'Collection areas' positions are given as the coordinates of the mid point.

TABLE 2: Chromosome numbers, ploidy levels, and mean C-values (Gbp and pg) from Vidic et al. [15] compared to the mean C- and Cx-values of the present study, and the locality code. Species are listed in alphabetical order.

Vidic et al. [15]					Present study						
Taxon	2n	Ploidy	2C (Gbp)	1C (pg)	Taxon mean values			Accession data			
					1C (pg)	S.D. (pg)	1Cx (pg)	Locality Code	1C (pg)	S.D. (pg)	N ^a
<i>Anemone nemorosa</i> L.	30	4x	10.80	5.52	19.479	0.359	9.740	48	19.073	0.267	5
								16	19.610	0.692	5
								46	19.755	0.058	5
<i>Arabis hirsuta</i> (L.) Scop.	32	4x	1.34	0.69	0.454		0.227	13	0.454	0.002	5
<i>Aremonia agrimonoides</i> (L.) DC.	35	5x	2.43	1.24							
<i>Aruncus dioicus</i> (Walter) Fernald	18	2x	0.42	0.21	0.174	0.002	0.174	27	0.173	0.001	3
								53	0.176	0.001	3
<i>Asarum europaeum</i> L.	26	2x	12.37	6.32	6.646	0.138	6.646	46	6.554	0.019	3
								53	6.579	0.016	3
								9	6.805	0.032	3
<i>Aster bellidiastrum</i> (L.) Scop. ^b	18	2x	3.12	1.60	1.517		1.517	5	1.517	0.008	3
<i>Astragalus glycyphyllos</i> L.	16	2x	1.63	0.83	0.751		0.751	26	0.751	0.002	3
<i>Betonica alopecuroides</i> L.	16	2x	4.43	2.26	2.421		2.421	37	2.421	0.051	3
<i>Biscutella laevigata</i> L. ^c	36	4x	3.62	1.85	1.828		0.914	35	1.828	0.004	3
<i>Buphthalmum salicifolium</i> L.	20	2x	4.72	2.41	2.490		2.490	19	2.490	0.001	3
<i>Campanula cespitosa</i> Scop.	34	2x	2.64	1.35	1.093	0.009	1.093	6	1.083	0.006	3
								36	1.086	0.005	3
								22	1.094	0.003	3
								20	1.100	0.001	3
								20	1.103	0.001	3
<i>Campanula rotundifolia</i> L.	34	2x	2.62	1.34	1.096		1.096	7	1.096	0.005	3
<i>Campanula thyrsoidea</i> L.	34	2x	5.25	2.68	2.365		2.365	37	2.365	0.005	3
<i>Carlina acaulis</i> L.	20	2x	10.27	5.25	5.094		5.094	14	5.094	0.023	3
<i>Centaurea fritschii</i> Hayek	20	2x	3.53	1.80	1.681 ^d		1.681	19	1.681	0.015	3
<i>Cirsium erisithales</i> (Jacq.) Scop.	34	2x	2.38	1.22	1.056		1.056	30	1.056	0.011	3
<i>Cruciata glabra</i> (L.) Ehrend.	22	2x	1.45	0.74	0.681		0.681	2	0.681	0.003	3
<i>Cyclamen purpurascens</i> Mill.	34	2x	6.03	3.08	3.303	0.002	3.303	9	3.302	0.004	3
								53	3.305	0.010	3
<i>Digitalis grandiflora</i> Mill.	56	2x	2.55	1.30	1.150		1.150	26	1.150	0.005	5
<i>Erysimum sylvestre</i> Scop.	14	2x	0.55	0.28							
<i>Eupatorium cannabinum</i> L.	20	2x	4.83	2.47	2.577	0.033	2.577	50	2.554	0.004	3
								11	2.601	0.005	3

TABLE 2: Continued.

