Theor Appl Genet (1992) 83:783–790

THEORETICAL AND APPLIED GENETICS

Genetic diversity at isozyme and RFLP loci in *Brassica campestris* as related to crop type and geographical origin

J. Mitchell McGrath* and Carlos F. Quiros

Department of Vegetable Crops, University of California-Davis, CA 95616, USA

Received July 26, 1991; Accepted September 3, 1991
Communicated by H. F. Linskens

Summary. Twenty accessions of *Brassica campestris*, representing the major crop types and their geographical origin, were tested for gene frequency at five isozyme and four RFLP loci. The majority of alleles (67%) were found in all geographic regions. Nearly 3 times more alleles were detected at RFLP loci than at isozyme loci. Genetic diversity among crop types (with the exception of turnip) was similar to diversity estimates of geographical regions, implying that crops used for similar purposes (i.e., oilseed or leafy vegetable) are derived from geographically differentiated populations. Geographically, Central Asian and Indian types showed the highest level of heterozygosity (excluding self-fertile sarson oilseed types), followed closely by European varieties, and Asian varieties showed the greatest gene diversity. Phenetic dendrograms indicated that sarson and Chinese cabbage have diverged considerably from other types, perhaps due to consequences of their breeding habit or origin.

Key words: Crop evolution – *Brassica rapa* – Rapeseed – Turnip

Introduction

Brassica campestris L. (syn. *B. rapa* Metz.) (2n = 20) contains a number of morphologically variable but inter-fertile subspecies. The classification of subspecies has generally followed patterns of crop use (e.g., seed-extracted oil, leafy vegetable and turnip fodder) and geography (e.g., Europe, India and Asia), and as a such may reflect the history of domestication. Some subspecies currently in cultivation, such as turnip and oilseed, were probably gathered and utilized by late neolithic peoples (Renfrew 1973; Prakash and Hinata 1980; Zohary and Hopf 1988). Many authors concur that the region ranging from temperate Europe to Western Siberia should be included in the primary center of origin (reviewed in Prakash and Hinata 1980), yet the origin of Asian subspecies has been largely enigmatic. Modern Asian leafy vegetables may have been derived from European introductions that were then subject to intensive breeding and selection by early Asian horticulturalists (Burkill 1930; cited in Prakash and Hinata 1980). Alternatively, the natural range of *B. campestris* may have extended across Asia and was subsequently divided into European and Asian races by geologic or climatic changes (e.g., Whyte 1983). However, the latter hypothesis was discounted by De Candolle (1886) due to the lack of wild or weedy *B. campestris* forms in Asia. Ancient writings suggest that a number of *B. campestris* forms were cultivated in China by 5000 B.C., among these turnip (ssp. *rapifera*), pak-choi (ssp. *chinesis*) and Chinese cabbage (ssp. *pekinensis*) (Li 1983).

Hybrids between *B. campestris* subspecies are often fully fertile (Sinskaia 1927; Olsson 1954; McGrath and Quiros 1991 a). However, in some subspecies hybrid combinations, notably those involving the Indian oilseeds sarson and toria, reductions in F_1 hybrid pollen and seed fertility have been reported (Olsson 1954; McGrath and Quiros 1991 a). Reciprocal differences in leaf morphology, most pronounced when turnip and pak-choi are crossed, are also evident in hybrids (McGrath and Quiros 1991 a). Novel characters and sterility often appear in F_2 populations derived from inter-subspecies hybrids (Sinskaia 1972, McGrath and Quiros 1991 a). Analyses of isozyme composition (Denford and Vaughan 1977) and restriction fragment length polymorphisms (RFLP) (Song et al. 1990) suggest that diversity is partitioned geographically into Asian and Indo-European groups. However, it is not yet clear how genetic diversity is partitioned between and within subspecies, and to what extent genetic diversity parallels morphological variation.

* To whom correspondence should be addressed.
Present address: Department of Biology, University of Michigan, Ann Arbor, MI 48109-1048, USA

