ITS sequence variation in wild species and cultivars of pea

Polans, N.O. and Saar, D.E. Dept. of Biological Sci. and Plant Molecular Biology Center
Northern Illinois Univ., Dekalb, IL

An often powerful approach to characterizing the relationships among plant taxa is to compare the nucleotide sequences of their ribosomal DNA. Nuclear ribosomal DNA (nrDNA) is organized as distinct chromosomal units that are repeated thousands of times in most higher plant genomes. Each of these units contains the genes that encode the 18S, 5.8S and 26S ribosomal RNA subunits, as well as several different spacer DNA regions. The nucleotide sequence variation found in both of the internal transcribed spacer regions (ITS-1 and ITS-2, Fig. 1) is routinely used for the systematic analysis of closely related taxa, at least in part due to the high rate of evolutionary change characterizing these DNA regions (1).

In our preliminary study of pea ITS regions (6), ITS-1 and ITS-2 DNA sequence variation was assessed for five pairs of wild and cultivated pea taxa selected to approximate the range of *Pisum.* The objective of that investigation was to examine the similarity of the sequences within paired accessions, the overall level of genetic variation found across the entire genus, and the topological relationships established among the five selected groups of taxa. It resulted in the following six observations: 1) very close genetic affinities throughout *Pisum*, with *P. fulvum* exhibiting the greatest degree of divergence, 2) support for the established taxonomic categories of the genus based upon identical or near identical sequences within group pairs, 3) the assignment of JI1794 as a "northern" *humile*, 4) the validity of northern and southern *humile* as closely-related, but distinct, lines, 5) the apparent independent evolution of a pea chromosomal translocation and 6) a close relationship between *elatius* and the cultivated *sativum.* Additionally, when *Vicia montbrettii* was included as an outgroup to *Pisum* in both the preliminary and present studies, phylogenetic analyses indicated that *P. fulvum* remained not only the most divergent pea taxon but also the most basal taxon relative to the *sativum* group (data not shown).

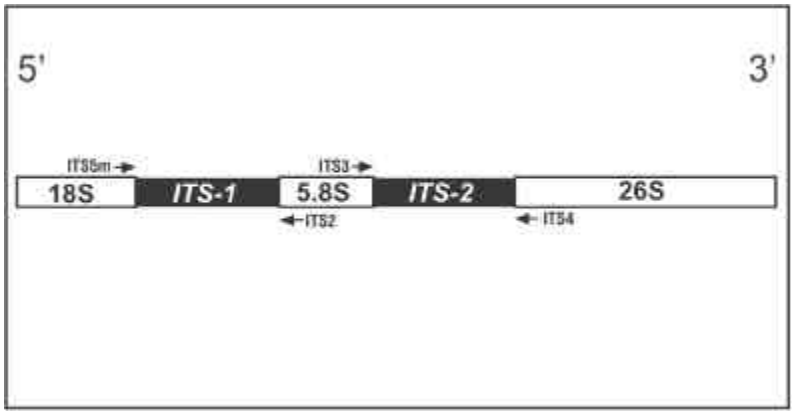

Fig. 1. The three coding and two internal transcribed spacer regions of the nuclear ribosomal DNA repeat unit of a typical angiosperm (not drawn to scale). Arrows indicate approximate locations of the four primers used for PCR amplification.

The goal of the present study is to extend the use of ITS variation as a comparative tool to an additional 55 wild and cultivated pea taxa, both to validate our preliminary findings among a more diverse sample of the genus and to include previously unexamined pea types in these analyses.

Materials and Methods

Pisum isolates 701-722 are from the Ben Ze'ev and Zohary (1973) collection (courtesy of J.G. Waines), JI accessions are from the John Innes collection (courtesy of M. Ambrose), cv. Alaska is from J. Mollema and Son, Inc. (Grand Rapids, MI) and cv. (Morse's) Progress #9 is from Ferry-Morse Seeds (Mountain View, CA). *P. sativum* Syriacum was graciously provided by R. Jorgensen, and accessions 82-14n, A1078-234 and PI 179449 were kindly provided to this project by G. Marx and N. Weeden.