Vidic et al. [15]					Present study						
Taxon	2n	Ploidy	2C (Gbp)	1C (pg)	Taxon mean values			Accession data			N ^a
					1C (pg)	S.D. (pg)	1Cx (pg)	Locality Code	1C (pg)	S.D. (pg)	
<i>Euphorbia amygdaloides</i> L.	20	2x	5.38	2.75	3.508		3.508	36	3.508	0.006	3
<i>Euphorbia cyparissias</i> L.	40	4x	2.17	1.11							
<i>Fragaria vesca</i> L.	14	2x	0.31	0.16	0.246	0.006	0.246	9	0.243	0.002	3
								27	0.254		1
<i>Galium anisophyllum</i> agg. <i>austriacum</i>	22	2x	1.00	0.51							
<i>Galium mollugo</i> L. ^e	44	4x	3.69	1.89	1.885		0.942	49	1.885	0.013	3
<i>Galium verum</i> L.	44	4x	3.80	1.94	1.887	0.047	0.943	4	1.775	0.001	3
								3	1.851	0.002	3
								24	1.869	0.004	3
								19	1.889	0.004	3
								11	1.894	0.006	3
								18	1.894	0.006	3
								32	1.914	0.004	3
								1	1.922	0.007	3
								49	1.929	0.004	3
								47	1.932	0.012	3
<i>Galium x pomeranicum</i>	44	4x	3.66	1.87	1.888	0.007	0.944	3	1.882	0.002	3
								32	1.885	0.010	3
								18	1.895	0.009	3
<i>Gentiana asclepiadea</i> L.	44	4x	6.45	3.30	3.140	0.023	1.570	53	3.124	0.003	3
								39	3.156	0.008	3
<i>Gentianella ciliata</i> (L.) Borkh.	44	4x	20.07	10.26	9.004		4.502	42	9.004	0.014	3
<i>Helleborus niger</i> L.	32	4x	29.87	15.27	14.728	0.185	7.364	5	14.552	0.054	3
								23	14.683	0.087	3
								9	14.686	0.168	3
								15	14.990	0.045	3
<i>Hepatica nobilis</i> Schreber	14	2x	34.45	17.61	16.709		16.709	12	16.709	0.157	3
<i>Hieracium murorum</i> L.	27	3x	9.02	4.61	5.706		3.804	37	5.706	0.045	3
<i>Hippocrepis comosa</i> L.	28	4x	3.73	1.91							
<i>Hypericum perforatum</i> L.	32	4x	1.59	0.81	0.783	0.065	0.392	19	0.737	0.004	3
								46	0.829	0.002	3
<i>Knautia drymeia</i> Heuff.	40	4x	6.38	3.26	7.031 ^f		3.516				
<i>Laserpitium peucedanoides</i> L.	22	2x	3.03	1.55							
<i>Lathyrus pratensis</i> L.	14 ^g	2x	8.89	4.54	5.951		2.975	44	5.951	0.010	3
<i>Leontodon hispidus</i> L.	14	2x	4.46	2.28	2.188	0.057	2.188	33	2.123	0.026	3
								33	2.212	0.002	3
								13	2.230	0.024	3
<i>Leontodon incanus</i> (L.) Schrank	8	2x	2.39	1.22	0.956		0.956	36	0.956	0.001	3
<i>Lotus corniculatus</i> L.	24	4x	2.48	1.27	1.235		0.618	13	1.235	0.003	3
<i>Medicago lupulina</i> L.	16	2x	1.35	0.69	0.649		0.649	13	0.649	0.003	3
<i>Minuartia gerardii</i> (Willd.)	24	2x	1.98	1.01	1.092		1.092	21	1.092	0.004	3
<i>Origanum vulgare</i> L.	30	2x	1.79	0.92	0.718		0.718	12	0.718	0.004	3
<i>Orobanche alba</i> Steph. ex Willd.	38	2x	6.03	3.08	2.979 ^h		2.979				
<i>Petrorhagia saxifraga</i> (L.) Link	60 ⁱ	4x	2.20	1.12	0.648		0.648	3	0.648	0.001	3
<i>Phyteuma orbiculare</i> L.	22	2x	2.26	1.16	1.136	0.024	1.136	5	1.119	0.001	3
								29	1.153	0.007	3

TABLE 2: Continued.

Vidic et al. [15]					Present study						
Taxon	2n	Ploidy	2C (Gbp)	1C (pg)	Taxon mean values			Accession data			
					1C (pg)	S.D. (pg)	1Cx (pg)	Locality Code	1C (pg)	S.D. (pg)	N ^a
<i>Pimpinella saxifraga</i> L.	20	2x	8.33	4.26	3.823		3.823	28	3.823	0.015	3
<i>Plantago major</i> L.	12	2x	1.39	0.71	0.712	0.002	0.712	53	0.710	0.000	3
								11	0.713	0.003	3
<i>Plantago media</i> L.	24	4x	5.11	2.61	2.782	0.084	1.391	12	2.727	0.006	3
								31	2.741	0.006	3
								50	2.878	0.006	3
<i>Polygala amara</i> ssp. <i>brachyptera</i>	34	2x	1.05	0.54	0.422		0.422	4	0.422	0.002	3
<i>Potentilla erecta</i> (L.) Räuschel	28	4x	0.89	0.46							
<i>Primula vulgaris</i> Huds.	22	2x	1.14	0.58	0.470	0.006	0.470	8	0.463	0.001	3
								53	0.471	0.002	3
								46	0.475	0.009	3
<i>Prunella vulgaris</i> L.	28	2x	1.27	0.65	0.650	0.016	0.650	51	0.639	0.003	5
								11	0.661	0.002	3
<i>Ranunculus nemorosus</i> DC.	16	2x	9.86	5.04	6.147		6.147	41	6.147	0.014	3
<i>Salvia glutinosa</i> L.	16	2x	2.31	1.18	1.071	0.023	1.071	53	1.055	0.001	3
								25	1.087	0.006	5
<i>Scabiosa columbaria</i> L.	16	2x	2.12	1.08	1.072 ^j		1.072				
<i>Silene alpestris</i> Jacq.	24	2x	2.64	1.35	2.210	0.003	2.210	37	2.208	0.004	5
								36	2.212	0.009	3
<i>Silene dioica</i> (L.) Clairv.	24	2x	5.40	2.76	2.734	0.003	2.734	34	2.732	0.012	3
								40	2.736	0.007	3
<i>Silene nutans</i> L.	24	2x	4.96	2.54	2.385	0.005	2.385	17	2.381	0.004	3
								12	2.388	0.003	3
<i>Silene pusilla</i> Waldst. & Kit. s. lat.	24	2x	2.82	1.44	1.324		1.324	37	1.324	0.002	3
<i>Solidago virgaurea</i> L.	18	2x	1.77	0.90	1.134		1.134	38	1.134	0.005	3
<i>Symphytum tuberosum</i> L.	64	8x	5.73	2.93	2.745		0.686	43	2.745	0.014	3
<i>Taraxacum officinale</i> Weber in Wiggers	16	2x	2.50	1.28	1.254		1.254	45	1.254	0.004	3
<i>Thlaspi praecox</i> Wulf.	14	2x	0.66	0.34	0.262	0.024	0.262	46	0.245	0.001	3
								48	0.279	0.002	3
<i>Tussilago farfara</i> L.	60	2x	3.78	1.93	1.736	0.048	1.736	12	1.702	0.005	3
								12	1.770	0.004	3
<i>Verbascum austriacum</i> Schott ex Roem & Schult. ^k	26	2x	0.87	0.44	0.357		0.357	3	0.357	0.003	5
<i>Veronica officinalis</i> L.	36	4x	2.31	1.18	0.901		0.451	30	0.901	0.006	3
<i>Vincetoxicum hirundinaria</i> Med.	22	2x	0.71	0.36	0.311	0.006	0.311	12	0.307	0.003	3
								52	0.316	0.001	3
<i>Viola hirta</i> L.	20 ^g	2x	1.70	0.87	1.511		0.755	10	1.511	0.069	3