Table 1. List of *B. campestris* accessions

Symbol	Variety	Subspecies	Area[a]	Accession		
				Use[b]	Number[c]	Source[d]
B. campestris L. (Syn. *B. rapa* Metz.)						
TCH	Torch	*olifera* (Metz.) Sinsk.	E	O	B200	DC
PTG	Purple Top White Globe	*rapifera* (Metz.) Sinsk.	E	T	B456	HM
WFD	White Flat Dutch	*rapifera*	E	T	6106	NSSL
YS	Yorii Spring	*rapifera*	A-N	T	B493	JSS
SHO	Shogoin	*rapifera*	A-N	T	B226	S
SHL	Shelgham	*rapifera*	I	T	103719	NSSL
RAB	Broccoli Raab	*perviridis* Bailey	E	L	–	HM
TND	Tendergreen	*perviridis*	A	L	B232	NK
LC	Lei Choi	*chinensis* (L.) Makino	A-S	L	B489	L
CPC	Canton Pak-choi	*chinensis*	A-S	L	–	S
KHC	Kwan Hoo Choi	*parachinensis* (Bailey) T.&L.	A-S	L	B233	RC
TT	Tah Tsai	*narinosa* Bailey	A-S	L	–	PW
WB	Wong Bok	*pekinensis* (Lour.) Olsson	A-N	L	B113	S
MAT	Matsushima	*pekinensis*	A-N	L	B488	JSS
MIZ	Mizuna	*japonica* Sieb. (syn. *nipposinica*)	A-N	L	B487	JSS
HTT	Hong Tsai Tai	*utilis* T.&L.	A-S	?	–	PW
T2	toria	*dichotoma* (Roxb.) Olsson	I	O	1186	DC
T1	toria	*dichotoma*	I	O	1390	DC
S1	yellow sarson	*trilocularis* (Roxb.) Olsson	I	O	1394	DC
S2	yellow sarson	*trilocularis*	I	O	1404	DC
	yellow sarson	*trilocularis*	I	O	1158	DC
	yellow sarson	*trilocularis*	I	O	1160	DC
	yellow sarson	*trilocularis*	I	O	1164	DC
	yellow sarson	*trilocularis*	I	O	1169	DC
	yellow sarson	*trilocularis*	I	O	1170	DC
	yellow sarson	*trilocularis*	I	O	1176	DC
	yellow sarson	*trilocularis*	I	O	1179	DC
	yellow sarson	*trilocularis*	I	O	1184	DC
	yellow sarson	*trilocularis*	I	O	1185	DC
	brown sarson	*dichotoma*	I	O	1165	DC
	brown sarson	*dichotoma*	I	O	1167	DC
	brown sarson	*dichotoma*	I	O	1180	DC
	brown sarson	*dichotoma*	I	O	1182	DC
B. tournefortii Gouan.			I	–	0850	DC
B. tournefortii			I	–	0575	DC

[a] Area: A-N = Northern Asia; A-S = Southern Asia; E = Europe; I = India
[b] Use: O = oilseed; T = turnip; L = leafy vegetable; ? = uncertain
[c] All D. Cohen material preceded by DC77–; (except B200)
[d] Seed sources: DC = D. Cohen, UC Davis; HM = Harris Moran Seed Co, Rochester N.Y.; INIA = Instituto Nacional de Investigacinones Agrarias, Madrid, Spain; JSS = Johnny's Selected Seeds, Albion, Maine; L = Lockhart Seed Co, Stockton, Calif.; NK = Northrup King Seed Co, Gilroy, Calif.; NSSL = National Seed Storage Lab, Fort Collins, Colo.; PW = Paul Williams, University of Wisconsin; RC = Redwood City Seed Co, Redwood City, Calif.; S = Sakata Seed Co, San Francisco, Calif.
The taxonomic authorities used in defining subspecies and varieties are listed in McGrath and Quiros (1991a) and follow the designations of Prakesh and Hinata (1980)

We report here results of a survey examining genetic variation both within and among *B. campestris* subspecies, and between *B. campestris* and *B. tournefortii*, perhaps the closest $n = 10$ relative of *B. campestris* (Denford and Vaughan 1977).

Materials and methods

A list of accessions and their sources used in this study is given in Table 1 (nomenclature follows that of Prakash and Hinata 1980). Isozyme analysis was performed according to methods previously described by Quiros et al. (1987). Informative zymograms were obtained for PGI (phosphoglucose isomerase), 6PGD (6-phosphoglucose dehydrogenase) and PGM (phosphoglucose mutase). DNA isolation and RFLP analysis, using cDNA clones as probes, was performed as described by McGrath and Quiros (1991b). Allele frequencies were determined for nine loci in 20 *B. campestris* subspecies and in two accessions of *B. tournefortii*. The genetics of these loci have been previously determined in *B. campestris* (Truco 1986; Quiros 1987; McGrath and Quiros 1991b). At least four independently segregating regions of the *B. campestris* genome were marked

Table 2. Allele frequencies for scored nine loci (as percentages). Varieties correspond to Table 1

