Theor Appl Genet (2002) 105:58–67
DOI 10.1007/s00122-002-0866-8

E. Potokina · F.R. Blattner · T. Alexandrova
K. Bachmann

AFLP diversity in the common vetch (*Vicia sativa* L.) on the world scale

Received: 20 August 2001 / Accepted: 19 October 2001 / Published online: 18 May 2002

Abstract The Vavilov Institute of Plant Industry (VIR) keeps a living seed collection of about 700 accessions of landraces and local cultivars of common vetch (*Vicia sativa* L.) that have been collected over a period of more than 50 years throughout the former USSR. Much of the material is available nowhere else. The collection of this economically important fodder crop is well adapted to the various growing regions of Russia and serves as a basis for the all domestic vetch breeding programs. Using AFLP as a DNA fingerprinting method we investigated 673 accessions from the VIR and compared their genetic variability with that of the worldwide vetch collection of the Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK), 450 accessions. The analysis is a first assessment of the intra-specific diversity of *V. sativa* stored ex situ on a scale of more than 1,000 accessions. Six primer combinations, which gave clear polymorphic amplification products with 96 test samples, were chosen from 111 primer combinations tested. The selected AFLP primers used to analyse the *V. sativa* intra-specific diversity resulted in 70 unequivocally recognizable polymorphic fragments. We found that all of the AFLP fragments generated can be detected with varying frequency throughout the entire distribution area of *V. sativa*. The difference in frequency of some AFLP fragments between the regions may amount to 90%. The arrangement of most of the accessions in all dendrograms reflects their geographical origin, with a differentiation between Russia, Western Europe, Turkey and Bulgaria, and the Mediterranean. The "Russian" genepool stored at the IPK is a limited and biased sample of the available diversity when compared to the material stored at the VIR. Approximately 10–15% of the accessions in each geographical group showed AFLP patterns that clustered with members of other groups. This appreciable overlap raises several questions: (1) to which degree is an AFLP pattern representative of the overall genetic similarity of the samples; (2) to which degree are samples collected at a site adaptively limited to that site? Since our data identify accessions with very similar AFLP patterns from very diverse geographic origins, a comparison of the agronomic performance of these accessions (possibly in the two regions) will provide important information for the utilization of ex situ germplasm collections.

Keywords AFLP · DNA fingerprinting · *Vicia* · Intra-specific diversity · Ex situ collections

Communicated by H.C. Becker

E. Potokina (✉)
Department of Taxonomy,
Vavilov Institute of Plant Industry (VIR),
Bolshaya Morskaya 42, St. Petersburg 19000, Russisa
e-mail: potokina@ipk-gatersleben.de
Tel.: +49-39482-5663, Fax: +49-39482-5595

F. R. Blattner
Department of Taxonomy,
Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK),
D-06466 Gatersleben, Germany

T. Alexandrova
Department of Legumes,
Vavilov Institute of Plant Industry (VIR),
Bolshaya Morskaya 42, St. Petersburg 190000, Russia

K. Bachmann
Department of Taxonomy,
Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK),
D-06466 Gatersleben, Germany

Present address:
E. Potokina, Department of Genebank,
Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK),
D-06466 Gatersleben, Germany

Introduction

Vicia sativa L. sensu stricto, the common vetch, an economicálly important fodder crop, is a member of the *V. sativa* aggregate, a closely related group of autogamous annual taxa. Some authors recognise eight species in this aggregate, others consider this aggregate to be one variable species, *V. sativa* sensu lato, with subspecies and varieties. The *V. sativa* aggregate seems most likely to have originated in southern Europe or Southwest Asia

(Maxted 1995). In the area of the former USSR, two members of the aggregate, *Vicia cordata* Wulf. and *Vicia macrocarpa* (Moris) Bertol., are absent, and *Vicia amphicarpa* Dorth. and *Vicia incisa* Bieb. reach the limit of their distribution areas, occurring as rare populations in some regions of Transcaucasia and the Crimea (Potokina 1997). Two most genotypically and phenotypically plastic species of the aggregate, *V. sativa* sensu stricto and *Vicia angustifolia* Reichard, have evolved rapidly away from the centre of origin, and are currently widespread in central and southern regions of the European part of Russia, in the Crimea, the Caucasus and Central Asia, occurring in Siberia and in the Far East as adventitious plants. These species are mostly present in the area of the former USSR as populations of weedy and cultivated forms.

Common vetch has been widely cultivated in Russia since ancient times. Some local traditional cultivars from southern central Russia and the Ural region are widely used (Stankevich and Repjev 1999). Each local cultivar usually consisted of five to eight botanical forms, which were well adapted to the local conditions and had a high productivity (Tupikova 1933). The traditional cultivars as well as wild and weedy populations of *V. sativa* have been collected over a period of more than 50 years by expeditions of the Vavilov Institute of Plant Industry (VIR). At present, the VIR keeps a living seed collection of about 700 accessions of landraces and local cultivars of common vetch from different geographical regions of Russia. This material serves as a basis for all vetch breeding programs in Russia. The effectiveness of the programs depends to a large extent on the genetic information available on the germplasm with which breeders work. In this study we use molecular fingerprint methods to characterize the vetch germplasm mantained at the VIR and at the German genebank at the Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK). We want to analyze the fingerprint patterns for signals indicating a taxonomic or geographic substructure and gain an impression of the effectiveness of the sampling methods and the genetic diversity present in the ex situ collections.

Materials and methods

The VIR collection

A total of 673 accessions of *V. sativa* from the VIR collection were analyzed. They included:

(1) cultivars in current use, which are indicated in the VIR passport data as Advanced Cultivars;
(2) landraces;
(3) farmers' old traditional cultivars being selected from local landraces. These cultivars usually have a local name, for example cultivar K-27793 was entered into the VIR collection in 1949 under the name "cv Saranskaya local" from the Saransk region (Mordoviya), where this cultivar is popular and widely cultivated;
(4) breeder resources being selected from local landraces by the regional breeding organisations;
(5) weedy populations.

Most of the accessions were true weedy populations, mainly from the Caucasus or the former Central Asian Republics. These are common weeds of cereal crops. They differ from cultivated and escaped forms of European Russia by their xeromorphic habit and narrower and shorter pods. The morphological difference is so evident that the forms were described as a separate subspecies *V. sativa* subsp. *linearifolia* Stankev. (Stankevich 1978). A special case are the weedy populations of *V. sativa* from the southeastern regions of the forest-steppe zone of Russia (Saratov and Penza regions). They are weeds of lentil crops, cultivated there from ancient times. As a result of the long-term process of selection, the forms with flattened seeds (having a size, shape and colour identical with that of lentil) are dominant among weedy *V. sativa* populations from this region. This form was also described as a separate taxon *V. sativa* var. *platysperma* Barul. (Barulina 1930).