^aN = number of runs^b*Aster bellidiastrum* = *Bellidiastrum michelii* Cass.^cCollected by C. König^d*Centaurea fritschii* substituted by *Centaurea scabiosa* L. ssp. *scabiosa*^eTetraploid *Galium mollugo* is now *G. album* Mill.^f*Knautia drymeia*: The ploidy level of Vidic et al. [15] should be reinvestigated. The present value is from Temsch and Greilhuber [20]^gThe present samples of *Lathyrus pratensis* (2n = 28) and *Viola hirta* (2n = 40) are tetraploid^hWeiss-Schneeweiss et al. [19]ⁱThe present sample of *Petrorhagia saxifraga* is diploid (2n = 30)^jThe present value is from Temsch and Greilhuber [20]^k*Verbascum austriacum* = *Verbascum chaixii* Vill. subsp. *austriacum* (Roem. & Schult.) Hayek

Origanum vulgare, *Petrorhagia saxifraga*, *Plantago major*, *Plantago media*, *Veronica officinalis*, and *Viola hirta*. Proliferating parts of these species were fixed in 3 : 1 methanol/acetic acid (3:1) overnight at 7°C, or in formaldehyde (FA, 4% in Sørensen buffer) for two hours followed by 3:1 fixation, and transferred to 96% ethanol for further storage. Feulgen stained mitotic configurations were analyzed for chromosome number (Table 2).

2.4. Data Analysis. Vidic et al. [15] described a decrease in species having large nuclear characters with decreasing distance of the sample plots to a source of heavy metal pollution. Large nuclear characters are the upper quartile of C-values, Cx-values, chromosome number, and mean chromosomal DNA content. In Vidic et al. [15] a randomization test was used to determine the probability for large nuclear characters equal or less than observed for each plot. The probability can, however, be calculated directly: The number of distinct samples of size K out of a lot of size N is given by the binomial coefficient $Q = \binom{N}{K} = N!/K!(N-K)!$. (A number lottery is a common example: There are $\binom{45}{6}$ distinct patterns of 6 numbers out of 45). If we only consider large nuclear characters, there are $\binom{L}{i}$ distinct patterns of size i for L large characters in the entire lot. The number of samples of size K with exactly i large characters is given by $Q_i = \binom{L}{i} \binom{N-L}{K-i}$, since each of the large character patterns can be combined with all possibilities to select $K-i$ small characters out of $N-L$ in the lot. The probability of finding exactly i large genomes in a sample of size K is $p_i = Q_i/Q$. This is the hypergeometric probability distribution. The probability of finding i or less large genomes (cumulative probability) as described in Vidic et al. [15] is $\sum_{k=0}^L p_k$. Obviously $\sum_{k=0}^L p_k = 1$, which simply means, the number of large genomes in the sample is less or equal to L , the number of large genomes in the lot.

Binomial coefficients are composed of the factorials of the occurring numbers. These can be very large numbers, which are likely to exceed the range of representable numbers in a computer system, or at least substantial loss of precision has to be expected. Therefore it is strongly recommended to use established statistics software to calculate such probability distributions. Here we used SAS 9.1 (SAS [27]). As an alternative to commercial software we made trial runs based on prime factor decomposition of the factorials to circumvent the problems with large numbers. We got the same results as in SAS. The accordance of our data with Vidic et al. [15] was visualized by a Bland-Altman-plot (Figure 1).

