

PHYLOGENETIC ANALYSES WITHIN THREE SECTIONS OF THE GENUS *Vicia*

TOMASZ SAKOWICZ¹ and TOMASZ CIEŚLIKOWSKI²

¹University of Łódź, Department of Cytogenetics and Plant Molecular Biology,
90-237 Łódź, Banacha 12/16, Poland, ²Centre for Medical Biology Polish
Academy of Sciences, 93-232 Łódź, Lodowa 106, Poland

Abstract: The averaged genomic similarities based on multilocus randomly amplified polymorphic DNA (RAPD) were calculated for eight species representing three sections of the genus *Vicia*: *faba*, *bithynica* and *narbonensis*. The frequency of appearance of the sequences corresponding to 25 decamers selected at random from genomes of different *Fabace* species was checked, and a high correlation with the frequency observed for *Vicia* allowed us to assume their similar weight in typing *Vicia* species. The RAPD-based similarity coefficients compared with those related to whole genome hybridization with barley rDNA and those based on restriction fragment length polymorphism (RFLP) revealed similar interspecies relationships. The averaged RAPD-based similarity coefficient (Pearson's) was 0.68 for all the species, and was section-specific: 0.43 (*bithynica*), 0.50 (*faba*) and 0.73 (*narbonensis*). The averaged similarity coefficient for *V. serratifolia* (0.63) placed it apart from the rest (0.75) of its section. The results correspond to the interspecies relationships built upon non-genetic data. The averaged similarity coefficient for particular RAPD was related to the presence and type of tandemly repeated motif in a primer: 0.7-0.8 for heterodimers (GC, AG, CA, GT, CT), 0.5-0.6 for homodimers (CC, GG) and 0.6 for no repeat, indicating the sensitivity of diversity range to the type of target sequences.

* Author for correspondence: e-mail: tomeksakowicz@wp.pl

Abbreviations used: AFLP – amplified fragment length polymorphism; BLAST – basic local alignment search tool; EDTA – ethylene diamine tetraacetic acid; HTGS – high throughput genomic sequences; ICARDA – International Center for Agricultural Research in the Dry Areas; ITS – internal transcribed spacers; Mb – mega base; Nr – ‘non redundant’; QTL – quantitative trait loci; RAPD – randomly amplified polymorphic DNA; RFLP – restriction fragment length polymorphism; UPGMA – unweighted pairs group method using mathematical averages

Key words: *Vicia*, Average similarity, Multilocus diversity, RAPD

INTRODUCTION

The genus *Vicia* belongs to the Legumes, one of the largest families of flowering plants, representing tremendous morphological, ecological and genetic diversity. It includes approximately 170 species [1]. About 40 species, mostly of Mediterranean origin, are cultivated worldwide [2], while about 10 cultured species (e.g. *V. sativa*, *V. villosa*, *V. faba*, *V. cracca*) are a significant source of fodders, green manures and forages. The extensive research into legume genomes carried out over the past decade not only yielded considerable sequence data, rich enough to perform whole genome comparative analyses, but it enabled the establishment of the methods suitable for this type of investigation [3-7]. These include detection of sequence linkage equilibria, RFLP, RAPD or microsatellite-based mapping, and identification of associations with morphological traits or isozyme patterns. To assess the general taxon internal diversity, one has to include overall sequence homologies and the conservation of their order (synteny). Colinear RFLP maps and a close phylogenetic relationship are good indicators of local colinearity, but detailed analysis of the targeted local region is necessary to reveal microcolinearity, which in plants is regarded as an especially useful tool for plant gene identification and study, and for the comparison of many plant genomes [6]. Microcolinearity is observed even between distantly related species; therefore, *loci* marker probes can be obtained allowing for comparative research on different individuals [8]. Elsewhere, instead of one-to-one colinearity, the distributed network of synteny was discovered [9-11]. Besides the number of comparative experimental data types, statistical methods of sequence analysis were proposed for searches of regions of homology [12, 13].

Several *Vicia* species characteristics (biochemical, molecular, karyotype, genetic etc.) have been described, leading to a general picture of the diversity within this genus [14-20]. Comparative analysis of the above-mentioned data sets revealed some discrepancies in the resulting species arrangements within the *Vicia* genus, and thus a need for a common strategy for data ordering stimulated the search for genomic markers related to the general taxon specificity.

The equivalence between taxonomic interspecies order and species-specific trait relationships and the identification and location of important genes and sequences has been investigated in particular detail in *V. faba* [21, 22]. Its large genome, one of the largest amongst the legumes at 13,000-26,000 Mb [23], seriously complicates the development of saturated linkage maps [24].

The conservation of numerous chromosomal regions in *Vicia* based on isozyme *loci* has been suggested [25]. The information concerning gene arrangement on the genome is recognized as crucial for genome structure evolution. Different kinds of DNA standard markers are considered especially useful for cross-referencing genetic information to reveal areas of homology and to establish

genomic rearrangements [3, 26]. There is particular interest in the RAPD markers tightly linked to genes or quantitative trait *loci* (QTLs) of agronomic significance (rust resistance, broomrape resistance, *Ascochyta fabae* resistance) [22, 27, 28]. However, ribotyping [29], RFLP analysis [17, 30] amplified fragment length polymorphism (AFLP) [31], microsatellites [21, 32], dispersed [33] and tandem repeats [34], and transposon distribution [34-37] are still of particular importance. Such a variety of genetic typing results encourages investigation of the extent of equivalence between the diversity related to the whole *Vicia* genome and the diversity related to non-genetic features, in particular the extent of equivalence between the whole genome-related similarity of the species and their similarity based on a limited genomic representation referring to a randomly selected set of *loci*.

In this study, we investigated *Vicia faba* and a group of its close relatives (eight species: *V. bithynica*, *V. faba*, *V. galilaea*, *V. hyaeniscyamus*, *V. johannis*, *V. kalakhensis*, *V. narbonensis* and *V. serratifolia*, representing the narbonensis, bithynica and faba sections) [38] with the aim of estimating the genetic diversity among the selected *Vicia* species, based on the averaged genomic relationships. We also examined the effectiveness of averaged simple genetic characteristics in the description of groups of species.

MATERIALS AND METHODS

Plant material and genomic DNA isolation

Plant material was obtained from the collections of the International Center for Agricultural Research in the Dry Areas (ICARDA) in Aleppo, Syria (seven species) and the Plant Breeding Station in Olsztyn, Poland (*V. faba*). All the accessions are listed in Tab. 1.

Tab. 1. *Vicia* cultivars and accessions used in this study.

Crop and IG number	Section	Species	Subtaxa	Origin
110 60059	narbonensis	<i>V. narbonensis</i>	var. <i>aegyptica</i>	Greece, Crete
932 60881	narbonensis	<i>V. bithynica</i>	-	Italy
4199 64148	bithynica	<i>V. bithynica</i>	-	Syria, Lattakia
2130 62079	narbonensis	<i>V. serratifolia</i>	-	Turkey
3119 63068	narbonensis	<i>V. johannis</i>	var. <i>procumbens</i>	Syria, Lattakia
4159 64108	narbonensis	<i>V. hyaeniscyamus</i>	-	Syria, Homs
4182 64131	narbonensis	<i>V. kalakhensis</i>	-	Syria, Homs
6031 32210	narbonensis	<i>V. galilaea</i>	-	Italy
8121 42116	narbonensis	<i>V. eristalioides</i>	-	Italy
32 717	faba	<i>V. faba</i>	subsp. <i>Minor</i>	Poland

The total genomic DNA was extracted from an approximately 5-mm long section of the apical root (meristem). The frozen meristem material (1 mg) was disrupted by grinding in liquid nitrogen prior to cell lysis, and digested by

RNase-A in a lysis buffer at 65°C. DNA extraction (DNeasy Plant Mini Kit, Qiagen) was followed by spectrophotometrical quantification of the DNA amount at 260 nm (Hitachi 2000). Subsequently, the DNA solutions originating from separate accessions were collected together to form the final mixtures, each consisting of 7-10 individual genomic DNAs.