	Variety																			
	TCH	PTG	WFD	YS	SHO	SHL	RAB	TND	LC	CPC	KHC	TT	WB	MAT	MIZ	HTT	T2	T1	S1	S2
ISOZYMES																				
n=	14	15	13	14	14	15	14	14	14	15	14	15	14	14	14	15	15	14	14	14
6Pgd-1																				
-1	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Pgi-2																				
* -1	100	100	100	36	79	83	93	89	86	100	54	47	39	32	32	93	60	36	.	.
-2	.	.	.	64	21	17	7	11	14	.	46	53	61	68	68	7	40	64	100	100
Pgm-1																				
* -1	79	83	89	43	96	87	50	61	75	80	89	97	.	.	38	70	93	43	100	100
-2	21	17	12	57	4	13	50	39	25	20	11	3	100	100	62	30	7	57	.	.
Pgm-2																				
-1	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Pgm-3																				
* -1	25	17	4	50	.	53	21	43	100	27	32	80	100	100	14	100	53	57	.	.
-2	32	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
-3	.	33	69	11	25	10	69	.	.	.	68	17	.	.	54	.	30	4	100	100
-4	43	50	27	39	75	37	11	57	.	73	.	3	.	.	32	.	17	39	.	.
cDNA																				
N=	15	16	12	16	8	16	13	16	16	16	14	16	12	16	16	16	16	15	6	6
BN8																				
-1	.	.	.	3	.	.	8	6	.	.	.	.	.	.	.	.	6	.	.	.
-2	.	.	.	41	.	19	.	6	19	.	.	.	71	72	6	.	22	13	.	.
-3	.	3	.	16	.	.	.	9	.	.	.	.	.	.	.	.	.	.	.	.
* -4	100	97	100	34	100	81	92	59	81	100	100	100	29	28	94	50	72	87	100	100
-5	.	.	.	6	.	.	.	19	.	.	.	.	.	.	.	50	.	.	.	.
BN33																				
-1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	13	9	.	.	.
-2	4	.	4	.	.	.	.	.	.	.	.	.	.	.	.	.	3	.	.	.
* -3	18	44	38	6	6	50	50	44	19	91	44	97	.	81	69	69	56	10	100	100
-4	8	13	13	.	.	.	.	.	.	.	.	.	.	.	.	.	3	4	.	.
-5	4	6	.	.	.	34	28	19	.	.	19	.	.	.	.	.	3	4	.	.
-6	64	38	46	69	75	9	22	38	.	.	38	.	.	.	19	19	16	80	.	.
-7	.	.	.	25	19	6	.	.	81	9	.	3	100	19	13	.	9	4	.	.
-8	4	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
BN55																				
-1	.	.	.	.	.	6	.	3	16	30	4	6	.	.	.	56	27	35	.	.
* -2	.	6	.	34	11	6	31	38	47	.	39	19	36	.	22	.	.	.	.	100
-3	.	6	.	3	22	3	4	16	19	70	43	.	23	100	9	.	.	.	.	.
-4	.	3	.	3	.	19	.	25	.	.	.	.	.	.	.	.	.	.	.	.
-5	.	6	.	.	.	9	.	.	.	.	4	75	.	.	.	44	7	.	.	.
-6	.	.	.	.	56	.	31	.	.	.	.	.	.	.	.	.	.	.	.	.
-7	.	.	.	56	11	3	.	19	.	.	.	.	.	.	53	.	27	15	100	.
-8	100	34	50	.	.	19	19	.	13	.	.	.	.	.	.	.	10	.	.	.
-9	.	44	50	3	.	34	15	.	6	.	11	.	.	.	16	.	30	50	.	.
-10	.	.	.	.	.	.	.	.	.	.	.	.	41	.	.	.	.	.	.	.
CA12																				
-1	.	3	.	.	.	.	.	.	.	.	.	.	.	.	50	.	.	.	.	.
-2	.	9	9	.	.	.	.	23	3	.	25	3	.	.	.	22	3	39	.	.
-3	5	47	.	28	50	3	.	.	.	.	.	16	33	28	19	.	9	4	.	.
-4	65	.	.	.	.	44	54	17	.	3	.	.	.	.	.	.	13	14	.	.
-5	10	.	41	3	.	13	.	13	59	22	.	3	.	.	.	50	6	4	.	.
-6	5	.	14	53	.	.	42	7	.	6	7	.	.	.	31	.	44	14	.	.
* -7	.	.	14	3	11	25	4	37	38	69	57	78	.	.	.	28	3	11	.	.
-8	.	41	23	13	39	16	.	.	.	.	11	.	67	72	.	.	.	.	.	.
-9	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	4	.	.
-10	10	.	.	.	.	.	.	3	.	.	.	.	.	.	.	.	22	11	100	100
-11	5	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H_o	.231	.319	.278	.430	.265	.390	.350	.431	.271	.197	.333	.190	.220	.172	.370	.293	.418	.381	.000	.000
A^o	2.4	2.7	2.2	3.0	2.1	3.1	2.4	3.0	2.1	1.8	2.3	2.1	1.6	1.4	2.3	1.9	3.4	3.0	1.0	1.0

* *Most frequent allele*

and defined by loci: (1) *Pgi-2*, (2) *CA12*, (3) *Pgm-3* and *BN55* (3 cM apart) and (4) *BN8* and *BN33* (24 cM apart) (McGrath and Quiros 1991 b). Map positions of *6Pgd-1*, *Pgm-1* and *Pgm-2* remain to be determined in *B. campestris*, however, *6Pgd-1* is syntenic with *BN33* in *B. oleracea* (McGrath et al. 1990). Allele frequency data were converted to Nei's genetic distance values with the aid of a computer program written by Dr. Kermit Ritland (University of Toronto, Department of Botany).

Results

Gene frequencies for nine loci (five isozyme and four RFLP) used for genetic diversity estimates are given in Table 2. Two loci, *6Pgd-1* and *Pgm-2*, were monomorphic in all *B. campestris* accessions, with the exception of a single-banded variant of the normally duplicated *6Pgd-1* locus (Quiros 1987) detected in one shelgham individual (accession 103714, data not shown). With the exception of 6PGD, *B. tournefortii* shared no alleles in common with *B. campestris*, and was excluded from further analyses. It was not clear whether 6PGD was duplicated in *B. tournefortii*, however both *B. tournefortii* accessions tested were homozygous for the same alleles over all loci.

Seven loci were polymorphic within *B. campestris* (Table 2). The average number of alleles detected at RFLP loci was 3 times greater than the average number detected at polymorphic isozyme loci (8.5 versus 2.7 alleles/locus, respectively), and was positively correlated with average heterozygosity (H_0, Table 2) within an accession ($R^2 = 0.84$). Individual accessions varied in the average number of alleles (A, Table 2) detected, ranging from a maximum of 3.44 alleles per locus in toria (accession 1186) to a minimum of 1.0 in sarson.