3. Results

3.1. Data. Table 2 contains data from 63 species, of which 60 are original and have been measured with flow cytometry and 3 further were taken from published sources (Temsch and Greilhuber [20], Weiss-Schneeweiss et al. [19]). These are 90% of the species investigated by Vidic et al. [15]. Chromosomes were counted in nine species in which more than one ploidy level occurs and a deviation between our data and those of Vidic et al. [15] was recognized. In three species the ploidy level given in Vidic et al. [15] and here in

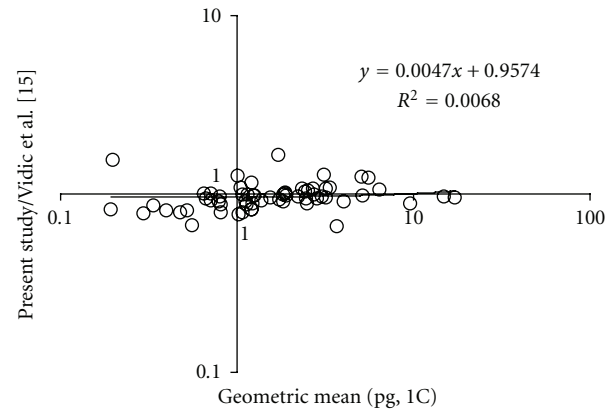


FIGURE 1: Modified Bland-Altman plot of the present data versus those of Vidic et al. [15] without the value of *Anemone nemorosa* and with incongruent ploidy levels corrected. The values are displayed on logarithmic axes.

Table 2 is not the same. Different ploidy levels (recognized by chromosome counts) between the two studies were found in the following taxa: *Lathyrus pratensis*, *Petrorhagia saxifraga*, and *Viola hirta* (Table 2). In this case, we corrected our C-values for ploidy level given for the Žerjav plants in the further calculations, assuming that the actual C-value of these plants closely approaches this estimate.

A comparison of the data of Vidic et al. [15] and those of the present study shows, besides a general agreement in many points, some striking deviations from the expectation. There is in the first instance *Anemone nemorosa*, of which the low 1C-value of 5.5 pg given in Vidic et al. [15] could not be confirmed with material from Žerjav and can only be a mismeasurement, because it is well below the 1C-value of diploids, which are expected to have about 10 pg (Zonneveld in Bennett and Leitch [17]). We also measured material from the Botanical Garden in Vienna (HBV) and from one accession in Lower Austria with congruent results (19.48 pg), very similar to the C-value of Zonneveld et al. [28] (1C = 19.05 pg). *Anemone nemorosa* was therefore excluded from some statistical analyses. Another case of deviation is *Knautia drymeia*. Vidic et al. [15] give 1C = 3.26 pg at $2n = 4x = 40$ (which is possibly an error and should read $2n = 20$), while we measure 1C = 7.031 pg at $2n = 4x = 40$ (Temsch and Greilhuber [20]).

The variation within the 1C-values ranged 114.6-fold from 0.17 pg in *Aruncus dioicus* to 19.48 pg in the tetraploid *Anemone nemorosa*. 1Cx-values varied 98.3-fold from 0.17 pg in *Aruncus dioicus* to 16.71 pg in *Hepatica nobilis*.

3.2. Comparison of the Data Sets. Among the 62 species for which agreement was established in regard to ploidy level, 38% agree within ± 0 to 5%, 24% within ± 5 to 10%, 19% within ± 10 to 20%, and 19% are more strongly deviating. The average ratio (present study/Vidic et al. [15]) over these species is 0.97. 14 C-values in Vidic et al. [15] are lower than the present ones. The correlation (C-values

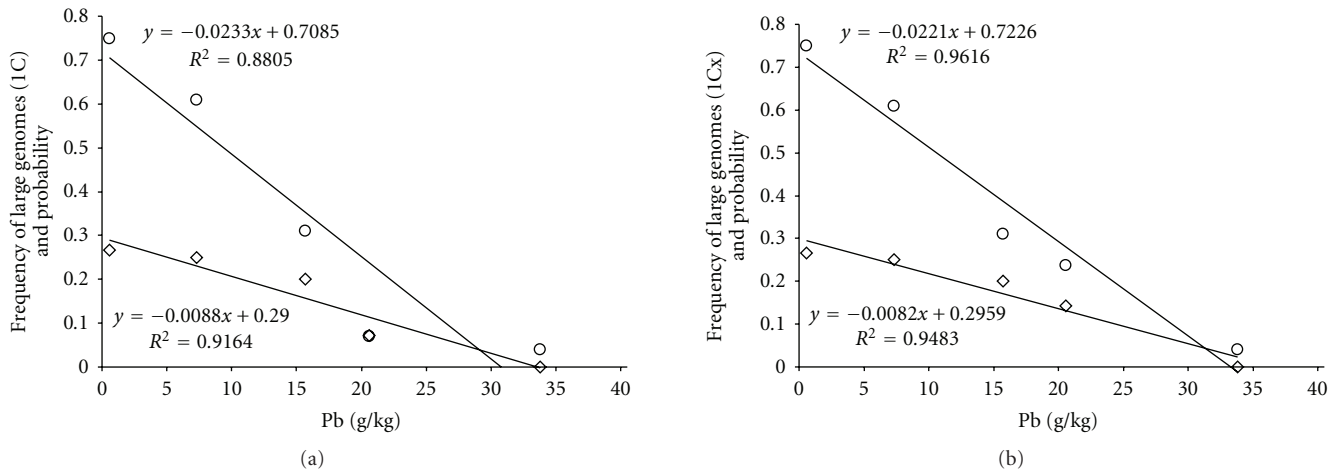


FIGURE 2: Correlation of lead concentration at the five test plots (from left to right A–E [15]) with (i) percent large genomes (diamonds) and (ii) probability of finding by chance that number or fewer large genomes (circles).

corrected for ploidy and without *A. nemorosa*) yielded a correlation coefficient $r = 0.9917$.