The structure of the material analysed in terms of the kind of genetic resources is presented in Table 1. This table gives some information about the development of the VIR collection since 1931 and reflects the steps of the breeding process for vetch in Russia.

Geographical distribution of the material

An origin of the accessions in the VIR passport data is indicated by the name of the administrative province of the former USSR, known as *Oblast* (Region), where the accession was collected. Vetch germplasm from 64 administrative regions was present in the analysed material. Most of the regions can be assigned to natural eco-geographical areas. Thus, all the accessions under study were divided into 11 eco-geographical groups by their origin. Figure 1 shows these 11 areas with the number of accessions from each area included in this study. Each of the 11 groups can be characterized by the locally dominant ecotypes of *V. sativa*. The detailed description of vetch ecotypes has been developed by Leokene (1978) in her eco-geographical classification of common vetch as a result of many years of field trials with a large common vetch collection comprising 1,500 variety samples from 35 countries. Eight eco-geographical groups and 22 ecotypes were described.

The material from the VIR examined by us comes from the western half of the former USSR. We have applied the following eco-geographical groupings.

(1) Baltic area (Latvia, Lithuania, Estonia and the Kaliningrad region of Russia): 95 accessions. West European eco-geographical group: Baltic ecotype.

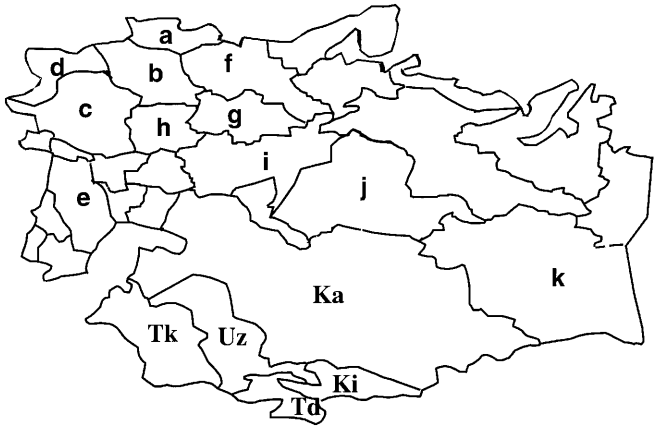

Fig. 1 Geographical distribution of the material analysed in the regions of the former USSR. Central Asian Republics: *Ka* – Kazakhstan; *Tk* – Turkmenia; *Uz* – Uzbekistan; *Td* – Tadzhikistan; *Ki* – Kirgizstan. Eco-geographical areas and number of the accessions from those areas included in the study: *a* Baltic, 98; *b* Byelorussia, 34; *c* Central Ukraine, 54; *d* Western Ukraine, 45; *e* Caucasus, 89; *f* Russia-North, 37; *g* Russia-Centre, 105; *h* Russia-South, 106; *i* Russia-South East, 43; *j* Russia-Ural, 34; *k* Russia-Siberia, 30

Table 1 Structure of the VIR collection of *V. sativa* accessions from the area of the former USSR in 1931–1997

Year	Weedy populations %	Landraces %	Advanced cultivars %	Local cultivars %	Breeder resources %	Number of accessions
1931	10.4	78.4	0	0	11.2	134
1950	5.1	79.9	4.1	3.1	7.7	413
1970	7.5	65.7	8.8	10.1	7.8	601
1997	13.4	58.4	14.8	9.6	7.0	676

Table 2 Accessions deposited at IPK included in the analysis

Taxon	Origin of accessions	Number of accessions analysed	Source	Year of receiving
V. sativa	Greece	107	Balkans, expedition	1941–1942
	Italy	27	Italy, expedition	1950
			South Italy, expedition	1980–1982
	Spain	61	Agriculture Inst., Cordoba	
			Spain, expedition	1978
	Turkey	19	M. Klinkowski, Anatolien, expedition	
	Iran	22	H. Kuckuck, FAO collecting in Iran	1952–1954
	Bulgaria	26	Agriculture Inst., Sofia	
	Germany	21	Plant Breed. Station, Halle	
			Inst. of Plant Industry, Weihenstephan	
	Poland	23	IHAB, Warshawa	
			Landraces collecting in Poland	1976
	Czechoslovakia	26	Landraces collecting in CSSR	1977, 1981
			Inst. of Plant Industry, Praga	
	Hungary	13	NIAVT Tapioszele	
	Yugoslavia	9	Agriculture Inst., Banja Luka	
	former USSR	43	Agriculture Inst., Sofia, Bulgarien	
			Vavilov Inst. of Plant Industry, Russia	
			Georgia, Caucasus, expedition	1981–1984
V. macrocarpa	Italy, Sicilia	5	Seed Experimental Station, Guyancourt, France	
V. cordata	Turkey	8	M. Klinkowski, Anatolien, expedition	
	Greece		Balkan, expedition	1941
V. segetalis	Turkey	10	M. Klinkowski, Anatolien, expedition	
	Greece		Balkan, expedition	1942
	Iran		H. Kuckuck, FAO collecting in Iran	1952–1954
	India		Tibet, expedition	1938–1939
V. angustifolia	the Netherlands	5	Arboretum Wageningen	
	Sweden		Bot. Gard. Univ., Lund	
	Finland		Bot. Gard. Univ., Helsinki	
	Portugal		Bot. Gard. Univ., Porto	
	Denmark		Bot. Gard. Univ., Kopenhagen	
	Halle		Bot. Gard. Univ., Halle	
V. pilosa	Crimea	1	Nikita Bot. Gard. Yalta, UdSSR	
V. amphicarpa	Turkey	2	Bot. Lab. Univ. Leicester, England	
V. incisa	Bulgaria	2	Forestry Inst., Sofia, Bulgarien	
	Slovakia		Czechoslovakia, expedition	1977
V. grandiflora var.grandiflora	Unknown	1	from Prof. Yamamoto, Kagawa Univ., Japan	
V. grandiflora var. *kitaibeliana*	Yugoslavia	2	Bot. Gard. Univ., Ljubljana	
	Hungary		Inst.of Plant Science, Budahalasz	

(2) Byelorussia: 34 accessions. Middle Russian eco-geographical group: Byeloruss ecotype.
(3) Central Ukraine: 54 accessions. Middle-Russian eco-geographical group: East Ukraine ecotype.
(4) Western Ukraine: 45 accessions. Middle-European eco-geographical group: West Ukraine ecotype.
(5) Caucasus: 87 accessions. Caucasian eco-geographic group: Svanetia, Dzhavakhet, Armenia ecotypes.
(6) Russia-North: 37 accessions. Middle-Russian eco-geographical group: North-west ecotype.
(7) Russia-Centre: 105 accessions. Middle Russian eco-geographical group: Kursk ecotype, North-west ecotype.
(8) Russia-South: 106 accessions. Middle-Russian eco-geographical group: Kursk ecotype.
(9) Russia-South-East: 43 accessions. South European eco-geographical group: South-east ecotype.