Ribotyping

Ribotyping was performed according to the procedure described above (with *Bam*HI digestion of the total DNA sample series). The same hybridization conditions were applied. The plasmid pTA71, including the barley 5.8S, 18S and 25S rDNA genes (without spacers), was used as a hybridization probe [39]. The barley genes are evolutionarily far from the equivalent sequences of particular *Vicia* species. Therefore, in numerical calculation, they might safely be assumed as approximately equidistant from each of them.

Southern hybridization analysis

Two endonucleases, *Bam*HI and *Bgl*III, were applied for the two separate DNA sample series digestions. Digested genomic DNA samples (8 µg per lane) were fractionated by electrophoresis in 1% (w/v) TAE-agarose gels (40 mM Tris, 1 mM EDTA, pH 8.0, 1.14 ml acetic acid/L), before alkaline transfer to nylon membranes (Hybond N+, Amersham). Then they were denatured, hybridized and detected under standard conditions [40]. The filters were immobilized to ³²P-labelled total DNA from *Vicia* species as probes (labelled by nick translation). Stringent washes were carried out at 65°C in 0.5 M sodium phosphate buffer, pH 7.5.

RAPD amplification

The amplification of genomic DNA was performed according to the procedure of Williams [41] with a minor modification related to the temperature profile. The reaction mixture included 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 2 mM MgCl₂ (Finnzyme, Finland), and 0.2 mM dNTP (Promega, USA). The aliquot of 25 µL contained the reaction mixture, 30 ng of primer (SIGMA-ARK, Germany), 5 ng of genomic DNA, and 2 units of thermostable polymerase Dynazyme™ (Finnzyme, Finland) and Milli-Q water (Millipore, Austria). The temperature profile included initial denaturation at 95°C for 1 min., followed by 35 cycles (denaturation at 94°C for 1 min., annealing for 1 min. and extension at 72°C for 2 min.) with a final extension step at 72°C for 10 min. The temperature of annealing was primer-specific and was each time 5°C below its melting temperature (supplied by the manufacturer). The amplification was performed in thin-wall vials (MJ Research, USA) in a Uno II thermocycler (Biometra, Germany). The amplification products were separated against a molecular-weight marker (100-bp DNA ladder, MBI Fermentas) in 1.5% agarose gel and TAE buffer (40 mM TRIS-acetate, 1 mM EDTA, pH 7.8) at 80 V in an MGU 602T electrophoresis unit (CBS Scientific, USA). The gel stained with ethidium

bromide was visualized under ultra-violet light and documented using a computer image system (Vilber Lourmat, France).

BLAST searches

A homology search for 25 decamer oligonucleotides that were applied as primers for RAPD was carried out to estimate the frequency of their appearance in the *Vicia* genome. The search was performed with the aid of the Basic Local Alignment Search Tool (BLAST) program, version 2.2.8 (NCBI: <http://www.ncbi.nlm.nih.gov/BLAST>). The search concerned the following databases: Nr ('Non redundant', i.e. all GenBank+EMBL+DDBJ+PDB sequences), 'Chromosome' and HTGS (Unfinished High Throughput Genomic Sequences; Sequences: phases 1 and 2, but not those of phase 0). The search limits were the same for all the search series ($E = 5000$). The common search was performed across the database, with a common wide taxonomy limit (e.g. *Viridiplanteae*, *Magnoliophyta*, *Fabaceae*, etc.) to ensure the same chance of selection for each species [42], and then the resulting hit numbers were split between species of interest. The species selected for comparison were *Vicia faba* (*Viciae*), *Pisum sativum* (*Viciae*), *Phaseolus vulgaris* (*Phaseoleae*) and *Glycine max* (*Phaseoleae*). An additional search concerning *Arabidopsis thaliana* (*Brassicaceae*) was performed across two databases, 'Non redundant' (Nr) and 'Chromosome'.

The averaged hit numbers (i.e. the frequency of the appearance of particular primer sequences in the genome) were designed for the two species (*G. max* and *P. sativum*) represented in the database by a relatively large number of sequences. Then, the interspecies differences in the hit numbers related to particular decamers were applied as the first approximation of the error related to average values. The averaged hit numbers accompanied by the appropriate errors were subsequently used to extrapolate the hit number value related to the *Vicia* genus. The correlation between hit numbers found for individual species was calculated using Statistica 5.1 software.

Computer-assisted electrophoretogram analyses

The numerical analysis of electrophoretic patterns was performed with the aid of the GelCompar programme (Applied Maths, Kortrijk, Belgium) applying product-moment correlation (Pearson's) and the unweighted pair group method using arithmetic averages (UPGMA) [43]. For the inter-track comparison, the similarity coefficient of Nei was applied. Only amplicons greater than 200 bp were analysed. The values of the similarity coefficient matrices (GelCompare software, 'Fine alignment') were transformed to Fisher z for further numerical analysis [43]. The dendrogram of generalized interspecies relationships was constructed by means of the UPGMA algorithm based on averaged Pearson's correlation coefficients (Statistica 5.1, Statsoft).

RESULTS

Genomic DNA mixtures of eight *Vicia* species were analyzed in a series of RAPD tests. Each mixture consisted of DNA isolates originating from several separate accessions. The amplification products were separated electrophoretically, then the electrophoretic pictures were subjected to computer assisted analysis as described in the Materials and Methods section. Finally, the resulting similarity matrices were applied for the common clustering analysis and for the calculation of the averaged interspecies diversity.

Accuracy of the RAPD test

The accuracy and repetitiveness of typing were verified for each species of interest. For each species, a set of three different DNA mixtures was used in tripled accuracy testing procedures including the isolation of meristemic DNA and amplification. R03, R07 and R08 were the three different RAPD markers applied for each mixture. The general accuracy was higher than 89%. The exceptions were *V. faba* for R03 and *V. narbonensis* for R08, for which the differences were much greater. These differences in the amplicon pools generated for particular species were limited to one or two bands. The greatest diversity, observed for *V. faba*, was additionally tested on the next six distinct genomic mixtures. As no further differences were found (data not shown), the observed discrepancies were interpreted as resulting from genome structure differences, and the experimental error was approximated as not exceeding the value related to molecular weight marker electrophoretic resolution.

The equivalence of primer-based representation of related genomes

The set of decamer primers selected at random was tested to check if it represents particular plant genomes with similar significance, prior to detailed interspecies diversity measurement. The frequency of their appearance in the plant genome of species closely related to *Vicia faba* (*Pisum sativum*, *Glycine max* and *Phaesolus vulgaris*) was compared [9, 11, 44-46].

The decamer distributions are shown in Fig. 1, which presents the comparison of the hit numbers found in the organellar, nuclear and transposon sequences. The latter class of sequences was distinguished because of the significance of the transposable element amplification for plant genome size variation [8, 33, 47] and because of the copy number of transposable elements in the *Vicia* genome, which reaches up to 10^6 [35]. The histogram of similarities found for most of the applied decamer sequences for the three species (*P. sativum*, *G. max*, *P. vulgaris*) and the genus *Vicia* illustrated the general correspondence existing between their genomes.