3.3. Probability Analysis. Vidic et al. [15] addressed the question of whether the degree of heavy metal pollution along a gradient has an influence on the C-value and the Cx-value spectra of the species living on five test plots. Since the number of species decreases with increasing pollution, the crucial point is whether the observed shift of the C-value composition towards lower C-values is stronger than expected under the condition of randomness. Vidic et al. [15] found a significant decrease of species with large genomes along increasing pollution using a lottery simulation technique. The present approach relies on the direct calculation of the probabilities of the occurrence of large nuclear characters.

3.3.1. Our Data Set Analysed Using Hypergeometric Distribution. There are minor differences between the values plotted in Vidic et al. [15, Figure 2(c)] and the results that we recalculated applying our method to the data of Vidic et al. [15]. We calculated the probability from the composition of the entire set (hypergeometric distribution), whereas Vidic et al. [15] counted the probability from randomly drawn samples. Their method is an estimate, strictly speaking, which would explain the deviation. It is unlikely that Vidic et al. [15] would obtain exactly the same results in case of a repeat of their method. In our calculation the probabilities for having found a random genome size spectrum decrease with increasing lead concentration at the plots. At the highest lead concentration, the probability level of 4.1% is reached with the C- and the Cx-value spectrum.

Using the limits for large genomes (upper quartile) from Vidic et al. [15] the result appears slightly more significant. This probability should not be misunderstood as significance of a statistical test. Rather it is a characterization of the detrimental effect of pollution on large genomes: a low probability makes it unlikely that pollution has no effect on

the phytocoenosis at this plot. The species and data set of Vidic et al. [15] analysed with hypergeometric distribution result in similar Figures (Figures 2(a) and 2(b)) as in [15, Figure 2], reaching 11.3% probability for the species composition at plot E [15] for 1Cx-values and 1.3% for 1C-values.

Regarding the frequency of large genomes, there are slight differences between the limits set by us and by Vidic et al. [15] because large genomes are defined as the upper quartile of the genome size distribution. In our data set, large genomes are those with $1C > 2.781$ pg and $1Cx > 2.576$ pg. The progressive elimination of large genomes (C- and Cx-values) with increasing heavy metal pollution and decreasing distance to the lead smelter is clear with both data sets (Figures 3(a)–3(c), Table 2), but much more convincing with the present 1Cx data. The deficit of large 1Cx-values at plot A in Vidic et al. [15] data seems to result from wrong ploidy levels in some species. The frequency of polyploid genomes also decreases but with less regularity (Figures 4(a) and 4(b)).

The correlations between heavy metal pollution at test plots on the one hand, and % species with large genomes ($P = .005$, $r = -0.974$) and the probabilities of hypergeometric distribution ($P = .003$, $r = -0.981$), on the other hand, are negative and statistically significant.

4. Discussion

Vidic et al. [15] demonstrated a strong negative correlation between heavy metal content of the soil and the proportion of plant species with large genomes along a gradient of pollution caused by the emissions of a lead smelter. The conclusion was that genome size is associated with differential survival of species in this extreme environment and that species with large genomes are at a disadvantage.

The present paper is in some respect a control investigation for the Žerjav study by Vidic et al. [15] using a different methodology for genome size measurement, that is, more precise flow cytometry instead of the error-susceptible