Table 3 List of primers and adapters used

Primer/adapters		Sequences
*Mse*I adapter		5′-GACGATGAGTCCTGAG-3′ 3′-TACTCAGGACTCAT-5′
M00 (universal primer)		GATGAGTCCTGAG TAA
*Mse*I + 1 primer	M02	M00 + C
*Mse*I + 3 primers	M47, M48, M49, M50, M51, M52, M53, M54,M55, M56, M57, M58, M59	M00 + CAA, +CAC, +CAG, +CAT, +CCA, +CCC, +CCG, +CCT, +CGA, +CGC, +CGG, +CGT, +CTA
*Mse*I + 1 primer	M03	M00 + G
Mse I + 3 primers	M63, M64, M65, M66, M67, M68, M69, M70, M71, M72, M73, M74, M75, M76, M77, M78	M00 + GAA, +GAC, +GAG, +GAT, +GCA, +GCC, +GCG, +GCT, +GGA, +GGC, +GGG, +GGT, +GTA, +GTC, +GTG, +GTT
*Eco*RI adapter		5′-CTCGTAGACTGCGTACC-3′ 3′-CTGACGCATGGTTAA-5′
E00 (universal primer)		GACTGCGTACCAATTC
*Eco*RI + primer	E01	E00 + A
*Eco*RI + 3 primers	E36, E37, E38, E39, E40, E41	E00 + ACC, +ACG, +ACT, +AGA, +AGC, +AGG

(10) Russia-Ural: 34 accessions. Middle Russian eco-geographical group: North-east ecotype.
(11) Russia-Siberia: 30 accessions. Nothern eco-geographical group and Middle Russian eco-geographical group: North-east ecotype.

IPK collection

Four hundred and fifteen accessions of *V. sativa* deposited at the IPK were analysed in comparison with 34 accessions of seven other species of the *V. sativa* aggregate and three accessions of *Vicia grandiflora* L. as an outgroup (Table 2). *V. sativa* accessions comprise mainly populations of landraces collected in Europe and Asia or acquired through a worldwide exchange of seeds with plant breeding, agricultural and horticultural institutions. The accessions were not assigned to a specific ecotype and geographical distribution was represented by the country of origin.

AFLP analysis

AFLPs (amplified fragment length polymorphism, Vos et al. 1995) were used for DNA fingerprinting. Seed material of 1,122 accessions was grown in pots at the IPK, Gatersleben, Germany. DNA was isolated with the NucleoSpin Plant kit (Macherey-Nagel, Düren, Germany) according to the instructions of the manufacturer. Young leaves from 8 to 15 individuals were used to extract bulk DNA for each accession. The AFLP technique of Vos et al. (1995) was employed with the following modifications: 4 µl of genomic DNA (150 ng) was added to 6 µl of the solution containing 0.5 pmol of *Eco*RI-adapters, 5 pmol of *Mse*I adapters, 1 µl of T4 DNA Ligase Buffer (10×), 10 mM of NaCl, 80 ng/µl of BSA. Next, 2 µl of the Restriction–Ligation Mix containing 1 U of *Mse*I, 4 U of *Eco*RI, 4 U of T4 DNA-ligase, 0.2 µl of T4 Ligase Buffer (10×), 50 mM of NaCl, 100 ng/µl of BSA was added, and the incubation was performed overnight at room temperature. After ligation, the reaction mixture was diluted to 50 µl with water, and used as template for the following amplification step as described by Vos et al. (1995). Adapters, *Mse*I site primers and the *Eco*RI site primers used are listed in Table 3. *Eco*RI primers used in the selective amplification were fluorescence labelled. The amplification products of each three differently labelled selective amplifications were pooled and electrophoresed on sequencing gels on a Perkin-Elmer ABI 377 automatic sequencer together with a size standard (Genescan-500 Rox, Applied Biosystems). To allow exact determination of fragments larger than 500 base pairs (bp) the size standard was supplemented with five additional DNA fragments ranging from 568 to 812 bp. Gels were analyzed with the GeneScan software (version 3.0; Perkin-Elmer ABI).

Table 4 Evaluation of 111 primer combinations based on 12 *V. sativa* accessions

Primer combination	E36	E37	E38	E39	E40	E41
M47	Poor	Fair	Poor	Fair	Fair	Fair
M48	Fair	Good	Good	Fair	Good	Fair
M49	Good	Fair	Fair	**Good**[b]	Fair	Fair
M50	Poor	Poor	Poor	Poor	Fair	Poor
M51	Poor	Poor	Fair	Fair	Poor	Poor
M52	Fair	**Good**	Fair	Fair	Good	**Good**
M53	Fair	Fair	**Good**	Good	Fair	Good
M54	**Good**	Good	Fair	Good	Fair	Poor
M55	nd[a]	nd	nd	**Good**	Fair	Poor
M56	nd	nd	nd	**Good**	**Good**	Poor
M57	nd	nd	nd	Fair	**Good**	**Good**
M58	nd	nd	nd	**Good**	**Good**	**Good**
M59	nd	nd	nd	Poor	Fair	**Good**
M63	Fair	Fair	Fair	nd	nd	nd
M64	Fair	Fair	Good	nd	nd	nd
M65	Good	Fair	Fair	nd	nd	nd
M66	Poor	Poor	Poor	nd	nd	nd
M67	Poor	Fair	Fair	nd	nd	nd
M68	**Good**	Good	**Good**	nd	nd	nd
M69	Good	Fair	Good	nd	nd	nd
M70	Fair	Good	Fair	nd	nd	nd
M71	nd	nd	nd	Poor	Poor	Poor
M72	nd	nd	nd	Fair	Fair	Fair
M72	nd	nd	nd	Fair	Poor	Fair
M74	nd	nd	nd	Fair	Poor	Poor
M75	nd	nd	nd	Fair	Poor	Poor
M76	nd	nd	nd	Poor	Fair	Poor
M77	nd	nd	nd	Poor	Poor	Poor
M78	nd	nd	nd	Poor	Poor	Poor