DNA extraction, PCR amplification, gel purification, and ITS primers (ITS2, ITS3, ITS4 and ITS5m) are described elsewhere (6). DNA sequencing is performed with either an Applied Biosystems model 373 DNA sequencer or a Beckman Coulter CEQ 2000 XL DNA analysis system. Forward and reverse DNA sequences are compared to resolve ambiguities using PC Gene software and the resulting sequences are aligned with the Clustal X computer program. Sequence data are analyzed using the PAUP computer package (7).

Results and Discussion

The pea 18S rRNA, ITS-1, 5.8S rRNA, ITS-2 and 26S rRNA regions examined in this study contain 27, 238, 164, 213 and 22 alignable base pairs (bp), respectively, totaling 664 bp (including 451 bp of spacer DNA) for all but one of the 65 plants analyzed. The only exception to these results involves a *P. sativum* Syriacum accession that contains an additional guanine at ITS-2 position number 582. Ambiguous or polymorphic pyrimidine and purine sites are denoted by the IUPAC/IUB symbols "Y" and "R," respectively. Of the 664 total bp sequenced for each of the individual plants, 640 (>96%) of these sites are constant among the 64 pea taxa. Of the 451 ITS bp sequenced, 428 (>94%) of these sites are constant. Only 24 of the total sites are polymorphic (and only 21 are parsimony informative), reaffirming both the very close evolutionary relationships that must exist within the genus and the limited ITS information available with which to differentiate pea taxa. In this study, ITS-1 contains 14 of the polymorphic sites, as compared with nine found for ITS-2 and one polymorphic site located just within the 5.8S rRNA coding region (Table 1).

Table 1. Variable ITS sites for wild and cultivated taxa of pea.

	Nucleotide Position*		
	ITS-1 ITS-2	Number	
	111111111122222445566666	Changes	GenBank
	011233334903346570300023	from	Accession
	358425895084607900001411	fulvum	Numbers
Pisum fulvum			
701	GTTGGGCACCGACTGTTCTTGAAG		AF305582
			AF305920
702	GTTGGGCACCGACTGTTCTTGAAG		AF305583
			AF305921
703	GTTGGGCACCGACTGTTCTTGAAG		AY143432
706	GTTGGGCACCGACTGTTCTTGAAG		AY143433
707	GTTGGGCACCGACTGTTCTTGAAG		AY143434
708	GTTGGGCACCGACTGTTCTTGAAG		AY143435
JI224	GTTGGGCACCGACTGTTCTTGAAG		AY143447
JI1006	GTTGGGCACCGACTGTTCTTGAAG		AY143451
Pisum sativum			
ssp. humile (northern)			
716	GTCGGGCGCTACCCACCCATGTAC	11	AF305586
			AF305924
JI1794	GTCGGGCGCTACCCACCCATGTAC	11	AF305587
			AF305925
ssp. humile (southern)			
711	RYCRAACGCTACCCACCCATGAAC	12	AY143436
712	RYCRGACGCTACCCACCCATGAAC	11	AF305584
			AF305922
713	RYCRAACGCTACCCAYCCATGAAC	12	AF305585
			AF305923