Six of the twenty accessions carried more then 50% of the 42 alleles scored. These were toria accessions 1186 and 1390, shelgham, 'Yorii Spring' and 'Purple Top White Globe' turnips and 'Tendergreen' (derived from Table 1). Sarson accessions were the least polymorphic (7/42 possible alleles scored). Remaining accessions had between 24% and 48% of the maximum number of alleles represented in their populations. Both accessions of spp. *pekinensis* (i.e., 'Wong Bok' and 'Matsushima') were among the least polymorphic.

The majority of alleles in this species (28/42, 67%) were distributed across the entire geographic range. Only five alleles were unique to particular accessions and were found in accessions from all geographic regions. In the Indian center, the *CA12* allele *9* was present at a low frequency (4%) in toria accession 1390. Chinese cabbage cv 'Wong Bok' contained the only Asian unique allele (*BN55* allele *10*), which occurred at a relatively high frequency (41%). Three unique alleles from the European center were peculiar to the turnip-rape oilseed variety 'Torch': *Pgm-3* allele *2*, *BN33* allele *8* and *CA12* allele *11*.

For each pair of accessions, Nei's genetic identity and gene diversity estimates (Nei 1987) were computed (Table 3). Mean genetic identity summed across all accessions was 0.722 (SD = 0.122, $n = 190$). Genetic identities ranged from 0.94 between European turnip varieties 'Purple Top White Globe' and 'White Flat Dutch' to 0.37 between Chinese cabbage 'Wong Bok' and sarson accession 1394. Similarly, gene diversity estimates (a measure of heterozygosity between accessions) also showed a wide range of values that paralleled genetic identities (Table 3). For instance, gene diversity was least between 'Purple Top White Globe' and 'White Flat Dutch' and greatest between 'Wong Bok' and sarson accession 1394.

A UPGMA (unweighted pair-group method with arithmetic mean) dendrogram (Nei 1987) was constructed to better visualize relationships between accessions (Fig. 1). As expected from gene diversity estimates, the European turnips 'Purple Top White Globe' and 'White Flat Dutch' were the most closely associated. Other relationships were not as clear, but in general European accessions grouped more closely with each other than they did with Asian accessions. Some Indian accessions (shelgham, toria accession 1186) were more closely allied with European types, while toria accession 1390 grouped

Table 3. Nei identity (above diagonal) and gene diversity (below diagonal) for varieties listed in Table 1

	TCH	PTG	WFD	YS	SHO	SHL	RAB	TND	LC	CPC	KHC	TT	WB	MAT	MIZ	HTT	T2	T1	S1	S2
TCH	-	.863	.861	.663	.797	.859	.825	.816	.710	.744	.726	.653	.453	.448	.660	.692	.784	.749	.516	.516
PTG	.101	-	.943	.709	.903	.906	.854	.858	.738	.830	.835	.751	.530	.559	.778	.743	.862	.786	.616	.624
WFD	.104	.040	-	.663	.838	.877	.875	.818	.746	.800	.858	.724	.443	.465	.759	.737	.859	.762	.648	.648
YS	.231	.184	.221	-	.752	.722	.718	.810	.711	.613	.724	.640	.758	.697	.828	.656	.810	.860	.609	.577
SHO	.153	.069	.118	.166	-	.808	.778	.823	.690	.783	.810	.681	.509	.490	.736	.639	.785	.772	.595	.595
SHL	.101	.062	.084	.164	.131	-	.876	.915	.841	.871	.857	.847	.590	.621	.766	.835	.917	.807	.637	.641
RAB	.126	.097	.087	.173	.155	.078	-	.850	.736	.775	.853	.728	.546	.569	.826	.727	.855	.757	.645	.687
TND	.130	.091	.121	.108	.120	.051	.093	-	.824	.886	.849	.790	.604	.632	.785	.843	.849	.829	.585	.613
LC	.217	.185	.184	.191	.227	.109	.183	.119	-	.785	.787	.787	.746	.641	.670	.842	.809	.724	.506	.567
CPC	.201	.129	.153	.272	.168	.097	.167	.087	.165	-	.830	.818	.477	.616	.731	.799	.814	.697	.584	.584
KHC	.198	.111	.099	.172	.134	.092	.097	.095	.149	.127	-	.865	.543	.607	.820	.751	.865	.797	.744	.797
TT	.274	.188	.212	.255	.247	.114	.202	.153	.165	.147	.103	-	.528	.605	.746	.842	.850	.708	.696	.720
WB	.424	.344	.418	.170	.372	.288	.326	.272	.192	.414	.332	.375	-	.847	.619	.576	.587	.651	.365	.410
MAT	.441	.335	.416	.220	.399	.277	.322	.265	.281	.313	.296	.324	.123	-	.671	.631	.627	.641	.460	.460
MIZ	.240	.146	.164	.104	.182	.145	.111	.130	.225	.197	.117	.187	.271	.245	-	.653	.847	.811	.806	.763
HTT	.228	.178	.188	.222	.260	.110	.186	.104	.113	.153	.171	.121	.316	.285	.233	-	.832	.741	.523	.523
T2	.151	.089	.095	.109	.145	.050	.090	.087	.129	.136	.086	.113	.285	.270	.093	.111	-	.841	.786	.747
T1	.177	.140	.161	.084	.156	.119	.155	.102	.188	.220	.131	.213	.248	.265	.118	.173	.095	-	.613	.592
S1	.432	.332	.310	.325	.358	.307	.305	.343	.433	.378	.226	.279	.568	.495	.175	.413	.191	.327	-	.889
S2	.432	.326	.310	.349	.358	.304	.271	.322	.380	.378	.182	.257	.528	.495	.209	.413	.221	.344	.111	-