[a] nd – not determined
[b] Bold – also screened with 96 accessions

Initially, 12 accessions were chosen to test the variation of primer combinations. With these accessions the polymorphism rates and the total number of peaks with the 111 primer combinations were evaluated (Table 4). The most-useful primer combinations were considered those having the highest polymorphism rate that also generate a reasonable number of clearly detectable total fragments. Using the results of the evaluation of 111 primer combinations based on 12 *V. sativa* accessions, 15 selected primer combinations were further screened with 96 representative accessions. As a result, the six most-polymorphic primer combinations

(E37-M52; E38-M53; E39-M49; E40-M56; E40-M58; E41-M57), producing clearly readable peaks, were selected for the subsequent analyses. AFLP analysis was conducted twice and only fragments were used that were unambiguously reproducible. Gels were scored for the presence (1) or absence (0) of peaks with the Genotyper software (version 2.1, Perkin-Elmer ABI) and the values were used to compile binary data matrices. A similarity index for all possible pairwise comparisons between accessions was calculated using Nei's genetic distances (Nei and Li 1979) and the Dice coefficient (Dice 1945). The similarity matrix was used to generate a phenogram by the unweighted pair group method with arithmetic averages (UPGMA) using NTSYS-pc version 1.70. Principal coordinate analysis (PCA) was conducted using the DOUBLE CENTER and EIGENVECTOR options of NTSYS-pc.

Results

The size of AFLP fragments was in a range from about 30 bp to 600 bp, but only fragments between 100 and 400 bp were taken into account to avoid scoring problems due to excess primer peaks near the front of the electrophoresed fragments and a decreasing signal for fragments longer than 450 bp. A total of 70 unequivocally recognisable polymorphic fragments were obtained with the six selected primer combinations within *V. sativa*. The number of polymorphic fragments generated by each primer varied from 6 to 16 among accessions of *V. sativa*.

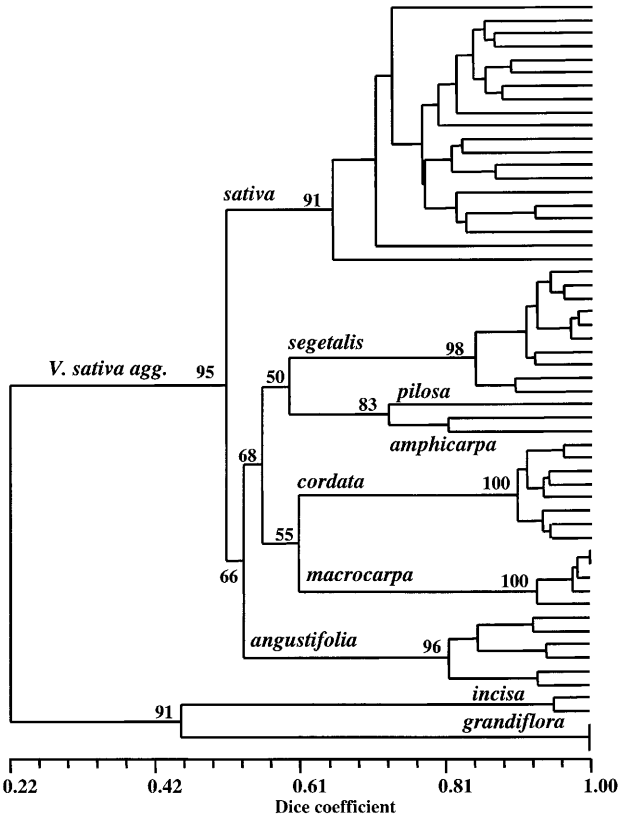

Fig. 2 UPGMA phenogram generated from AFLP analysis of 57 accessions from eight species of the *V. sativa* aggregate. Bootstrap value with UPGM clustering after 500 bootstrap re-samples

V. sativa intraspecific diversity relative to the *V. sativa* aggregate

Fourty four of 70 AFLP fragments occurring within *V. sativa* were also observed among the other species of the aggregate. A total of 103 polymorphic AFLP fragments were revealed among seven species of the *V. sativa* aggregate and with *V. grandiflora* L. as outgroup. Figure 2 shows the UPGMA phenogram generated from the Dice distance matrix over all the species studied. *V. sativa* senso stricto (s.str.) is represented in the UPGMA tree by accessions with the most-different AFLP patterns. The lowest value of genetic similarity found for intraspecific diversity within *V. sativa* is 0.64. Genetic similarities among the taxa of the aggregate ranged from 0.22 (between *V. incisa* and other taxa of the aggregate) up to 0.59 (between *V. cordata* and *V. macrocarpa*). The close relationship between *Vicia pilosa* and *Vicia amphicarpa* (0.72) supports the suggestion that the two taxa represent the same species (Potokina 1997). The phenogram reflects the relationship between the species and corresponds to the previous taxonomic interpretation of the aggregate (Potokina et al. 2000). Although the taxonomical status of members of the *V. sativa* aggregate is still debatable, all of them can be easily recognized by several specific AFLP fragments. Due to this specificity of the AFLP patterns, any accession with unclear taxonomical identification can be exactly identified by the molecular fingerprinting method. As an example, 11 accessions of the *V. sativa* aggregate from the IPK collection having no precise taxonomical identification were analysed. Six of them were defined as *V. sativa* s. str. (V-1668, V-1797, V-1798, V-1800, V-1834, V-1894) and four accessions belong to *V. macrocarpa* (Moris) Bertol. (V-1897, V-2614, V-845, V-2065).

Geographical differentiation within the *V. sativa* gene pool

Sixty five out of 70 AFLP fragments generated within *V. sativa* can be detected with varying frequency throughout the entire distribution area of the species. Five fragments occurred as unique peaks in several accessions of the Ukraine, Slovakia, Spain and Greece. We found that the different geographical regions could be characterized by frequency compositions of AFLP fragments rather than by the presence of "private" AFLP fragments. The difference in the frequency of some AFLP fragments between the regions may amount to 90% (Table 5). The degree of similarity between the frequency compositions of the geographical regions seems directly related to the proximity between them. The UPGMA tree generated from Nei's genetic distances of the AFLP fragment frequencies among 22 geographical regions is shown in Fig. 3. The correlation between genetic and geographic distances is evident.