The correlation between *Vicia* and *P. vulgaris* was high ($r^2 = 0.680$, $p < 0.0001$), similar to that for *P. sativum* ($r^2 = 0.833$, $p < 0.0001$) and for *G. max* ($r^2 = 0.755$, $p < 0.0001$). A high correlation was also found between *Vicia* and *A. thaliana* ($r^2 = 0.654$, $p < 0.0004$, and $r^2 = 0.768$, $p < 0.0001$ for the search across the 'Chromosome' database). As the homologues to the RAPD markers resulting

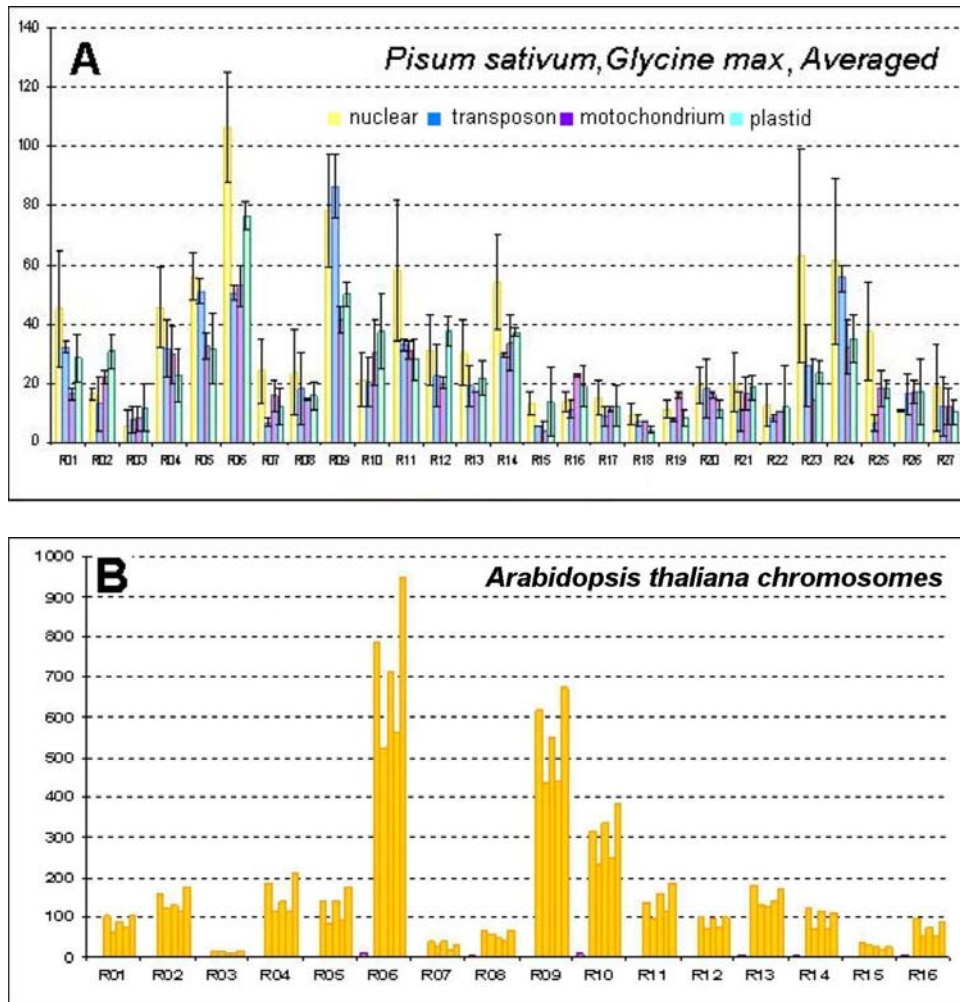


Fig. 1. The averaged frequency of appearance of decamer sequences applied as primers for RAPD, calculated for: A - two close relatives of *V. faba* (*P. sativum*, *G. max*); B - *A. thaliana*. The search was performed across the 'Chromosome' database, related to primers from R01 up to R16 and 'split' to separate chromosomes. The search was carried out using BLASTN 2.2.8 across the NCBI sequence database Nr (the number of descriptions = 1000; number of alignments = 1000; E = 5000). The calculation error was designed as the difference between individual hit numbers.

from the databank searches are mostly genic, the RAPD results might represent the relationships referring to the coding regions of the genome. That is also suggested by the high correlation with *A. thaliana* sequences [5, 48].

To estimate the hit frequency for *Vicia* (the sequences of which are relatively infrequent in the databases), the averaged number of hits was calculated for the

relatively close species *G. max* and *P. sativum*, sequences of which are substantially present in the databases. The correlation between the hit numbers for these species was very high ($r^2 = 0.919$, $p < 0.0001$), between the averaged (over *G. max* and *P. sativum*) hit frequencies and frequencies for *Vicia*, 0.807 ($p < 0.000001$). The error of the predicted hit number distribution in *Vicia* was estimated by the differences in the corresponding hit frequencies in *G. max* and *P. sativum* (Fig. 1).

The described correspondence in hit frequencies observed for different species suggests that despite the relatively limited data for particular *Vicia* species (Fig. 1), similar rules of genomic sequence distribution should be valid for the analysed genus, and that in the genomes more closely related to *Vicia*, there exists a similarly significant degree of correlation. That corresponds with the known simultaneous multi-gene inter-genome homology [25] and suggests that a random set of decamer sequences should possess similar significance for the comparisons of such genomes.

RAPD-based genetic diversity

For six oligonucleotides (R01, R03, R08, R15, R18 and R19), the averaged amplicon numbers were higher than six per track, while for four (R20, R21, R23 and R27), they were lower than three (Tab. 2). The amplicon length generally did not exceed 2 kb, and mostly ranged from 200 bp to 700 bp. The widest range of band length (up to 450 bp) was observed for R01, R03, R07, R18 and R15, while the narrowest (up to 300 bp) was for R06, R11, R16, R20, R22 and R27. For only a few decamers was an amplicon common to all the species observed. For 84.5% (21/25 tests) of the decamers, products common to the entire narbonensis section were generated, while for 16.5% (4/25 tests) of the decamers, no common band was observed in any group bigger than two species. The dendrograms formed as a result of these analyses were strongly diverse in their shape and global homology level. About 20% (5 tests) of them were stair-shaped, while about 80% (20 tests) formed more or less distinct clusters. In about 20% of the results, the greatest distance in the interspecies homology level exceeded 80%.

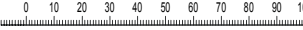
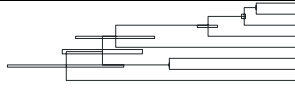
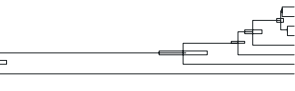
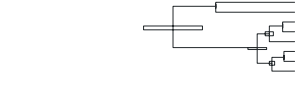
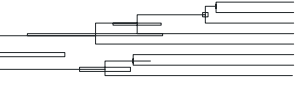
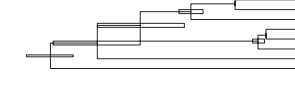
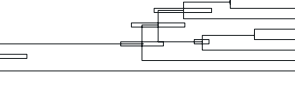
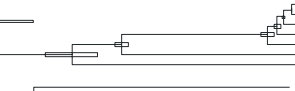
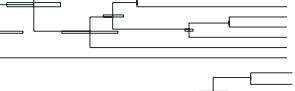

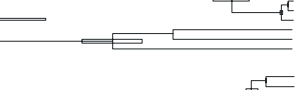
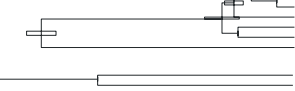
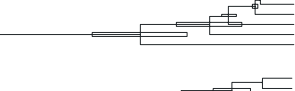

For most of the RAPD tests (Tab. 2), all six species of the narbonensis section were rather similar and their electrophoretic patterns formed common clusters, while the species representing the faba and bithynica sections remained separated. The internal homology level of narbonensis ranged between 95% and 30% (in stair-shaped dendrograms). *V. bithynica* and *V. faba* were gathered out of the narbonensis section, or the difference related to them was lower than the interspecies differences inside this section. These two species were occasionally included into the narbonensis section. They were sometimes clustered together.