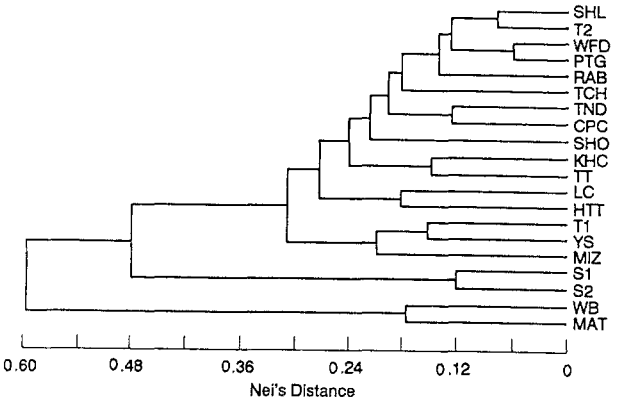

Fig. 1. UPGMA dendrogram constructed from gene diversity estimates within the *B. campestris* listed in Table 1

Table 4. Gene diversity statistics for *B. campestris* examined by marker class (isozyme or RFLP), crop type or geographical origin

Group	H_T	H_S	D_{ST}	G_{ST}	R_{ST}
Marker class					
All polymorphic isozymes	0.518	0.286	0.232	0.448	0.854
Pgi-2	0.467	0.253	0.214	0.458	0.891
Pgm-1	0.430	0.252	0.178	0.414	0.745
Pgm-3	0.658	0.354	0.304	0.462	0.905
All cDNA clones	0.684	0.408	0.276	0.403	0.711
BN8	0.337	0.210	0.127	0.378	0.640
BN33	0.659	0.397	0.262	0.397	0.694
BN55	0.869	0.497	0.372	0.428	0.788
CA12	0.874	0.531	0.343	0.392	0.680
Crop type					
Oilseed	0.405	0.206	0.199	0.491	1.206
Leafy vegetable	0.463	0.281	0.182	0.393	0.728
Turnip	0.435	0.336	0.099	0.228	0.368
Geographical origin					
Europe	0.364	0.294	0.069	0.191	0.314
India	0.403	0.238	0.166	0.411	0.871
India minus sarson	0.455	0.396	0.059	0.129	0.222
Asia	0.476	0.288	0.188	0.394	0.716
North Asia	0.475	0.288	0.187	0.394	0.780
South Asia	0.370	0.257	0.113	0.306	0.552
Asia minus *pekinensis*	0.449	0.293	0.155	0.346	0.606
All accessions and loci	0.477	0.277	0.120	0.419	0.760

more distantly from European types. All Asian varieties were also removed from the European cluster, and no clear patterns emerged among the Asian varieties. For instance, ssp. *chinensis* types were scattered among other Asian accessions. Also, while European turnip types clustered together, this association was absent among Asian turnip types. Both accessions of Chinese cabbage clustered together as did the sarson accessions.

Gene diversity statistics were calculated for a number of different combinations; either on the basis of (1) marker class (i.e., isozyme or RFLP), (2) crop type or (3) geographical origin (Table 4). Briefly, allele frequencies were used to calculate average heterozygosity estimates assuming Hardy-Wienberg equilibria. Total observed heterozygosity (i.e., gene diversity) between populations (H_T) was partitioned as within population diversity (H_S) and between population diversity (D_{ST}); where $H_T = H_S + D_{ST}$ (Nei 1987). Gene diversity between populations was expressed relative to total population diversity (as $G_{ST} = D_{ST}/H_T$) or relative to within sub-population gene diversity [as $R_{ST} = D_m/H_S$; where $D_m = sD_{ST}/(s\text{-}1)$ and s is the number of populations, Nei 1987].

For each of the seven polymorphic markers, variation within populations (H_S) was greater than variation between populations (D_{ST}, Table 4). Overall, intra-population diversity values (H_S) for RFLP (i.e., cDNA) markers were 50% higher than those of polymorphic isozymes. In terms of their ability to discriminate between populations, isozyme markers generally showed slightly higher gene differentiation values between populations (G_{ST} and R_{ST}) relative to cDNA markers. It was interesting that estimates for the linked markers *BN55* and *Pgm-3* (3 cM apart) showed higher levels of inter-population differentiation. Conversely, linked markers *BN8* and *BN33* (24 cM apart) were among the least differentiated between populations.

The distribution of gene diversity was assessed by crop use (i.e., oilseed, leafy vegetable or turnip). The lowest inter-population differentiation within a crop type was among turnip varieties ($R_{ST} = 0.37$, Table 4). The most highly differentiated types were oilseeds, whose value was 3 times that of turnip ($R_{ST} = 1.21$). These results probably reflected the influence of comparing the single subspecies of turnip versus an amalgamation of subspecies for oilseeds (e.g., ssp. *olifera*, ssp. *dichotoma* and ssp. *trilocularis*) and leafy vegetables.