Table 5 Frequencies of the AFLP fragments generated by the E41-M57 primer combination within the *V. sativa* gene pool from 22 world regions

Region	Number of accessions	AFLP fragments (bp)												
		108.68	109.86	113	117.11	161	178	180	202	268.9	293	319	368	369
Baltic	98	0.938	0.255	1.000	0.684	0.061	0.765	0.877	0.071	0.633	0.020	0.786	0.020	0.245
Byelorussia	34	0.911	0.471	1.000	0.706	0.147	0.912	0.824	0.147	0.735	0.058	0.882	0	0.294
Central Ukraine	54	0.926	0.296	0.963	0.833	0.167	0.833	0.907	0.333	0.592	0.018	0.778	0	0.259
Western Ukraine	45	0.956	0.378	0.956	0.844	0.156	0.756	0.933	0.200	0.689	0.022	0.822	0.022	0.356
Caucasus	89	0.876	0.618	0.640	0.528	0.124	0.876	0.584	0.022	0.427	0.011	0.449	0.011	0.101
Russia, North	37	0.889	0.278	0.944	0.861	0.167	0.944	0.833	0.028	0.694	0.056	0.972	0	0.209
Russia, Centre	105	0.962	0.429	0.895	0.829	0.152	0.886	0.867	0.105	0.610	0.028	0.857	0.038	0.209
Russia, South	106	0.915	0.387	0.821	0.802	0.123	0.868	0.802	0.151	0.557	0.028	0.839	0	0.321
Russia, South-East	43	0.860	0.419	0.744	0.605	0.349	0.860	0.697	0.186	0.628	0	0.744	0.023	0.326
Russia, Ural	34	1.000	0.441	0.676	0.765	0.206	0.882	0.852	0.059	0.676	0.029	0.853	0	0.441
Russia, Sibiria	30	0.933	0.200	0.933	0.700	0.133	0.867	0.867	0	0.733	0	0.867	0	0.267
Poland	23	0.782	0.130	0.652	0.522	0.130	0.696	0.783	0	0.261	0.435	0.565	0	0.478
Germany	21	0.900	0.150	0.700	0.450	0.150	0.800	0.950	0	0.500	0	0.700	0	0.300
Czechoslovakia	26	0.846	0.269	0.615	0.615	0.077	0.538	0.692	0	0.692	0.192	0.808	0	0.038
Greece	107	0.604	0.405	0.707	0.113	0.038	0.915	0.906	0	0.717	0.018	0.585	0.235	0.160
Spain	61	0.836	0.689	0.919	0.081	0.049	0.984	0.984	0	0.885	0	0.819	0.082	0.065
Italy	27	0.815	0.629	0.815	0.407	0.074	0.926	0.778	0	0.815	0.037	0.852	0.074	0
Bulgaria	26	0.880	0.200	0.600	0.280	0.280	1.000	0.440	0	0.760	0	0.200	0.120	0
Yugoslavia	9	1.000	0	0.889	0.555	0	0.889	0.333	0	0.667	0.111	0.889	0.222	0.111
Hungary	13	1.000	0.461	0.615	0.615	0.154	0.461	0.692	0	0.769	0.231	0.923	0	0.231
Turkey	19	0.842	0.526	0.210	0.368	0.052	0.684	0.632	0	0.737	0	0.579	0	0.105
Iran	22	0.454	0.955	0.909	0.727	0.045	0.773	0.364	0	1.000	0	0.273	0.591	0

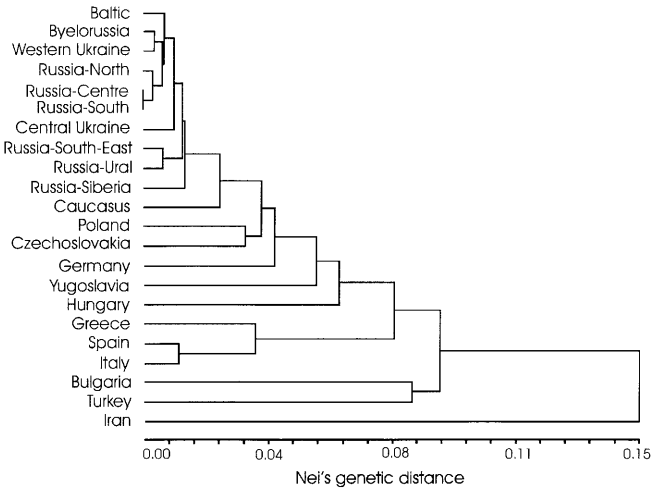

Fig. 3 UPGMA phenogram generated from the AFLP fragment frequencies within 22 geographical regions of the *V. sativa* distribution area

V. sativa gene pools of European countries are closely related to each other and to those from the Mediterranean basin. Populations of Turkey and Bulgaria clustered separately on the UPGMA tree. The most-distant cluster was formed by the Iranian accessions.

Comparison of the genetic composition of the accessions

The similarity matrix was calculated for all possible pairwise comparisons between the accessions using the Dice coefficient, and then was used to generate an UPGMA phenogram and to perform a Principal Coordinate Analysis. At first glance, the PCA plot (Fig. 4) does not reveal any obvious structure. When the major geographical groups of the UPGMA analysis are highlighted, a clear geographical component in the PCA plot becomes obvious with Russia and neighbouring states, Western Europe, Turkey and Bulgaria, and the Mediterranean accessions contributing characteristic components of the total variation. The PCA plot shows a good relation with the frequency UPGMA phenogram (Fig. 3) except for the Iranian accessions which are dispersed throughout the Turkey/Bulgarian and Mediterranean clusters (not separately shown, see open circles in Figs. 4C and D). Figure 4 shows that about 10–15% of the accessions in each geographical group showed AFLP patterns that clustered with members of other groups. Thus, having the same geographical origin the accessions can be characterized by the clearly different genetic compositions. For instance, 69 accessions from the VIR collection fall outside of the "Russian" cluster showing closer relations to the accessions from Turkey/Bulgaria or the Mediterranean cluster.