Three types of gathering patterns distinguishing separate sections were observed (Tab. 2):

- i. In 14 experiments, the species of the narbonensis section were gathered together, while *V. faba* and *V. bithynica* clustered together out of this section, or were weakly associated with it (R05-R08, R11, R15, R16, R18-R23, R25).

Tab. 2. A comparison of dendrogram topologies. The dendrogram errors are indicated as narrow rectangles. Some repeated motifs in the primer sequences are underlined. The species names are abbreviated as follows: *V. bithynica* – B, *V. faba* – F, *V. narbonensis* – N, *V. haenicum* – H, *V. galilaea* – G, *V. kalakhensis* – K, *V. johannis* – J, *V. serratifolia* – S.

RAPD	Sequence	Number of polymorphic bands	Dendrogram structure	Species
R01	GTCC <u>ACAC</u> GG	27		G S B H J F K N
R02	GATCCGGAGC	26		N H K F J S G B
R03	GTGCGCGACC	29		K N G H J S B F
R04	<u>GAGAGCCT</u> CG	16		S K B H J N G F
R05	GCTCTCAGCG	27		F S B K J H G N
R06	<u>CTGCTGCT</u> G	24		H N G S J B K F
R07	<u>CGGCGGACTA</u>	25		N J H G K B F S
R08	<u>GCCCTCGGAT</u>	26		K H N G J F S B
R09	<u>GTGTGAGAGA</u>	14		B H N K S J G F
R10	GAT <u>CGATCGA</u>	16		K F G H J S B N
R11	<u>CCAGTGCATG</u>	14		N J K H G F S F
R12	<u>GCCTAGATCG</u>	19		N B H F J G K S

RAPD	Sequence	Number of polymorphic bands	Dendrogram structure	
				Species
R13	<u>CTCTCGGACA</u>	9		K H G J B N S F
R14	<u>GCCAGCTGAC</u>	15		K J G N H S B F
R15	<u>CGACTGCGCC</u>	21		F K B G J H N S
R16	<u>GTCCGATCGC</u>	10		K J H S G B F N
R18	<u>TGACGCGCT</u>	26		K S G H J F N B
R19	<u>ACGCGCTAGC</u>	10		J K G H N F S B
R20	<u>CCCTAGCGCG</u>	12		B K N G H F J S
R21	<u>TGACCGCCGG</u>	10		F G K N J H B S
R22	<u>ACCGTCGGCA</u>	12		K J G H N F S B
R23	<u>CTGCAGGCAG</u>	21		H K G J N F S B
R24	<u>GCTTGACCCG</u>	26		J N H F G K S B
R25	<u>CTACCGTGGC</u>	18		F B J N G H S
R27	<u>CAGCGCTGGT</u>	13		J N H F B K S G

- ii. In 6 experiments, only one of these two species was separated from the narbonensis section; the separation of *V. faba* and *V. bithynica* was distinct in three tests each for R01, R03 and R14 and R02, R09 and R24, respectively.
- iii. In 9 experiments, *V. serratifolia* was also separated from the narbonensis section, usually either with *V. bithynica* or *V. faba* (R02, R07, R08, R10, R12, R20-R23).

Interestingly, in 3 RAPD tests (R01, R10 and R16), *V. narbonensis* was found outside its own section. In one test (R16), *V. bithynica* and *V. faba* were also more similar to *V. narbonensis* and clustered altogether out of the narbonensis section. In one test (R12), *V. serratifolia* was significantly excluded from the narbonensis section, while in eight subsequent tests, it was clustered together with *V. bithynica* and *V. faba*. A high homology level (>60%) between the species (*V. narbonensis*, *V. johannis*, *V. hyaeniscyamus*, *V. kalakhensis*, *V. galilaea* and *V. serratifolia*) was observed for 14 out of the 25 RAPD tests, while it was low (<50%) for 9.

Ribotyping

Hybridization with a plasmid pTA71 including barley rDNA as a probe showed interspecies relationships very similar to those obtained from RAPD analysis. The experiments revealed the strong internal consistency of the narbonensis section (more than 95% homology), the separate nature of the faba section (25% homology) and the completely separate features of the *V. bithynica* pattern (less than 10% homology with the rest of the species). Also, an uncertain position for *V. serratifolia* was observed (Fig. 2A).

RFLP-based interspecies diversity

Out of the pool of eight cross-hybridization experiments among all the *Vicia* genomic DNAs, two were arbitrarily selected for numerical validation of the RAPD-based results. The total genomic DNA of *V. johannis* was used as a probe for the total genomic *Vicia* DNAs digested with *Bam*H. In the second experiment, the total genomic DNA from *V. narbonensis* was used as a probe for a Southern hybridization experiment with the total genomic DNA digested with *Bgl*II endonuclease. Both experiments yielded a similar series of bands common to all the species included in the narbonensis section. For *V. johannis* hybridization, there were two common bands, while for *V. narbonensis* hybridization, there were four (Fig. 2B and C). The dendrograms generated for each experiment were similarly stair-shaped; however, in both, the most distant species was *V. bithynica*, then *V. faba*, while the minor differences appeared inside the narbonensis section. While, according to *V. johannis*-based hybridization, the closest relationship linked *V. galilaea* and *V. kalakhensis*, in *V. narbonensis*-based hybridization, *V. galilaea* was the nearest neighbour of *V. johannis*. In both tests, *V. narbonensis* was placed between the narbonensis section and *V. serratifolia* (data not shown).

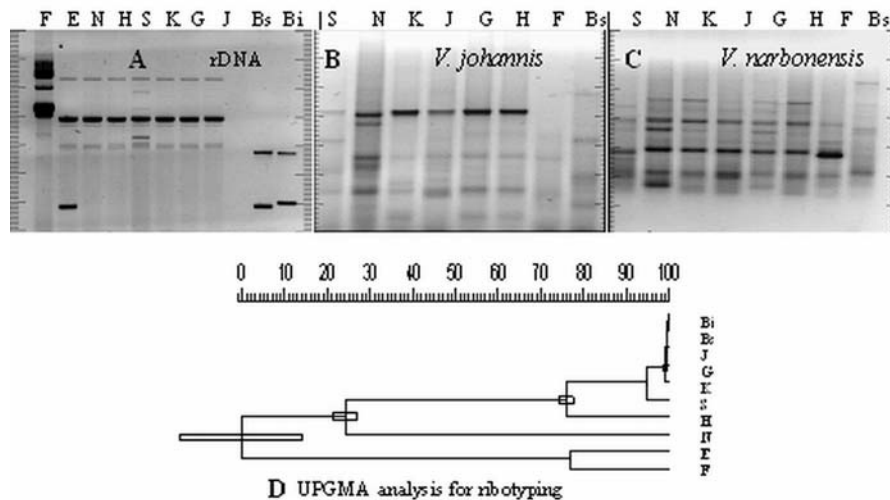


Fig. 2. The results of ribotyping and two whole-genome Southern hybridization analyses. The species names are shown at the top of each gel, with the symbols as follows: B - *V. bithynica*, E - *V. eristaloides*, F - *V. faba*, N - *V. narbonensis*, H - *V. hyanicyamus*, G - *V. galilaea*, K - *V. kalakhensis*, J - *V. johannis*, S - *V. serratifolia*, B(i) - *V. bithynica* (Italy), B(s) - *V. bithynica* (Syria). A - Ribotyping. Hybridization with a plasmid pTA71 including barley rDNA as a probe. B - RFLP analysis. Southern hybridization of the total genomic DNA from *V. johannis* with the total genomic *Vicia* DNA digested with *Bam*HI. C - RFLP analysis. Southern hybridization of the total genomic DNA from *V. narbonensis* with the total genomic DNA digested with *Bgl*III. The scale on the two edges of each picture was used for the normalization procedure. D - The dendrogram correlations (UPGMA) among the species. The scales shown above the dendrogram represent the level of homology between the investigated probes. The values of the clustering errors are shown as narrow rectangles positioned at appropriate dendrogram nodes. *V. eristaloides* is not included in the cluster analysis, as it was not included in the other experiments.