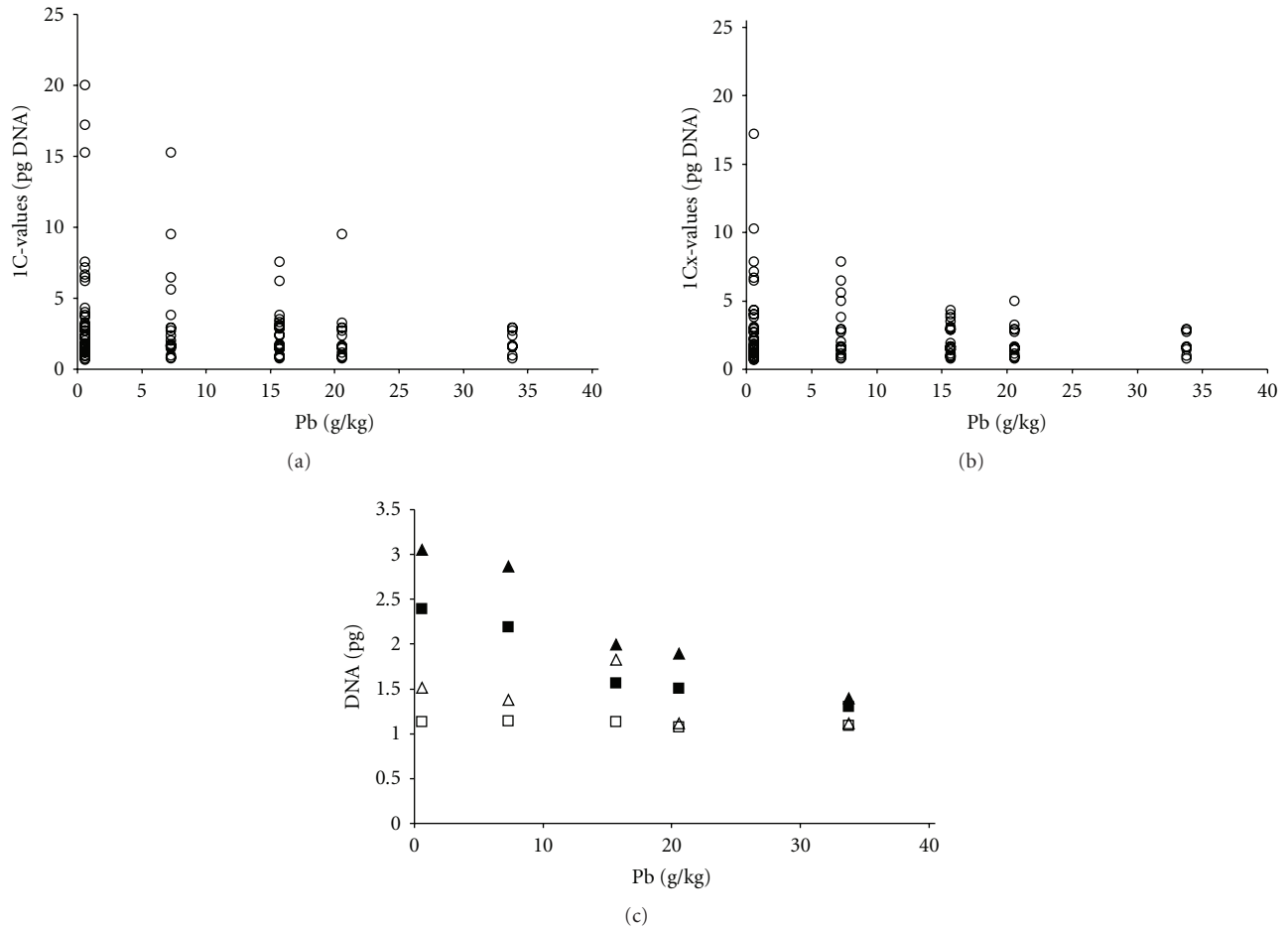


FIGURE 3: (a) Distribution of the 1C-values and (b) 1Cx-values of the taxa as they occur at the sample plots, and (c) the median 1Cx-values (empty squares), the mean 1Cx-values (filled squares), the median 1C-values (empty triangles), and the mean 1C-values (filled triangles).

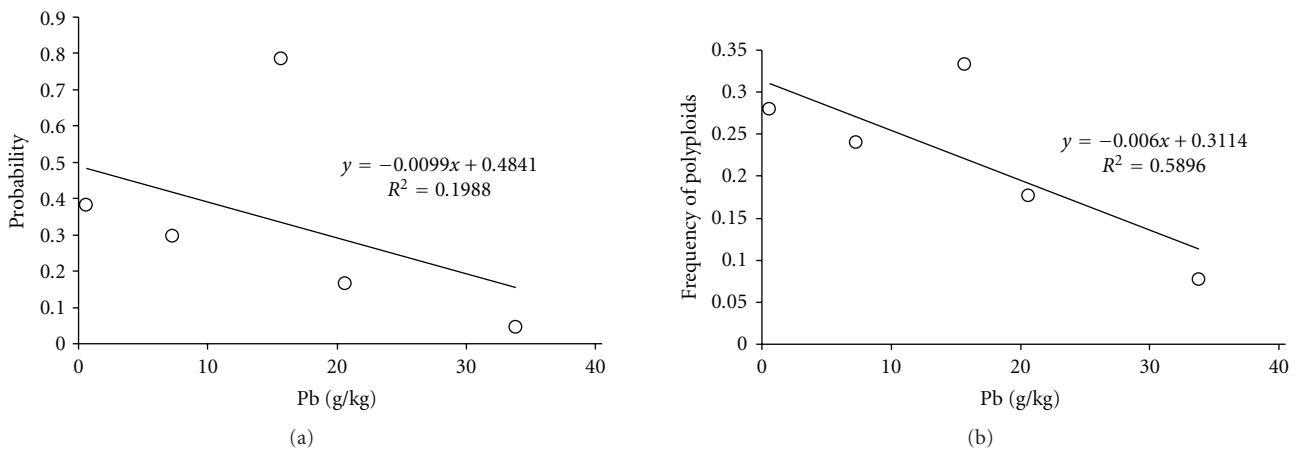


FIGURE 4: (a) Probability of finding by chance that number or fewer polyploids, and (b) percent polyploids at the five test plots.

Feulgen densitometry, and for calculation of probabilities for large genome frequencies, that is, a hypergeometric distribution method instead of a simulation method. Our material originated mostly not from Žerjav. However, the results are representative for the Žerjav locality because genome sizes are generally very stable within a species.

Our results confirm the findings of Vidic et al. [15] and exhibit in some regard even clearer trends. The species characterized by large genomes (i.e., upper quartile of the genome sizes) out of the complete species spectrum are increasingly removed in parallel with heavy metal pollution until none of these species is left in the most polluted site.

This clearer trend, compared to Vidic et al. [15], may be the consequence of removing a few possible errors from the Žerjav data set (mainly correction of a mismeasurement in *Anemone nemorosa* and possible misidentification of ploidy level in *Knautia drymeia*), and a general improvement of the genome size data using a more precise measurement method. Notably, *Anemone nemorosa* ranks highest in genome size and occurs only at the control site (plot A), while it is only fifth-highest in the data set of Vidic et al. [15].