Gene diversity was also considered relative to the geographic distribution of subspecies (see Table 1). Accessions were placed in one of three groups; Europe, India and adjacent regions to the northwest or Asia. Some within-area comparisons were also made. Overall, gene diversity estimates were highest between the Indian accessions, intermediate within the Asian group and lowest among the European types (Table 4). However, exclusion of sarson types from the Indian group drastically altered this hierarchy, and the Indian group was the least differentiated relative to the others (Table 4). This result indicated the influence of allele fixation and loss of heterozygosity due to the inbreeding habit of sarson on the calculated levels of gene diversity. Thus, with the exclusion of sarson as a special case, the most highly heterozygous types (i.e., those with the lowest gene diversity or least amount of allele fixation) were present in India,

followed relatively closely by those of the European division, and least heterozygous (i.e., more highly differentiated) types in Asia.

The Asian group was further subdivided geographically following historical records (Li 1969; Li 1983 see Table 1). Between Northern and Southern groups, gene diversity was lower among accessions considered of southern origin (morphologically most of these accessions could be considered related to the pak-choi group of Asian leafy vegetables) and higher among the more crop-diversified northern group (e.g., turnip, Chinese cabbage and spp. *japonica*). Since the ssp. *pekinensis* group clustered at an unexpected position in the dendrogram (Fig. 1), the exclusion of this group from analysis of gene diversity among Asian types did not appear to alter the results as dramatically as did the exclusion of sarson from the Indian group (Table 4).

Sarson is unique among *B. campestris* in that it is self compatible. However, the name is applied loosely and there are two recognized forms at an agricultural level, brown sarson and yellow sarson, differentiated primarily by seed coat color. The genetic structure of sarson was examined further. Six individuals from each of 15 accessions obtained as yellow sarson (including accessions 1394 and 1404) were tested for homozygosity at isozyme loci. Four of these accessions (designated as brown sarson in Table 1) appeared to carry multiple alleles at most loci, contrary to the inbreeding habit of sarson, and were not considered typical of ssp. *trilocularis*. In the remaining 11 accessions (yellow sarson in Table 1), each allele was homozygous at all loci, and only two alleles at each locus were detected among all accessions. Four of these yellow sarson accessions (1164, 1185, 1394, and 1404) were also tested at RFLP loci. Again, each was homozygous over all loci and only two alleles at a locus were detected. The only differences detected between accessions 1164 or 1185, and 1394 (Table 2) were the substitution of *Pgm3* allele *1* and *CA12* allele *6* in accession 1185. Amendment of the dendrogram (Fig. 1) with these additional sarson accessions did not influence the tree topology (data not shown).

Discussion

All forms of *B. campestris* studied were monomorphic for the duplicated locus *6Pgd-1* (Quiros 1987) as well as for *Pgm-2*. These results confirm that these widely distributed and morphologically diverse subspecies should be united at the species rank. It is interesting that the *Pgm-2* locus (the most anodic band) showed no variation among these varieties, a result also observed by Chen et al. (1990). In relation to the two polymorphic PGM loci present in *B. oleracea* (Arus and Orton 1983), *B. campestris* appears to have duplicated one or more PGM genes. Chen et al. (1990) were also able to detect a fourth PGM locus in *B. campestris* (in crosses between ssp. *olifera* and sarson), and it is possible that we scored their locus as an allele of *Pgm-3* unique to the European oilseed variety 'Torch'. If so, this additional locus represents a ssp. *olifera*-specific PGM gene duplication.

Three isozyme loci were polymorphic within *B. campestris* (*Pgi-2, Pgm-1* and *Pgm-3*) and showed an average of 2.7 alleles per locus. Each of the four cDNA clones used for RFLP analysis were also polymorphic, and the average number of alleles observed was nearly 3 times higher than that of the polymorphic isozymes. This result likely reflects the nature of the molecule detected, where isozymes are assayed for a functional protein and RFLPs would detect differences at the DNA level up to several kilobases from the coding region corresponding to the cDNA clone. No monomorphic RFLP markers were tested. While additional cDNA markers appeared single copy in two of these accessions (i.e., 'Yorii Spring' and 'Kwan Hoo Choi'), no polymorphisms were detected, and their genetic basis remains uncertain (McGrath and Quiros, unpublished). Thus, these results likely overestimate the absolute level of gene diversity within this species but mirror relative relationships between subspecies. As *B. campestris* shows a high level of gene duplication (McGrath and Quiros 1991 b), it was important that RFLP loci used in this study were demonstrably single copy loci to avoid comparisons between paralogous genes in different subspecies. While each of these loci have been mapped in one inter-subspecies cross (McGrath and Quiros 1991 b), fixed heterozygous gene duplications in some accessions could also explain observed allele frequencies of 50% (e.g., *Pgm-1* in 'Raab' and *BN8* in 'Hong Tsai Tai'). At least for the more frequent alleles (which factor strongly in diversity estimates) the present sample size appears adequate to draw preliminary conclusions regarding the distribution of genetic diversity across this eclectic range of *B. campestris* subspecies. Although the number of plants per accession tested here would not have been sufficient to detect rare alleles with high confidence, comparison of the isozyme gene frequency values of 'Torch' observed here agree well with a more extensive survey of this variety by Truco (1986).

Most alleles are present in more than one crop type and in more than one geographical center, even with the highly polymorphic cDNA loci. With the exception of sarson and Chinese cabbage, a continuum of isozyme and RFLP variability existed across the species. On a gross level, genetic variability tended to follow geography. For instance, European varieties were more related with each other, along with some Indian accessions, than they were with the Asian group. The Asian types tended not to fall into well-defined groups and were interspersed with other Indian varieties. Both sarson and Chinese cabbage were exceptional in that they grouped distantly from all other subspecies.