Comparison of genepools of "Russian" landraces and cultivars maintained at VIR and IPK

Fourty three accessions of the IPK collection originating from the area of the former USSR have mainly been obtained:

(1) during collection missions through the Caucasus (10 accessions),

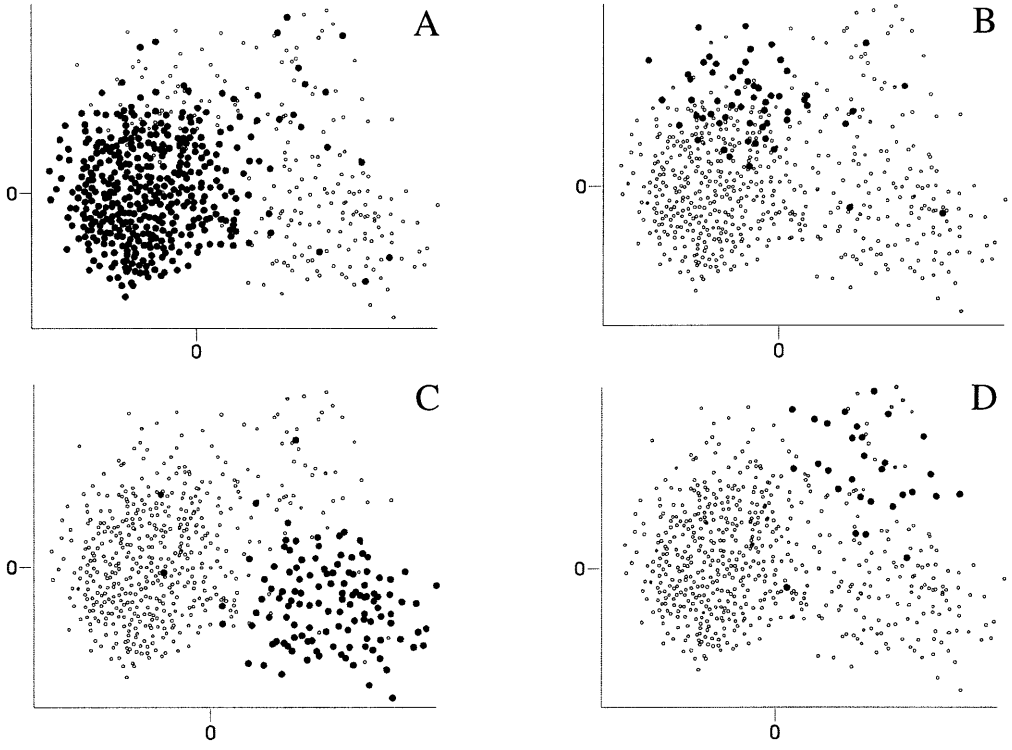

Fig. 4A–D Plot of the first two principal co-ordinates computed from the Dice similarity matrix among 1,088 accessions of *V. sativa* from the VIR and IPK genebanks. **A** – former USSR; **B** – Poland, Chechoslovakia, and Germany; **C** – Greece, Spain, and Italy; **D** – Turkey and Bulgaria

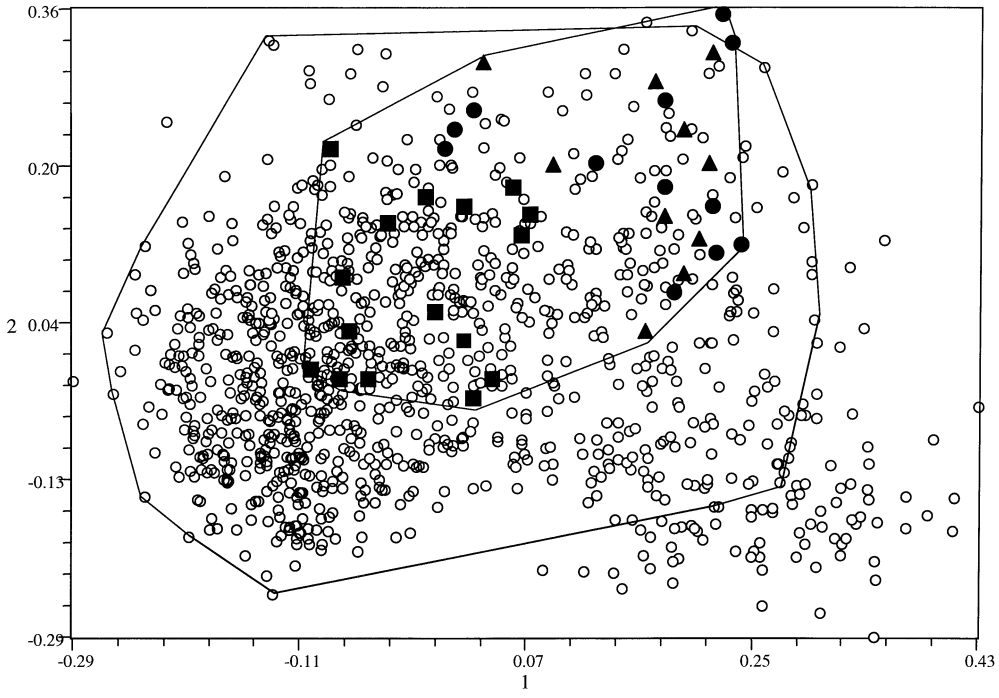

Fig. 5 Comparison of genetic diversity of "Russian" vetch landraces and cultivars stored at IPK (*inner circle*) in the scale of the VIR collection. IPK accessions: ▲ – from the Caucasus; ● – from the Institute of Agriculture, Sofia, Bulgaria; ■ – from the Vavilov Institute of Plant Industry (VIR), St. Petersburg, Russia

(2) via the Institute of Agriculture, Sofia, Bulgaria (12 accessions), and
(3) from the Vavilov Institute of Plant Industry (VIR), St. Petersburg, Russia (15 accessions).

Figure 5 highlights these accessions within the PCA plot of the total vetch sample. The genepool of "Russian" landraces and cultivars from the VIR collection (Fig. 4A) was found to be more diverse, and the IPK collection essentially constitutes a biased sample of the total variety. There is an overall bias of the "Russian" genepool mantained at IPK in the direction of the Turkey/Bulgarian cluster (Fig. 4D). This is mainly due to accessions received via Bulgaria, where the material had been reproduced (closed circles in Fig. 5). As shown in Fig. 5, the material acquired directly from the VIR reflects more closely the sampling of the VIR collection, but also is not truly representative for the "Russian" component of the VIR collection. Accessions from the Caucasus (closed triangles in Fig. 5) clustered outside of the "Russian" range of variation. This was also true for 26% of the Caucasus accessions from the VIR collection.