Vidic et al. [15] considered only C-values, while we also consider Cx-values. With these, we find the same trends as with C-values inasmuch as the frequency of large monoploid genomes and the hypergeometrically distributed probability decrease continuously from plot A to plot E. The frequency of polyploid genomes also decreases finally, but plot C is the highest and violates any significance. The observed trends, therefore, cannot be attributed either to the C-value or to the Cx-value alone.

The positive correlation between radiosensitivity and genome size in plants described by Sparrow and Miksche [29] seems to be most relevant for the present results, because direct genotoxic effects are involved in both studies. A high frequency of chromosome aberrations has been observed by Druškovič [30] in plants from the Death Valley, Dolina Smrti at Žerjav. It seems plausible to us that DNA damage with its various consequences for cell division and cell function is the primary cause for the exclusion of large genomes at the more heavily polluted plots. Once a nucleus is hit, the damage is realized irrespective whether the nucleus is big or small, but large nuclei and especially mitotically and meiotically active ones are a larger target for damage and receive more hits.

Apart from chemical stress, also natural stress such as extreme temperatures and low precipitation seem to favour plants with small genomes rather than large ones in certain floras, whereas the species with largest genomes occur preferably in mesic areas (Knight and Ackerly [31]). This and other evidence led Knight et al. [14] to propose the Large Genome Constraint Hypothesis, which predicts that “plants with large genomes are at a disadvantage under all aspects of evolution, ecology and phenotype because large genomes are inflated with unnecessary junk DNA whose replication and maintenance imposes a load to the organism” (Knight et al. [14]).

In line with this hypothesis is the observation of Vinogradov [32] that red-list species more often have large genomes than non-endangered species. Gruner et al. [33] report slower root growth in species with large genomes, which may be one factor of risk for endangered plant species. The analyses by Bennett et al. [34] in weeds and nonweeds showed that weeds and especially the most aggressive weeds have smaller genomes than nonweeds, which indicates that competitive stress brings plants with larger genomes into a disadvantage.

Šmarda et al. [35] investigated differential seedling survival in a tetraploid *Festuca pallens* population with up to 1.189-fold genome size variation at constant chromosome number of $2n = 28$. In this case, during development of the plants those with the lowest and the highest C-values were eliminated under conditions of intraspecific competition

for resources. This case of stabilizing selection could be explained by a disadvantage for karyotypes with too many deletions or duplications and is of a different character than the selection phenomena observed under chemical or ecological stress.

Acknowledgments

The authors firstly thank Marina Dermastia and Tatjana Vidic for kindly supporting this investigation with plant material (*Anemone nemorosa*) and for valuable discussions and Jasna Dolenc Koce (Ljubljana) and Markus Koch (Heidelberg) for seeds of *Thlaspi praecox*. The authors also thank Friedrich Ehrendorfer, Irmgard Greilhuber, Josef Greimler, Walter Gutermann, Elvira Hörandl, Christiane König, Harald Niklfeld, and Hermann Voglmayr for providing material and for helping with the identification, and Martina Mittlböck and Harald Heinzel for the valuable discussion on data analysis.

References

- [1] B. J. M. Zonneveld, “New record holders for maximum genome size in eudicots and monocots,” *Journal of Botany*, vol. 2010, Article ID 527357, 4 pages, 2010.
- [2] C. A. Thomas Jr., “The genetic organization of chromosomes,” *Annual Review of Genetics*, vol. 5, pp. 237–256, 1971.
- [3] M. D. Bennett, “The duration of meiosis,” *Proceedings of the Royal Society of London, Series B*, vol. 178, pp. 277–299, 1971.
- [4] M. D. Bennett, “Nuclear DNA content and minimum generation time in herbaceous plants,” *Proceedings of the Royal Society of London. Series B*, vol. 181, no. 63, pp. 109–135, 1972.
- [5] T. R. Gregory, “The C-value enigma in plants and animals: a review of parallels and an appeal for partnership,” *Annals of Botany*, vol. 95, no. 1, pp. 133–146, 2005.
- [6] J. Greilhuber, J. Doležel, M. A. Lysák, and M. D. Bennett, “The origin, evolution and proposed stabilization of the terms ‘Genome Size’ and ‘C-Value’ to describe nuclear DNA contents,” *Annals of Botany*, vol. 95, no. 1, pp. 255–260, 2005.
- [7] C. A. Knight and J. M. Beaulieu, “Genome size scaling through phenotype space,” *Annals of Botany*, vol. 101, no. 6, pp. 759–766, 2008.
- [8] D. Francis, M. S. Davies, and P. W. Barlow, “A strong nucleotypic effect on the cell cycle regardless of ploidy level,” *Annals of Botany*, vol. 101, no. 6, pp. 747–757, 2008.
- [9] M. D. Bennett and H. Rees, “Induced and developmental variation in chromosomes of meristematic cells,” *Chromosoma*, vol. 27, no. 2, pp. 226–244, 1969.
- [10] D. Francis, *The Plant Cell Cycle and Its Interfaces*, Wiley-VCH, Weinheim, Germany, 2001.
- [11] R. S. Verma and M. S. Lin, “The duration of DNA synthetic (S) period in *Zea mays*: a genetic control,” *Theoretical and Applied Genetics*, vol. 54, no. 6, pp. 277–282, 1979.
- [12] R. S. Verma, “The duration of G₁, S, G₂, and mitosis at four different temperatures in *Zea mays* L. as measured with ³H-thymidine,” *Cytologia*, vol. 45, pp. 327–333, 1980.
- [13] J. Greilhuber, “Chromosomes of the monocotyledons (general aspects),” in *Monocotyledons: Systematics and Evolution*, P. J. Rudall, D. F. Cribb, and C. J. Humphries, Eds., pp. 379–414, Royal Botanic Gardens, Kew, Surrey, UK, 1995.
- [14] C. A. Knight, N. A. Molinari, and D. A. Petrov, “The large genome constraint hypothesis: evolution, ecology and