Considering crop type as the unit of analysis, oilseeds show nearly 3 times greater gene diversity than do turnip types. This is remarkable since these turnips enjoy a wider distribution than do the oilseeds. The comparison is misleading, however, and illustrates the effect of allele fixation on values of genetic diversity in out-crossing taxa. Unlike all other *B. campestris* subspecies, the present evidence suggests that the self-fertile Indian oilseed sarson is predominantly inbred. Thus, inherent assumptions of random mating made by most genetic diversity algorithms are violated (Nei 1987; Felsenstein 1981). Similarly, the single accession of a European oilseed tested probably violates this assumption as well, as it was derived from Polish introductions to Canada for breeding low erucic acid cultivars (Downey et al. 1975) and has been driven to fixation for alleles of *BN8* and *BN55*. Peculiarities of its breeding history may have increased the frequency of three rare allelic variants undetected in other subspecies. Of the three oilseed types, only toria seems to have escaped this reduction of heterozygosity. Since each allele in sarson was also present in toria, inbreeding sarson types may form a monophyletic group derived from a single self-compatible 'toria-like' individual. However, chloroplast DNA restriction patterns suggest sarson has diverged from both European and Asian cultivars (Palmer et al. 1983).

Asian leafy vegetables show remarkable morphological diversity. From the analysis of gene diversity here it is difficult to discriminate the pattern of their diversity. Estimates of genetic variability were slightly higher for leafy vegetables as a crop type (i.e., including Asian and European types) relative to the estimates for southern Asia (i.e., primarily pak choi type vegetables). While Chinese cabbage types (which are the primary north Asian leafy vegetables) contribute slightly to the overall diversity within the Asian group, it appears they have suffered a reduction of heterozygosity. The appearance of Chinese cabbage in the historical record occurs after turnip in northern China and pak choi in the south were important agricultural crops, and occurred in a region of overlap (Li 1969; Li 1981). Meager evidence suggests that Chinese cabbage may have resulted from an inter-subspecies hybrid of turnip and pak-choi (Li 1981; Song et al. 1990), or perhaps as a variant of an Asian turnip. In support of the latter interpretation, alleles of *CA12* are not shared between ssp. *chinensis* (pak choi) and Chinese cabbage, but are shared between Asian turnips and Chinese cabbage. Also, inter-subspecies hybrids of turnip and Chinese cabbage are more vigorous than either parent (and have been used as a high yield forage crop, Werner 1984), unlike hybrids between either pak choi and turnip or Chinese cabbage (McGrath and Quiros 1991 a).

Turnip types are distributed across the Old World and, as one of the few native temperate root crops, may have been among the first *B. campestris* types utilized. Based on morphology, Sinskaia (1928) classified turnip varieties into seven groups based on regional variation from Europe to Central Asia. Contemporary turnips were considered either as originating in Central Asia from 'primitive types' or derived from two independent domestication events in Europe and Asia (Sinskaia 1928). However, Nishi (1980) considered the migration of turnip to Asia as an agricultural crop that entered China through Siberia and Mongolia. Independent migrations of turnip to Japan, one from China and another from Europe via Siberia or Manchuria, suggest a further complexity in tracing the evolution of turnip (Aoba 1970, cited in Nishi 1980), and by extension all *B. campestris* types.

Crop gene diversity estimates summed across all geographic regions were lowest for the turnip types, relative to leafy vegetables and oilseeds. Thus, turnips appear to be less diversified within a region than other crop types. From the varieties tested here, the Indian turnip shelgham appeared to be minimally affected by directed breeding efforts. Variation in the presence of an enlarged hypocotyl, in root color and size, a variant phenotype of the highly conserved *6Pgd-1* phenotype, a putative self-compatible plant, and a high level of heterozygosity were observed. The other turnip types tested were morphologically uniform within a variety but still retained a high level of biochemical variation.

With the exception of turnip, partitioning genetic diversity into crop type alone is a less useful predictor of subspecies relationships than considering genetic diversity by geographical regions. Given that gene diversity estimates reflect relative levels of heterozygosity and that decreasing heterozygosity through genetic drift or by directed breeding operations leads to increasing gene diversity estimates, it seems reasonable to surmise that regions with the lowest gene diversity contain the most heterozygous forms, i.e., those approaching a level that could be considered a 'primitive' stage. Following this line of reasoning, varieties of the Central Asia/India subgroup have the lowest between-group gene diversities (excluding sarson), followed closely by Europe and more distantly by the Far East. Since at least one accession of toria and shelgham appear to be closely related and few genes control the enlarged hypocotyl of turnip (Ragionieri 1920, cited in Yarnell 1956; McGrath and Quiros 1991 a), variation in natural populations for the presence or absence of the turnip character may have characterized stocks from which modern *B. campestris* was derived.

The mode of domestication (e.g., single origin with subsequent migration versus multiple origins from a widely dispersed wild stock), the timing of domestication (early or recent) and the distribution of genetic variability in *B. campestris* bear on our understanding of the agricultural origins and conservation of *Brassica* germplasm. A linear mode of domestication for *B. campestris* subspecies has been suggested by Song et al. (1990). From an origin in Europe, germ plasm is suggested to have migrated first to India and then to Asia. Selection of new crop types and their development was then performed by indigenous cultures after each successive migration. An alternative scheme, whereby indigenous cultures would each have access to primitive germplasm (perhaps centered in Central Asia) may also explain the utilization and domestication of these subspecies. Our results suggest neither model adequately explains the distribution of genetic diversity, but instead supports particular aspects of both models.