Primer-specific interspecies characteristics

Two numerical values were calculated for each similarity matrix to characterize and compare the interspecies relationships based on particular RAPD typing: the average of the Pearson's correlation coefficient values (C_{sim}^k) and the variance related to this average. All the RAPD tests were divided into five groups according to the repeated motif present in the primer. Then, the C_{sim}^k values were placed on the diagram as a function of the approximated genomic primer frequency, v_k (Tab. 3). The mean of these C_{sim}^k values was similar in subsequent v_k intervals, suggesting the independence of interspecies relationships from primer genomic frequencies ($p < 0.01$). However, for particular types of primers, such adependence appeared to be significantly differentiated (Tab. 3). The differences might reflect the rule that different sequences evolve at different

rates, with the genes more highly conserved than repetitive sequences [8, 48]. When tandem GC, AG, GT, CA or CT repeats characterized the primer, a higher value of the similarity coefficient (0.7-0.8) and a rather wide range of variance (0.1-0.6) were observed.

Tab. 3. Global diversity for the analyzed group of *Vicia* species representing the faba, bithynica and narbonensis sections. The species are described with the same abbreviations as before (Fig. 2).

	<i>S</i>	<i>N</i>	<i>K</i>	<i>J</i>	<i>H</i>	<i>G</i>	<i>F</i>	<i>B</i>	Average interspecies similarity/diversity
$C_{sim\ ij}$ – RAPD similarity coefficient									
<i>S</i>	1.00								
<i>N</i>	0.64	1.00							
<i>K</i>	0.68	0.80	1.00						N = 25
<i>J</i>	0.67	0.85	0.83	1.00					
<i>H</i>	0.73	0.89	0.84	0.91	1.00				$\langle C_{sim} \rangle = 0.68$
<i>G</i>	0.68	0.80	0.86	0.83	0.85	1.00			0.80 ⁿ
<i>F</i>	0.47	0.49	0.49	0.48	0.53	0.54	1.00		0.50 ^f
<i>B</i>	0.46	0.38	0.37	0.40	0.44	0.48	0.50	1.00	0.43 ^b
Δ_{ij} – dispersion of RAPD similarity coefficient									
<i>S</i>	0.00								
<i>N</i>	0.11	0.00							
<i>K</i>	0.10	0.13	0.00						N = 25
<i>J</i>	0.10	0.11	0.12	0.00					
<i>H</i>	0.11	0.13	0.10	0.10	0.00				$\langle \Delta \rangle = 0.09$
<i>G</i>	0.09	0.10	0.12	0.11	0.11	0.00			
<i>F</i>	0.08	0.08	0.06	0.08	0.08	0.08	0.00		
<i>B</i>	0.08	0.07	0.06	0.08	0.09	0.07	0.07	0.00	
$C_{sim\ ij}$ – RFLP similarity coefficient									
<i>S</i>	1.00								
<i>N</i>	0.64	1.00							N = 2
<i>K</i>	0.71	0.83	1.00						$\langle \Delta \rangle = 0.71$
<i>J</i>	0.94	0.82	0.90	1.00					
<i>H</i>	0.66	0.84	0.94	0.87	1.00				$\langle C_{sim} \rangle = 0.56$
<i>G</i>	0.68	0.75	0.99	0.91	0.99	1.00			0.89 ⁿ
<i>F</i>	0.47	0.46	0.62	0.53	0.57	0.49	1.00		0.49 ^f
<i>B</i>	0.60	0.15	-0.94	-0.14	0.75	0.12	0.23	1.00	0.05 ^b
RFLP similarity coefficient (ribotyping)									
									N = 1
									$\langle C_{sim} \rangle = 64.1$

^a $C_{sim\ ij}$ is an averaged value of Pearson's correlation coefficients calculated over 25 RAPDs. ^b Δ_{ij} is a standard error related to $C_{sim\ ij}$. ^c $\langle C_{sim} \rangle$ and $\langle \Delta \rangle$ are the averaged values over the independent values of the presented matrices of $C_{sim\ ij}$ and Δ_{ij} , respectively. ^dthe superscripts n, f, and b refer to the sections narbonensis, faba and bithynica. ^eN refers to the number of the experimental series.

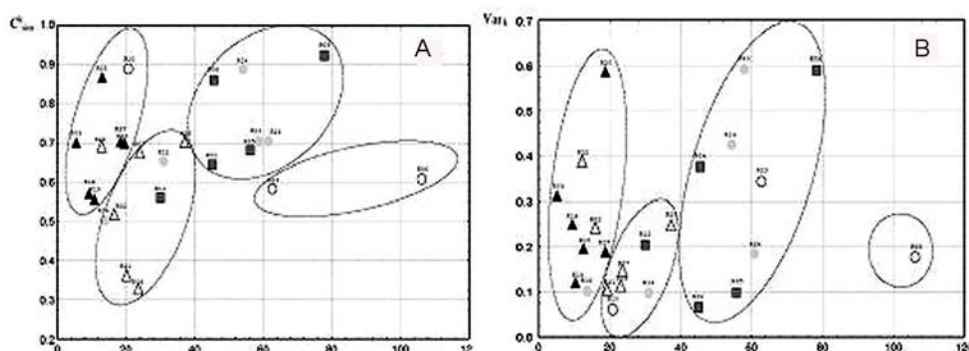


Fig 3. Interspecific genetic similarity/diversity for *Vicia* species. A - Summarized interspecies genetic similarity calculated in particular RAPD analyses for 8 *Vicia* species. The ranges of similarity/diversity common to particular C_{sim}^k types are indicated with the dotted ellipsoid curves. B - Summarized interspecies genetic diversity in particular RAPD analyses for 8 *Vicia* species. C_{sim}^k is an averaged value of all coefficients for the similarity matrix found in one test. The v_k describes the frequency of appearance of a decamer sequence in the *Vicia* genome. Five separate types were distinguished, depending on the presence of a specific motif: (○) decamers, which include the tandemly repeated tetramers or trimers GCAG, GATC or CTG (R06, R10, R23); (△) decamers, with CC or GG repeats (R02, R07, R08, R21, R22, R25); (▲) decamers, with GC repeats (R03, R15, R18, R19, R20, R27); (■) decamers, including other types of tandemly repeated dimer, i.e. CA, CT, GT or AG (R01, R04, R05, R09, R13); and (●) decamers without any tandem repeat (R11, R12, R14, R16, R24).

When a primer included CC or GG repeated motifs, a lower C_{sim}^k value (0.5-0.6) was accompanied by a rather moderate value (0.1-0.4) of the variance. The primer sequences containing trimer or tetramer tandem repeats had rather high (0.9) or moderate (0.6) C_{sim}^k values and a lower variance (0.1-0.4), while primers with no repeats gave a higher (0.6-0.9) C_{sim}^k and wider (0.1-0.6) range of variance (Tab. 3). Generally, the observed interspecies similarity was highest for the *loci* related to a marker containing tandemly repeated heterodimers.

Averaged genomic similarity between *Vicia* species

The averaged similarity related to a random set of decamers (possessing diversified genomic frequency) also noticeably separated the sections bithynica (0.43) and faba (0.50) from narbonensis (0.73), and the corresponding averaged diversity was also section-specific (bithynica: 0.07, faba: 0.08, and narbonensis: 0.1). For the individual species of the narbonensis section, the similarity coefficients were as follows: *V. serratifolia* 0.63, *V. kalakhensis* 0.73, *V. narbonensis* 0.73, *V. galilea* 0.75, *V. johannis* 0.76 and *V. hayaeniscyamus* 0.78 (Tab. 3, Fig. 4A). The final dendrogram structure revealed similar interspecies relationships when particular correlation matrices were weighted by the relative genomic frequencies of the appropriate primers (Fig. 4A, top right corner).