	Nucleotide Position*		
	ITS-1 ITS-2	Number of Base	
	11111111112222244556666	Changes	GenBank
	01123333490334657030023	from	Accession
	35842589508460790001411	fulvum	Numbers
714	RYCGGACGCTACCCACCCATGAAC	11	AY143437
ssp. elatius			
721	GCCGTACGYTACCCACCCATGTAC	14	AF305588 AF305926
722	GCCGTACGYTACCCACCCATGTAC	14	AF305589 AF305927
723	GCCGAACGCTACCCACCCATGTAC	14	AY143438
JI64	GCCGGACGCTACCCACCCATGTAC	13	AY143442
JI261	GCCGAACGCTACCCACCCATGTAC	14	AY143450
JI2201	GCCGAACGCTACCCACCCATGTAC	14	AY143455
ssp. abyssinicum			
JI2	GCCGAACGCTACCCACCCATGTAC	14	AY143441
JI130	GCCGGACGCTACCCACCCATGTAC	13	AY143444
JI225	GCCGGACGTTACCCACCCATGTAC	14	AY143448
JI2202	GCCGGACGTTACCCACCCATGTAC	14	AY143456
ssp. sativum			
JI196 Georgia	GCCGAAYGCTACCCACCCATGTAC	14	AY143463
JI228 Bolivia	RCCGAACGCTACCCACCCATGTAC	14	AY143466
JI245 Russia	GCCGAACGYTAYCCACCCATGTAC	14	AY143467
JI1035 Turkey	GCCGAACGCTACCCACCCATGTAC	14	AY143473
JI1057 Antioquia I Chilena	GCCGAACGCTACCCACCCATGTAC	14	AY143474
JI1345 Mongolia	GCCGAACGYTACCCACCCATGTAC	14	AY143476
JI1428(P. tibetanicum)	GCCGAACGYTACCCACCCATGTAC	14	AY143478
JI1835 Spain	GCCGAACGYTACCCACCCATGTAC	14	AY143481
JI2116(P. speciosum)	GCCGAACGCTACCCACCCATGTAC	14	AY143482
JI2124 ponderosum	GCCGAACGCTACCCACCCATGTAC	14	AY143483
JI2265 Primitive Albanian	GCCGAAYGYTACCCACCCATGTAC	14	AY143484
JI2385(P. sp. Yemen)	GCCGGACGCTACCCACCCATGTAC	14	AY143485
82-14n	GCCGAACGCTACCCACCCATGTAC	14	AY143457
JI185 Wiraig	GCCGAACGTTAYCCACCCATRTAC	15	AY143462
JI263 Balkans	ACCGAACGYTAYCCACCCATGTAC	15	AY143469
JI264 Greece	RCCGAACGTTAYCCACCCATGTAC	15	AY143470
JI711 Austrian Winter	ACCGAACGCTACCCACCCATGTAC	15	AF305590 AF305929
JI787 Minerva	GCCGAATGYTACCCACCCATGTAC	15	AY143471
JI1372 Mummy Pea	ACCGAACGYTACCCACCCATGTAC	15	AY143477
JI1758 Nepal	GCCGAACGTTAYCCACCCATRTAC	15	AY143480
JI2438 Partridge	ACCGAACGYTAYCCACCCATGTAC	15	AY143486
Alaska	ACCGAACGYTACCCACCCATGTAC	15	AF305202 AF305928
PI179449	RCCGAACGTTACCCACCCATGTAC	15	AY143440
Syriacum	GCCGAAYGTTACCCACCCATGTAC	15	AY143459
JI85 Afghanistan	ACCGAACGTTACCCACCCATGTAC	16	AY143443
JI156 Sudan	ACCGAACGTTACCCACCCATGTAC	16	AY143445
JI159 Ethiopia	ACCGAACGTTAYCCACCCATRTAC	16	AY143460

	Nucleotide Position*		
	ITS-1 ITS-2	Number of Base	
	11111111112222244556666	Changes	GenBank
	01123333490334657030023	from	Accession
	35842589508460790001411	fulvum	Numbers
JI181 Keerau Pea	GCCGAACGTTATCCACCCATRTAC	16	AY143461
JI207 choresmicum	ACCGAACGTTAYCCACCCATRTAC	16	AY143464
JI209 arvense	ACCGAACGTTAYCCACCCATGTAC	16	AY143465
JI250 (P. jomardii)	ACCGAACGTTAYCCACCCATGTAC	16	AY143468
JI1578 China	ACCGAACGTTAYCCACCCATGTAC	16	AY143479
Progress#9	ACCGAACGTTAYCCACCCATGTAC	16	AY143458
A1078-234	ACCGAACGTTACCCACCCATGTAC	16	AY143439
JI1033 India	GCCGAACGTTATCCACCCATATAC	17	AY143472
JI1089 Syriacum	ACCGAACGTTATCCACCCATRTAC	17	AY143475
Inconsistent assignments:			
JI241 (1)	ACCGGACGTTACCCACCCATGTAC	15	AY143449
JI198 (2)	GCCGAACGTTACCCACCCATGTAC	15	AY143446
JI1398 (2)	ACCGAACGTTACCCACCCATGTAC	16	AY143453
JI1096 (3)	ATCGAACGCTACTCACCTACGTTC	18	AY143452
JI2055 (3)	GTCGAACGCTACTCACCTACGTTC	17	AY143454

*** In the 5'->3' direction (see Fig. 1) beginning with those bases nearest primer ITS5m. Position 267 is assigned to the 5.8S rRNA coding region.**

(1) JI241 is listed as ssp. humile, but it displays ssp. sativum ITS characteristics.