Acknowledgements. We are indebted to D. Cohen, P. Williams and C. Gomez-Campo for gifts of seed. We also thank S. Kianian, V. D'Antonio and P. This for their technical assistance, and L. van Raamsdonk and R. Raguso for critical reviews. This

work was supported by USDA competitive grants 88-37262-4018 and 90-37140-5425 to CFQ.

References

Arus P, Orton TJ (1983) Inheritance and linkage relationships of isozyme loci in *Brassica oleracea*. J Hered 74:405–412

Chen BY, Heneen WK, Simonsen V (1990) Genetics of isozyme loci in *Brassica campestris* L. and in the progeny of a trigenomic hybrid between *B. napus* L. and *Brassica campestris* L. Genome 33:433–440

De Candolle A (1886) Origin of cultivated plants. Hafner Publ Co, New York 1964 reprint

Denford KE, Vaughan JG (1977) A comparative study of certain seed isoenzymes in the ten chromosome complex of *Brassica campestris* and its allies. Ann Bot 41:411–418

Downey RK, Stringam GR, McGregor DI, Stefansson BR (1975) Breeding rapeseed and mustard crops. In: Harapiak JT (ed) Oilseed and pulse crops in Western Canada: a symposium. Western Co-operative Fertilizers, Calgary, pp 157–183

Felsenstein J (1981) Evolutionary trees from gene frequencies and quantitative characters: finding maximum likelihood estimates. Evolution 35:1229–1242

Li CW (1981) The origin, evolution, taxonomy and hybridization of Chinese cabbage. In: Talekar NS, Griggs TD (eds) Chinese cabbage: Proc. 1st Int Symp Asian Veg Res Dev Center. Taiwan, pp 3–9

Li H-L (1969) The vegetables of ancient China. Econ Bot 23:253–260

Li H-L (1983) The domestication of plants in China: ecogeographical considerations. In: Keightley DN (ed) The origins of Chinese civilization. University of California Press, Berkeley, California, pp 21–63

McGrath JM, Quiros CF (1991a) Inter-subspecies hybrids and their progeny in *Brassica campestris* L. J Am Soc Hortic Sci 116:349–355

McGrath JM, Quiros CF (1991b) Inheritance of isozyme and RFLP markers in *Brassica campestris* and comparison with *B. oleracea*. Theor Appl Genet 82:668–673

McGrath JM, Quiros CF, Harada JJ, Landry BS (1990) Identification of *Brassica oleracea* monosomic alien chromosome addition lines with molecular markers reveals extensive gene duplication. Mol Gen Genet 223:198–204

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

Nishi S (1980) Differentiation of *Brassica* crops in Asia and the breeding of 'Hakuran', a newly synthesized leafy vegetable. In: Tsunoda S, Hinata K, Gomez-Campo C (eds) *Brassica* crops and wild allies: biology and breeding. Jpn Sci Soc Press, Tokyo, pp 133–150

Olsson G (1954) Crosses within the *campestris* group of the genus *Brassica*. Hereditas 40:398–418

Palmer JD, Shields CR, Cohen DB, Orton TJ (1983) Chloroplast DNA evolution and the origin of amphidiploid *Brassica* species. Theor Appl Genet 65:181–189

Prakash S, Hinata K (1980) Taxonomy, cytogenetics and origin of crop *Brassicas*, a review. Opera Bot 55:1–57

Quiros CF (1987) Duplicated isozyme loci and cellular compartmentalization of their products in *Brassica*. Cruciferae Newsl 12:24

Quiros CF, Ochoa O, Kianian S, Douches D (1987) Analysis of the *Brassica oleracea* genome by the generation of *B. campestris-oleracea* chromosome addition lines. Theor Appl Genet 74:758–766

Renfrew JM (1973) Paleoethnobotany: the prehistoric food plants of the Near East and Europe. Columbia Universiy Press, New York

Sinskaia EN (1927) Geno-systemical investigations of cultivated *Brassica*. Bull Appl Bot Plant Breed 17:1–166

Sinskaia EN (1928) The oleiferous plants and root crops of the family Cruciferae. Bull Appl Bot Plant Breed 19:1–648

Song KM, Osborn TC, Williams PH (1990) *Brassica* taxonomy based on nuclear restriction fragment polymorphisms (RFLPs). 3. Genome relationships in *Brassica* and related genera and the origin of *B. oleracea* and *B. rapa* (syn. *campestris*). Theor Appl Genet 79:497–506

Truco MJ (1986) Estudie de la variabilitat isoenzimatica en *Brassica campestris*. Thesis, University of Barcelona, Spain

Werner K (1984) New turnip greens variety wins grower fans. Am Vegetable Grower 32:16–17

Whyte RO (1983) The evolution of the Chinese environment. In: Keightley DN (ed) The origins of Chinese civilization. University of California Press, Berkeley, Calif., pp 3–19

Yarnell SH (1956) Cytogenetics of the vegetable crops. II. Crucifers. Bot Rev 22:81–166

Zohary D, Hopf M (1988) Domestication of plants in the old world. Clarendon Press, Oxford