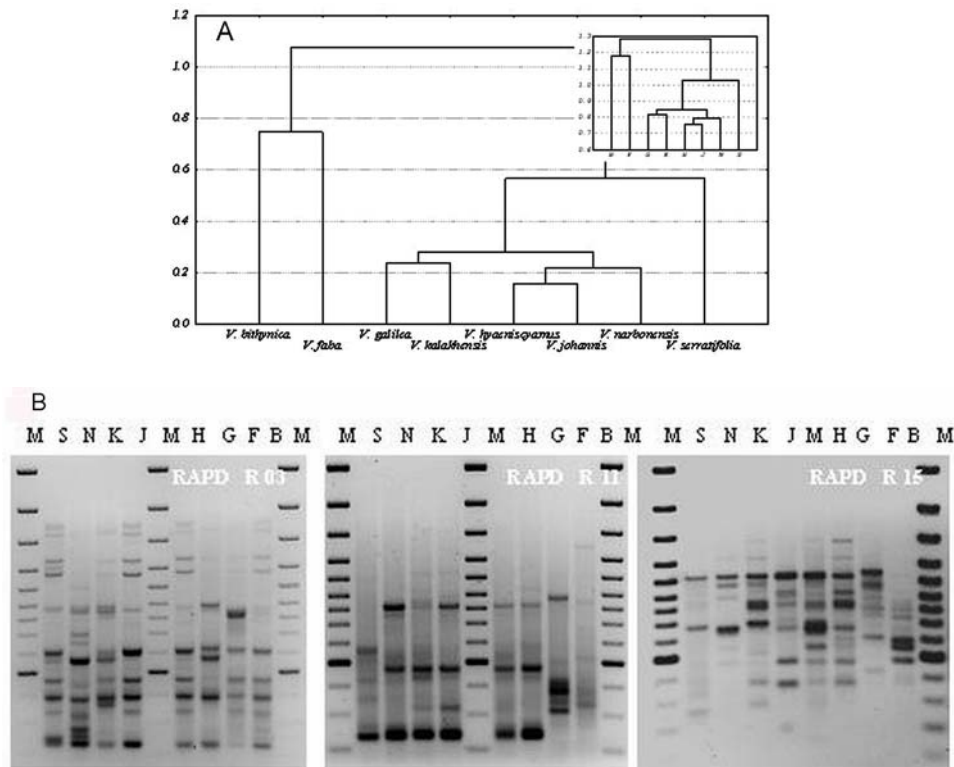


Fig. 4. Taxonomic relationships among the *Vicia* species – the final dendrogram. A - The averaged relationship between eight *Vicia* species found on the basis of 25 RAPD analyses. For calculations, the UPGMA (unweighted pair group method using arithmetic averages) and euclidean distances were applied. In the top right corner, the dendrogram is 'weighted' by primer genomic frequencies. B - Three RAPDs illustrating three points in A. R11 represents low C^k and high interspecies diversity, R16 represents low C^k and low diversity, while R03, located close to the center of the diagram, represents moderate values. The species names are abbreviated as in (Fig. 2), M – molecular weight marker.

DISCUSSION

Microcolinearity, apparent in numerous plant genomes and recognized as especially advantageous for their investigation [8], allows us to establish the general features of the genome structure by means of comparative sequence analyses. It is expressed by a relatively high degree of gene order conservation observed between numerous species closely related to *Vicia*, as in, e.g. [9, 46, 49] but also with *Arabidopsis* [12]. Strong similarities appearing between the DNA fingerprints of *Vicia* species published so far also seem to be the consequence of the above-mentioned colinearity, related to the entire genome. According to some RAPD analyses, to identify particular legume species, more than four different primer sequences are needed [28]. Similarly, as shown in

RFLP experiments [17], to assess a clear relationship between different *Vicia* accessions, the simultaneous analysis of several patterns was necessary. The set of decamers described here was arranged to cover a possibly wide range of different types of genomic *loci*, but simultaneously to possess different GC percentages and different genomic frequencies. It was found that for 25 decamer sequences applied as RAPD primers, despite the observed particular differences, the average value of the interspecies relationships resembled that based on the taxonomically important phenotypic, biochemical or molecular data.

Averaged genomic similarity between *Vicia* species

The search was focused on the generalized interspecies similarity to overcome the differences generated by clustering results related to the particular RAPD, and effective relationships found among the species strongly resembled the established taxonomic relationships (Fig. 3A). The averaged cluster analysis revealed separated clusters of *V. faba* and *V. bithynica*, while all the species forming the narbonensis section were gathered together. This appeared to be in general accordance with our ribotyping-based UPGMA cluster analysis, but also with the RFLP results (Fig. 2A). The average RAPD analysis seemed to confirm the association of *V. serratifolia* with the narbonensis section. However, in all three hybridization tests (RFLP and ribotyping), *V. serratifolia* was found to be much less related to the section. The relationships between the remaining species of the section were different according to each analysis, while the ribotyping resulted in their nearly 100% identity. The RFLPs showed a much weaker association between the species, and in each RFLP, different species were out of the complex (Fig. 2B, 2C). The RAPD-based observation corresponds with the fingerprinting results of other authors, e.g. with those presented by Van de Ven [17]. However, according to him, two other species (*V. galilaea* and *V. hayaeniscyamus*) were found to be especially similar. Some RAPD tests resulted in different relative positions of *V. faba*, *V. bithynica* or *V. serratifolia* to the narbonensis section. These results are consistent with the results of some other authors, indicating that *V. narbonensis* might be found quite apart from the species of its own section. The resultant correspondence of averaged similarity coefficients found among eight *Vicia* species for the two tests (RAPD- and RFLP-based Southern blot hybridization) suggests quite a wide-range genome similarity, and a kind of equivalence between these two kinds of typing, i.e. hexamer- and decamer-based (Tab. 3 and Fig. 3).

Relationships with a karyotype differences

The presented analysis referred to the species whose 1C ranged from 4.58 pg (*V. bithynica*) to 13.35 pg (*V. faba*) [50, 51] with a mean of 7.96 pg. Cluster analysis (Fig. 3A) gathered together four species with similar genome sizes: *V. hayaeniscyamus*, *V. johannis*, *V. narbonensis* and *V. serratifolia*; 7.90 pg, 7.08 pg, 7.28 pg and 7.83 pg, respectively. A greater similarity was also shown between the two species relatively distant according to genome size:

V. kalakhensis and *V. galilaea* (9.18 pg and 6.52 pg). The observed discrepancy between the RAPD clustering of *Vicia* species and their appropriate 1C values might be related to different evolutionary paths of particular species [47, 52].

In *Vicia*, apart from variations in chromosome number, there is a large-scale variation in chromosome size, nuclear DNA content, repetitive and non-repetitive DNA sequences and chromosomal repatterning in non-nuclear and nuclear chromosomes. The results of Raina and Ogihara [29], and literature therein, concerning rDNA repeat variability and length heterogeneity place *V. faba* and *V. bithynica* outside the narbonensis complex. In the presented ribosomal hybridization study, *V. bithynica* turned out to be the most distant from the narbonensis section (less than 10% similarity), while the second, *V. faba*, was about 30% similar, and the third, *V. serratifolia*, about 75%. The internal homology of the remaining five species of the narbonensis section was 95%.