- phenotype," *Annals of Botany*, vol. 95, no. 1, pp. 177–190, 2005.
- [15] T. Vidic, J. Greilhuber, B. Vilhar, and M. Dermastia, "Selective significance of genome size in a plant community with heavy metal pollution," *Ecological Applications*, vol. 19, no. 6, pp. 1515–1521, 2009.
- [16] J. Greilhuber, E. M. Temsch, and J. Loureiro, "Nuclear DNA content measurement," in *Flow Cytometry with Plant Cells. Analysis of Genes, Chromosomes, and Genomes*, J. Doležel, J. Greilhuber, and J. Suda, Eds., pp. 67–101, Wiley-VCH, Weinheim, Germany, 2007.
- [17] M. D. Bennett and I. J. Leitch, "Plant DNA C-values Database," October 2005, <http://www.kew.org/cval/homepage.html>, release 4.0.
- [18] M. A. Fischer, K. Oswald, and W. Adler, *Exkursionsflora für Österreich, Liechtenstein und Südtirol*, Biologiezentrum der Oberösterreichischen Landesmuseen, Linz, Austria, 3rd edition, 2008.
- [19] H. Weiss-Schneeweiss, J. Greilhuber, and G. M. Schneeweiss, "Genome size evolution in holoparasitic *Orobanchaceae* and related genera," *American Journal of Botany*, vol. 93, no. 1, pp. 148–156, 2006.
- [20] E. M. Temsch and J. Greilhuber, "Genome size in Dipsacaceae and *Morina longifolia* (Morinaceae)," , in press, 2010. *Plant Systematics and Evolution*.
- [21] D. W. Galbraith, K. R. Harkins, J. M. Maddox, et al., "Rapid flow cytometric analysis of the cell cycle in intact plant tissues," *Science*, vol. 220, no. 4601, pp. 1049–1051, 1983.
- [22] F. J. Otto, H. Oldiges, W. Goehde, and V. K. Jain, "Flow cytometric measurement of nuclear DNA content variations as a potential in vivo mutagenicity test," *Cytometry*, vol. 2, no. 3, pp. 189–191, 1981.
- [23] J. Greilhuber and I. Ebert, "Genome size variation in *Pisum sativum*," *Genome*, vol. 37, no. 4, pp. 646–655, 1994.
- [24] J. Doležel, J. Greilhuber, S. Lucretti et al., "Plant genome size estimation by flow cytometry: inter-laboratory comparison," *Annals of Botany*, vol. 82, supplement A, pp. 17–26, 1998.
- [25] E. M. Temsch, J. Greilhuber, and R. Krisai, "Genome size in liverworts," *Preslia*, vol. 82, pp. 63–80, 2010.
- [26] J. Greilhuber and E. M. Temsch, "Feulgen densitometry: some observations relevant to best practice in quantitative nuclear DNA content determination," *Acta Botanica Croatica*, vol. 60, no. 2, pp. 285–298, 2001.
- [27] SAS, "The SAS statistics software," 2010, <http://www.sas.com/>.
- [28] B. J. M. Zonneveld, I. J. Leitch, and M. D. Bennett, "First nuclear DNA amounts in more than 300 angiosperms," *Annals of Botany*, vol. 96, no. 2, pp. 229–244, 2005.
- [29] H. A. Sparrow and J. P. Miksche, "Correlation of nuclear volume and DNA content with higher plant tolerance to chronic radiation," *Science*, vol. 134, no. 3474, pp. 282–283, 1961.
- [30] B. Drušković, *Uticaj zagađjenja sredine na genetske promene u biljnim populacijama*, Ph.D. thesis, University of Novi Sad, Novi Sad, Serbia, 1984.
- [31] C. A. Knight and D. D. Ackerly, "Variation in nuclear DNA content across environmental gradients: a quantile regression analysis," *Ecology Letters*, vol. 5, no. 1, pp. 66–76, 2002.
- [32] A. E. Vinogradov, "Selfish DNA is maladaptive: evidence from the plant Red List," *Trends in Genetics*, vol. 19, no. 11, pp. 609–614, 2003.
- [33] A. Gruner, N. Hoverter, T. Smith, and C. A. Knight, "Genome size is a strong predictor of root meristem growth rate," *Journal of Botany*, vol. 2010, Article ID 390414, 4 pages, 2010.
- [34] M. D. Bennett, I. J. Leitch, and L. Hanson, "DNA amounts in two samples of angiosperm weeds," *Annals of Botany*, vol. 82, supplement A, pp. 121–134, 1998.
- [35] P. Šmarda, P. Bureš, L. Horová, and O. Rotreklová, "Intrapopulation genome size dynamics in *Festuca pallens*," *Annals of Botany*, vol. 102, no. 4, pp. 599–607, 2008.