Among the germplasm received from the VIR and currently maintained at the IPK, five accessions carry the precise cultivar names: Bogorodickaja 1, Cesisskaja Mestnaja, Isumrudnaja, Kamalinskaja A-611, Saranskaja Mestnaja. We compared the accessions with those from the VIR with the same names, assuming that they should be identical. Since a bulk analysis was conducted, it was expected that two independent analyses of the same accession with different sets of 10 to 15 plants could show some difference in the bulk AFLP pattern of the accession. Therefore a control experiment was performed with four accessions of the VIR collection (K-29453, K-25929, K-35203, K-3138). The Dice similarity coefficient was calculated for the AFLP patterns of two independent bulks extracted from each accession. It varied between 0.946 and 0.983 (a complete correspondence of AFLP patterns in terms of Dice coefficient means 1.0). A similar comparison of the six identically named accessions from the VIR and the IPK showed significantly larger differences (0.780–0.882).

Discussion

The *V. sativa* aggregate is a classical sample of a complex of well-separated taxa and derived forms, representing various degrees of phylogenetic divergence (Hanelt and Mettin 1989; Potokina et al. 2000). At the DNA level (RAPD, AFLP) their divergence is more evident (Potokina et al. 1999, 2000; Fig. 2 this study) than the divergence in morphological differences.

Here, we have estimated the intra-specific diversity of the widely distributed species *V. sativa* based on highly representative material (1,088 accessions from throughout the distribution range) and compared this diversity with that of closely related, but phylogenetically separate taxa of the *V. sativa* aggregate. AFLP patterns could clearly distinguish the various species of the *V. sativa* aggregate and, within the limits of the restricted samples from the various species apart from *V. sativa*, we have found potentially diagnostic AFLP bands. At least within the comparative framework of this study, the AFLP patterns are sufficiently informative to permit the assignment of accessions with uncertain taxonomical status to species within the aggregate (see above).

The situation is profoundly different within the strictly defined species *V. sativa*, where no clear intra-specific

patterns were found. Eco-geographical groups within *V. sativa* could only statistically be circumscribed by differences in frequencies of dominant alleles in various parts of the area rather than by the presence of area-specific alleles. Almost all alleles can be found throughout the whole area of distribution of the species, but with different probabilities (Table 5). The lack of clear intra-specific differentiation within *V. sativa* might be explained by three biological reasons. (1) The species has undergone a severe reduction in the number of individuals sometime in the past (bottleneck), followed by a relatively fast spread into its extant distribution area, without regaining clear differences up to now. (2) The rather high probability of cross-pollination (up to 10%; Hanelt and Mettin 1989) results in gene flow between populations of the species, which prevents the development of clearly differentiated groups within the species. (3) Breeding efforts together with trading of seed materials from one area to the other could mask or even eliminate area differences when cross-pollination between local and introduced material occurs. Possibly all three possibilities contribute to the pattern. Domestication of the wild progenitors of crop species normally results in a severe reduction of genetic variation and a fast spread of the cultivated plants throughout the area suitable for cultivation, where gene flow from the crop to wild or weedy local populations mixes their gene pools. Therefore different frequencies of alleles, as found in *V. sativa*, could reflect the weak eco-geographical differentiation within a species (Fig. 3). As a result, each eco-geographical region of the *V. sativa* distribution is only characterised by a set of frequent alleles, accompanied by some rare "foreign" alleles (Fig. 4).

Apart from biological reasons, methodological properties of the marker system used also contribute to the results we obtained. The analysis of an unusually large sample of material (1,088 accessions from one species) with a variable history of ex situ conservation has revealed features that would not have been detected with a smaller sample or with material with a common ex situ history. The results raise questions both about the interpretation of AFLP data and about the methods of ex situ conservation. Since the questions about the molecular method have a bearing on the questions about ex situ storage and vice versa, it will need additional information to separate these two factors. The importance of our data therefore lies also in the strikingly unexpected features and the potential consequences of the questions they raise.

In the following discussion we have to realize that, here, we are dealing specifically with the variation in AFLP amplification patterns. We do not know to which degree these are representative for the overall genetic similarity among the accessions. In the published literature, much more emphasis has been placed on applying the AFLP method than to investigating what precisely it reveals. The best information comes from investigations in which AFLP fragments have been cloned and sequenced or used as probes for Southern analysis. As expected, AFLP fragments can be amplified from single-copy or repetitive-DNA fractions (Meksem et al. 1995; Cho et al. 1996). Cloning AFLP fragments usually involves pre-selection of bands, and it will need additional information to determine the relative contribution of repetitive and single-copy DNA in AFLP patterns. There are indications that AFLP bands preferably are amplified from repetitive DNA and therefore are a biased sample of heritable polymorphisms. Since AFLPs have usually fixed positions on genetic maps, they could represent individual variants of degenerate repetitive elements. On genetic maps of various plant species, AFLP markers often are unevenly spaced, and their clustering differs from that of single-copy RFLP markers (Becker et al. 1995; Maheswaran et al. 1997). In the linkage map of *Arabidopsis thaliana*, AFLP markers were found to cluster more densly around the centromeric regions (Alonso-Blanco et al. 1998). Another indication of the biased distribution of AFLP fragments is their frequency distribution in larger samples of plants. These are regularly bimodal with a large fraction of fragments present only in one or a few specimens and a smaller peak of fragments that are present in (nearly) all of the plants. An example is the distribution in accessions of *A. thaliana* (Miyashita et al. 1999). We have found numerous other examples in our own work (unpublished data). Figure 6 shows the corresponding distribution in our 1,095 *Vicia* accessions. If many of the AFLP fragments only occur in one or a few plants in a large sample, this points towards a high mutation rate and a low information content of a large fraction of the AFLP bands. Such rare polymorphisms would add random noise when AFLP patterns are analysed for overall similarity, and there is a possibility that this highly variable fraction changes quickly with time (and behaves differently in plant species depending on the amount and structure of their repetitive DNA fractions). This could explain some of the results of this study.