According to the results reported by other authors, the analysis of *rrn*-related internal transcribed spacers (ITS) does not allow clarification of the phylogenetic relationships within the narbonensis complex, suggesting the possible relative conservativeness of the section [18]. This ITS alignment located *V. kalakhensis* further from the other members of the group and slightly distinguished the *V. serratifolia* species. Inside the section, the karyotype analysis gathered together *V. galilaea* and *V. johannis*, separately from both *V. narbonensis* and *V. kalakhensis*. It also widened the individual distance of *V. hyaniscyamus* [18].

Relationships with phenotype diversity

The question to what extent the similarity between the investigated species might correspond to relationships visible at the proteome level is partially answered by several papers analysing the polymorphism of albumins, isoenzymes (dehydrogenases, superoxide dismutase) or aspartate aminotransferase [15, 16]. The analysis involved allele frequencies at eight polymorphic *loci*. According to the cited authors, the variation within species revealed by UPGMA cluster analysis performed on a set of the narbonensis section species along with *V. faba* and *V. bithynica* allowed the recognition of *V. serratifolia* as placed peripherally in the narbonensis section (Nei's genetic identity < 0.7). Surprisingly, *V. narbonensis* was located apart from the rest of the species of the section, but not as distantly as *V. serratifolia* (genetic identity > 0.7). The overall diversity calculation exposed the greatest distances within the narbonensis section for *V. faba* (genetic identity < 0.4) and *V. bithynica* (< 0.5). Furthermore, similar interspecies relationships were revealed by orthozyme data. They separated both *V. faba* and *V. bithynica* from the narbonensis section, while morphological features separated only *V. bithynica* from the section [15]. Hence, the presented interspecies similarity averaged over a random set of genomic *loci* remained in agreement with that related to *loci* of species-specific enzymes and with the established taxonomic relationships (Fig. 4).

CONCLUSIONS

The averaged RAPD-based interspecies relationships are in agreement with RFLP-based and with rDNA-based typing results, confirming that averaged values found on the basis of 25 markers are quite a good approximation of interspecies relationships in *Vicia*. This suggests that interspecies relationships visible in analyses of rDNA emulate those resulting from a series of short random sequences. When included into analysis, the primer genomic frequencies did not change the observed interspecies relationships.

The results validate the micro- and colinearity known for *Fabaceae* and are congruent with data known from molecular and morphological taxonomy records. The averaged similarity coefficient and its variance, both referring to the particular RAPD test, produce an informative visualization of interspecies relationships. The primer motif influences the result of interspecies relationships, but in most cases keeps/maintains the relationships between the sections.

Genome description averaged over random *loci* proved to be an effective and informative tool for interspecies analysis in *Vicia*.

Acknowledgements. We would like to thank Professor Maria Olszewska from the University of Łódź for her help, encouragement and care. This study was supported by the University of Łódź, grant No 505/040424, and partially by the Centre for Medical Biology PAS.

REFERENCES

1. Allkin, R., Goyder, D.J., Bisby, F.A. and White, R.J. Names and synonymus of species and subspecies in the *Viciae*. Issue 3 *Viciae* Database Project Publication No 7 Southampton, 1986.
2. Harlan, J.R. Theory and dynamics of grassland agriculture. (Van Nostrand, D. Ed), Princeton, New Jersey, USA, 1956.
3. Buntier, J.B., Sørensen, A.P. and Peleman, J.D. Haplotype diversity: the link between statistical and biological association. **Trends Plant Sci.** 10 (2005) 1360-1385.
4. Gepts, P., Beavis, W.D., Brummer, E.Ch., Shoemaker, R.C., Stalker, H.T., Weeden, N.F. and Young, N.D. Legumes as a model plant family. Genomics for food and feed report of the cross-legume advances through genomics conference. **Plant Physiol.** 137 (2005) 1228-1235.
5. Lee, J.M., Grant, D., Vallejos, C.E. and Shoemaker, R.C. Genome organization in dicots. II. Arabidopsis as a 'bridging species' to resolve genome evolution events among legumes. **Theor. Appl. Genet.** 103 (2001) 765-773.
6. Paterson, A.H., Bowers, J.E., Burow, M.D., Draye, X., Elsik, Ch.G., Jiang, Ch-X., Katsar, C.S., Lan, T.-H., Lin, Y.-R. and Wright, R.J. Comparative genomics of plant chromosomes. **Plant Cell** 12 (2000) 1021-1029.

7. Zhu, H., Choi, H.-K., Cook, D.R. and Shoemaker, R.C. Bridging model and crop legumes through comparative genomics. **Plant Physiol.** 137 (2005) 1189-1196.
8. Bennetzen, J.L. Comparative sequence analysis of plant nuclear genomes: microcolinearity and its many exceptions. **Plant Cell** 12 (2000) 1021-1029.
9. Choi, H.K., Mun, J.H., Kim, D.J., Zhu, H., Baek, J.M., Mudge, J., Roe, B., Ellis, N., Doyle, J., Kiss, G.B., Young, N.D. and Cook, D.R. Estimating genome conservation between crop and model legume species. **Proc. Natl. Acad. Sci. USA** 101 (2004) 15289-15294.
10. McCouch, S.R. Genomics and synteny. **Plant Physiol.** 125 (2001) 152-155.
11. Yan, H.H., Mudge, J., Kim, D.-J., Shoemaker, R.C. and Young, N.D. Comparative physical mapping reveals features of microsynteny between Glycine max, Medicago truncatula, and Arabidopsis thaliana. **Genome** 47 (2004) 141-155.
12. Grant, D., Cregan, P. and Shoemaker, R.C. Genome organization in dicots: genome duplication in Arabidopsis and synteny between soybean and Arabidopsis. **Proc. Natl. Acad. Sci. USA** 97 (2000) 4168-4173.
13. Frediani, M., Caputo, P., Venora, G., Ravelli, C., Ambrosio, M. and Cremonini, R. Nuclear DNA contents, rDNAs, and karyotype evolution in *Vicia* subgenus *Vicia*: II. Section Peregrinae. **Protoplasma** 226 (2005) 181-190.
14. Fuchs, J., Strehl, S., Brandes, A., Schweizer, D. and Schubert, I. Molecular-cytogenetic characterization of the *Vicia faba* genome, heterochromatin differentiation, replication patterns and sequence localization. **Chromosome Res.** 6 (1998) 219-230.
15. Leht, M. and Jaaska, V. Cladistic and phenetic analysis of relationships in *Vicia* subgenus *Vicia* (Fabaceae) by morphology and isozymes. **Plant Syst. Evol.** 232 (2001) 237-260.
16. Przybylska, J. and Zimniak-Przybylska, Z. Electrophoretic seed albumin patterns and species relationships in *Vicia* sect *Faba* (Fabaceae). **Plant. Syst. Evol.** 198 (1995) 179-194.
17. Van de Ven, W.T.G., Duncan, N., Ramsay, G., Philips, M., Powell, M. and Waugh, R. Taxonomic relationships between *V. faba* and relatives based on nuclear and mitochondrial RFLPs and PCR analysis. **Theor. Appl. Genet.** 86 (1993) 71-80.
18. Venora, G., Blangifirti, S., Frediani, M., Maggini, F., Castiglione, M.R. and Cremonini, R. Nuclear DNA contents, rDNAs, chromatin organization, and karyotype evolution in *Vicia* sect *Faba*. **Protoplasma** 213 (2000) 118-125.
19. Wojciechowski, M.F. In Advances in legume systematics. Part 10. Higher level systematics. Edited by Klitgaard, B.B. and Bruneau, A. Royal Botanic Gardens, Kew, U.K. (2003) 5-35.
20. Raina, S.N. Genome organization and evolution in the genus *Vicia*. In: Biological approaches and evolutionary trends in plants. Academic Press, New York, NY (1990) 183-201.

21. Pozarkova, D., Kublizkova, A., Roman, B., Torres, A.M., Lucretti, S., Lysak, M., Dolezel, J. and Macas, J. Development and characterization of microsatellite markers from chromosome 1-specific DNA libraries of *Vicia faba*. **Biol. Plant.** 45 (2002) 337-345.
22. Román, B., Torres, A.M., Rubiales, D., Cubero, J.I. and Satovic, Z. Mapping of quantitative trait loci controlling broomrape (*Orobanche crenata* Forsk) resistance in *faba* bean (*Vicia faba* L). **Genome** 45 (2002) 1057-1063.
23. Bennett, M.D. and Leitch, I.J. Angiosperm DNA C-values database (release5.0 dec.2004) <http://www.rbgekew.org.uk/cval/homepage.html> (2004).
24. Román, B., Satovic, Z., Pozarkova, D., Macas, J., Dolezel, J. and Torres, A.M. Development of a composite map in *Vicia faba*, breeding applications and future prospects. **Theor. Appl. Genet.** 108 (2004) 1079-1088.
25. Torres, A.M., Weeden, N.F. and Martín, A. Linkage among isozyme, RFLP and RAPD markers in *Vicia faba*. **Theor. Appl. Genet.** 85 (1993) 937-945.
26. Heslop-Harrison, J.S. Comparative genome organization in plants: from sequence and markers to chromatin and chromosomes. **Plant Cell** 12 (2000) 617-635.
27. Avila, C.M., Sillero, J.C., Rubiales, D., Moreno, M.T. and Torres, A.M. Identification of RAPD markers linked to the Uvf-1 gene conferring hypersensitive resistance against rust (*Uromyces Viciae-fabae*) in *Vicia faba* L. **Theor. Appl. Genet.** 107 (2003) 353-358.
28. Weder, J.K.P. Identification of Food and Feed Legumes by RAPD-PCR. **Lebensm-Wiss u-Technol.** 35 (2002) 504-511.
29. Raina, S.N. and Ogihara, A. Ribosomal DNA repeat unit polymorphism in 49 *Vicia* species. **Theor. Appl. Genet.** 90 (1995) 477-486.
30. Potokina, E., Tomooka, N., Vaughan, D.A., Alexandrowa, T. and Xu, R.Q. Phylogeny of *Vicia* subgenus *Vicia* (Fabaceae) based on analysis of RAPDs and RFLP of PCR-amplified chloroplast genes. **Genet. Res. Crop. Evol.** 46 (1999) 149-161.
31. Potokina, E., Blattner, F.R., Alexandrowa, T. and Bachmann, K. AFLP diversity in the common vetch (*Vicia sativa* L) on the world scale. **Theor. Appl. Genet.** 105 (2002) 58-67.
32. Peakall, R., Gilmore, S., Keys, W., Morgante, M. and Rafalski, A. Cross-Species amplification of soybean (*Glycine max*) simple sequence repeats (SSRs) within the genus and other legume genera: implications for the transferability of SSRs in plants. **Mol. Biol. Evol.** 15 (1998) 1275-1287.
33. Frediani, M., Galati, M.T., Maggini, F., Galasso, I., Minelli, S., Ceccarelli, M. and Cionini, P.G. A family of dispersed repeats in the genome of *V. faba*: structure, chromosomal organization, redundancy modulation, and evolution. **Chromosoma** 108 (1999) 317-324.
34. Macas, J., Navratilova, A. and Meszaros, T. Sequence subfamilies of satellite repeats related to rDNA intergenic spacer are differentially amplified on *Vicia sativa* chromosomes. **Chromosoma** 112 (2003) 152-158.

35. Pearce, S.R., Harrison, G., Li, D., Heslop-Harrison, J.S., Flavell, A. and Kumar, A. The Ty-1-copia group retrotransposons in *Vicia* species: copy number, sequence heterogeneity and chromosomal localization. **Mol. Gen. Genet.** 250 (1996) 305-315.
36. Naumann, P., Koblizkova, A., Navratilova, A. and Macas, J. Significant expansion of *Vicia pannonica* genome size mediated by amplification of a single type of giant retroelement. **Genetics** 173 (2006) 1047-1056.
37. Macas, J., Neumann, P. and Pozarkova, D. Zaba: a novel miniature transposable element present in genomes of legume plants. **Mol. Gen. Genomics** 269 (2003) 624-631.
38. Maxted, N., Khettab, M.A. and Bisby, F.A. The newly discovered relatives of *Vicia faba* do little to resolve the enigma of its origin. **Bot. Chron.** 10 (1991) 435-465.
39. Gerlach, W.L. and Bedbrook, J.R. Cloning and characterization of ribosomal RNA genes from wheat and barley. **Nucleic Acids Res.** 8 (1979) 1869-1885.
40. Sambrook, J., Fritsch, E.F. and Maniatis, T. **Molecular Cloning A Laboratory Manual**, 2nd Ed. Edited by Cold Spring Harbor Laboratory Press, New York, 1989.
41. Williams, J.G.K., Kubelik, A.R., Livak, A.J., Rafalski, J.A. and Tingey, S.V. DNA polymorphism amplified by arbitrary primers are useful as genetic markers. **Nucleic Acids Res.** 18 (1990) 6531-6535.
42. Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman D.J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. **Nucleic Acids Res.** 25 (1997) 3389-3402.
43. Sokal, R.R. and Sneath, P.H.A., **Principles of Numerical Taxonomy**. WH Freeman & Co, San Francisco, CA USA, 1963.
44. Francki, M.G. and Mullan, D.J. Application of comparative genomics to narrowleafed lupin (*Lupinus angustifolius* L.) using sequence information from soybean and Arabidopsis. **Genome** 47 (2004) 623-632.
45. Hecht, V., Foucher, F., Ferrandiz, C., Macknight, R., Vardy, M.E., Ellis, N., Beltran, J.P., Rameau, C. and Weller, J.L. Conservation of Arabidopsis flowering genes in model legumes. **Plant Physiol.** 137 (2005) 142-1434.
46. Kalo, P., Seres, A., Taylor, S.A., Jakab, J., Kevei, Z., Kereszt, A., Endre, G., Ellis, T.H.N. and Kiss, G.B. Comparative mapping between *Medicago sativa* and *Pisum sativum*. **Mol. Genet. Genomics** 272 (2004) 235-246.
47. Hill, P., Burford, D., Martin, D.M. and Flavell, A.J. Retrotransposon populations of *Vicia* species with varying genome size. **Mol. Genet. Genomics** 273 (2005) 371-381.
48. SanMiguel, P. and Bennetzen, J.L. Evidence that recent increase in maize genome size was caused by the massive amplification of intergene retrotransposons. **Ann. Bot.** 82 Suppl A (1998) 37-44.

49. Menacio-Hautea, D., Fatokum, C.A., Kumar, L., Danesh, D. and Young, N.D. Comparative genome analysis of mungbean (*Vigna radiata* L. Wilczek) and cowpea (*V. unguiculata*) using RFLP analysis. **Theor. Appl. Genet.** 86 (1993) 797-810.
50. Bennett, M.D. and Smith, J.B. Nuclear DNA amounts in angiosperms. **Phil. Trans. of the Royal Society of London B** 334 (1991) 309-345.
51. Olszewska, M.J. and Osiecka, R. The relationship between 2C DNA content, life-cycle type, systematic position and the dynamics of DNA endoreplication in parenchyma nuclei during growth and differentiation of roots in some dicotyledonous-herbaceous species. **Biochemie und Physiologie der Pflanzen** 178 (1983) 581-599.
52. Naranjo, C.A., Ferrari, M.R., Palermo, A.M. and Poggio, L. Karyotype, DNA content and meiotic behaviour in five south american species of *Vicia* (Fabaceae). **Ann. Bot.** 82 (1998) 757-764.