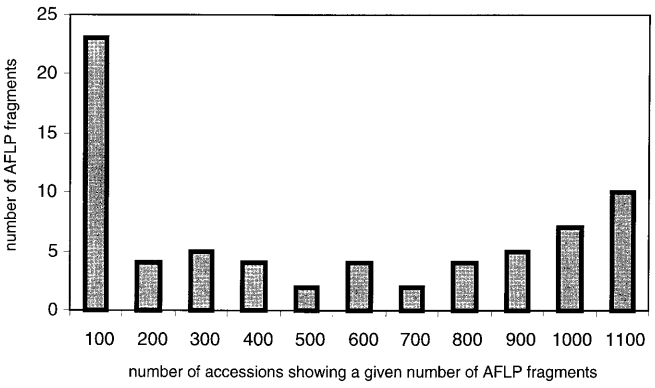

Fig. 6 Frequency of occurrence of 70 AFLP bands among 1,088 accessions of *V. sativa* s.str. Note that the largest group consists of AFLP bands that occur in less than 10% of all accessions

Since the accessions of *V. sativa* are not genetically uniform and bulk analyses were made, we cannot expect absolutely identical AFLP patterns even from repeated

samples from one accession. Comparative analyses of five accessions maintained both at the VIR and at the IPK shows that they differed significantly more (Dice similarity 0.882–0.780) than repeat determinations on accesions at one site (0.946–0.983). The history of these accessions is documented. On average, they had been reproduced three times at the VIR before they were sent to Gatersleben. Since then, they have been reproduced about six times at the VIR and at Gatersleben so that the samples that we compared are 12 reproduction cycles apart. It is possible that the decreased similarity in AFLP patterns is the result of sampling variation and/or the evolution of the AFLP patterns in the nearly 40 years since the VIR and Gatersleben lots of the same accessions were separated. It is also disturbing that the "Russian" sample of former VIR accessions at the IPK is not representative of the "Russian" sample of the VIR. Here we may also deal with an original sampling bias (not a conscious selection of accessions, as far as we can determine) and/or directed evolution of the AFLP pattern during local reproduction. An even stronger indication for such an effect comes from the VIR accessions that reached the IPK via Bulgaria, where they had been reproduced locally.

These observations suggest two possibilities, not necessarily exclusive, that need to be considered. One is that (a large fraction of) AFLPs undergo very fast evolution by mutation so that each individual is likely to harbor some unique AFLP fragments. The other is that handling of genetically polymorphic accessions during reproduction may involve severe sampling effects and inadvertent selection. Such selection might not directly affect the AFLP polymorphisms but be a hitchhiking effect of linked adaptively effective genes.

In order to separate these effects, an experimental study of the rate of AFLP evolution under the conditions of seed storage and reproduction will be necessary. With the magnitude of this effect known, the potentially much more disturbing effects of selection during seed reproduction can be studied independently. The results have a direct and important bearing on the practice of ex situ seed storage and the analysis of genetic integrity with AFLPs. Our observations are a strong reminder that the magnitude of genetic variation varies very much with the fraction of the genome that is examined, and variation in any one marker system may not be representative for agronomically significant variation. While these observations strengthen the use of AFLPs for the assay of genetic identity in clonal and inbred accessions, they show that the method can not yet be routinely applied as a meaningful measure of overall genetic differences or similarities. It will be very interesting to check morpho-physiological characteristics for some of the accessions that on the basis of their AFLP pattern cluster with material of another region of origin in order to determine how reliably AFLP patterns reflect local adaptation or distribution history.

Acknowledgments We are grateful to the International Plant Genetic Resources Institute who provided financial support for the study through the award of a Vavilov-Frankel Fellowship 1999 to E.P. Thanks also go to the Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany, for support during the Fellowship of E.P., and to the Vavilov Institute of Plant Industry (VIR), St. Petersburg, Russia, for their permission to undertake a fellowship in Germany.

References

Alonso-Blanco C, Peeters AJM, Koorneef M et al. (1998) Development of an AFLP based linkage map of Ler, Col, and Cvi *Arabidopsis thaliana* ecotypes and construction of a Ler/Cvi recombinant inbred population. Plant J 14:259–271

Barulina EI (1930) Lentils of U.S.S.R. and other countries (in Russian). Bull Appl Bot Genet Plant Breed 40

Becker J, Vos P, Kuiper M, Salamini F, Heun M (1995) Combined mapping of AFLP and RFLP markers in barley. Mol Gen Genet 249:65–73

Cho YG, Blair MW, Parnaud O, McCouch SR (1996) Cloning and mapping of variety-specific rice genomic DNA sequence: amplified fragment length polymorphisms (AFLP) from silver-stained polyacrylamide gels. Genome 39:373–378

Dice LK (1945) Measures of the amount of ecological associations between species. Ecology 26:297–302

Hanelt P, Mettin D (1989) Biosystematics of the genus *Vicia* L. (Leguminosae). Annu Rep Ecol Syst 20:199–223

Leokene LV (1978) Eco-geographical classification of common vetch varieties. Bull Appl Bot Genet Plant Breed 63:108–122

Maheswaran M, Subudhi PK, Nandi S et al. (1997) Polymorphism, distribution, and segregation of AFLP markers in a doubled haploid rice population. Theor Appl Genet 94:39–45

Maxted N (1995) An eco-geographical study of *Vicia* subgenus *Vicia*. Systematic and eco-geographic studies on crop genpools. 8. International Plant Genetic Resources Institute, Rome, Italy

Meksem K, Leister D, Peleman J, Zabeau M, Salamini F, Gebhardt C (1995) A high-resolution map of the vicinity of the R1 locus on chromosome V of potato based on RFLP and AFLP markers. Mol Gen Genet 249:74–81

Miyashita NT, Kawabe A, Innan H (1999) DNA variation in the wild plant *Arabidopsis thaliana* revealed by amplified fragment length polymorphism analysis. Genetics 152:1723–1731

Nei M, Li W (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA 76:5269–5273

Potokina EK (1997) *Vicia sativa* L. aggregate (Fabaceae) in the flora of former USSR. Genet Resour Crop Evol 44:199–209

Potokina E, Tomooka N, Vaughan DA, Alexandrova T, Xu RQ (1999) Phylogeny of *Vicia* subgenus *Vicia* (Fabaceae) based on analysis of RAPDs and RFLPs of PCR-amplified chloroplast genes. Genet Resour Crop Evol 46:149–161

Potokina EK, Vaughan DA, Eggi EE, Tomooka N (2000) Population diversity of the *Vicia sativa* agg. (*Fabaceae*) in the flora of the former USSR deduced from RAPD and seed protein analyses. Genet Resour Crop Evol 47:171–183

Stankevich AK (1978) On clarification of *Vicia sativa* L. taxonomy. Bull Vavilov Inst Plant Industry 81:3–11

Stankevich AK, Repjev SI (1999) Vetch. In: Flora of cultivated plants, vol. 4 (in Russian). Sanct-Petersburg SSC-VIR

Tupikova AY (1933) Vetch. In: Plant industry of U.S.S.R. (in Russian) Leningrad, Moscow, U.S.S.R. pp 451–463

Vos P, Hogers R, Bleeker M, Reijans M, Lee T, Hornes M, Frijters A, Pot J et al. (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 16:8186