(2) JI198 and JI1398 are listed as ssp. elatius, but they display ssp. sativum ITS characteristics.

(3) JI1096 and JI2055 are listed as ssp. elatius, but they display unique ITS variation at several sites

Parentheses around four JI accessions indicate taxonomic nomenclature not supported in this table.

A compilation of the 24 variable nrDNA sites is delineated for all 65 pea taxa in Table 1, accompanied by corresponding GenBank accession numbers for the retrieval of complete sequences. The table is organized in accordance with the two commonly recognized species of pea (2-4), the more divergent *P. fulvum* and the typically cultivated *P. sativum.* The former is represented by eight identical nrDNA sequences, while the latter is differentiated as four subspecies: *humile, elatius, abyssinicum* and *sativum.* Subspecies *humile* is further subdivided by northern and southern populations as described by (2). There are five pea accessions characterized as questionable taxonomic assignments solely based on their nrDNA variation, and there are also differences distinguishing from one another the two "Syriacum" accessions surveyed. The four subspecies and 52 assigned accessions of *P. sativum* are further arranged in Table 1 by the number of unambiguous base changes each possesses relative to the invariant *P. fulvum* accessions. The number of base differences separating *fulvum* from the 52 *sativum* accessions ranges from 11 to 17, with 10 of these sites being unique to *fulvum.* JI1096, an *elatius* accession displaying unique ITS variation at several sites, shows 18 base differences with *fulvum*. The subdivisions of *P. sativum* are listed in the following order based on their base pair differences with *fulvum*: northern *humile* (11 base changes), southern *humile* (11-12 base changes), *elatius* and *abyssinicum* (13-14 base changes each), and *sativum* (14-17 base changes). Named cultivars of *sativum* usually display 15 or 16 base changes.

A Neighbor Joining (NJ) distance analysis of these data is presented in Fig. 2 to provide a basic illustration of the associations suggested in Table 1, while also including such influences as the multiple polymorphisms

found at ITS-1 sites 132 and 234. No attempt is made, however, to infer evolutionary relationships among the 65 taxa, given the relatively few parsimony informative sites available to the analysis. In the figure, only *fulvum*, northern and southern *humile* and a pair of *elatius* accessions maintain distinct group associations. Ten of the 21 parsimony informative sites differentiate *fulvum* from the much larger *sativum* ingroup. Within *sativum*, the two northern *humile* accessions display completely identical nucleotide sequences (at 664 sites), while the southern *humile* differ at a single site and show ambiguity at several others. Only two *elatius* accessions (JI 1096 and JI 2055), displaying four unique sites and the largest overall numbers of sequence differences with *fulvum*, group separately from the remaining *sativum* subspecies. These remaining accessions group roughly based on possessing 14, 15 or 16 base differences with *fulvum*. Most of the other *elatius* and all four *abyssinicum* are found in the first group, along with approximately a dozen *sativum* and the single questionable *humile* accession. The latter two groups principally comprise *sativum*, including most of the named cultivars.

According to Fig. 2, *elatius* and *abyssinicum* are the closest taxa to the cultivated *sativa*, despite the fact that northern *humile* has been postulated the closest wild progenitor of the cultivated pea based in part on a shared chromosomal translocation (2) and detailed chloroplast studies (5). Other, larger data sets (not shown) place northern *humile* closer to *sativum*, but they do not support northern *humile* as the taxon closest to the cultivars. Thus, the present study largely supports the conclusions from our previous work (6): generally very close relationships within *Pisum*, with *P. fulvum* clearly displaying the greatest divergence; JI 1794 classified as a "northern" *humile*; northern and southern *humile* as closely-related, but distinct, taxa; and the independent evolution of a pea chromosomal translocation. The study also supports distinct taxonomic categories for *fulvum* and for northern and southern *humile*; however, the ITS sequence variation obtained from this investigation is too limited to separate unambiguously the very close relationships among *elatius*, *abyssinicum* and *sativum*. Further efforts are needed to resolve these relationships and to clarify the taxonomic assignments of the few questionable accessions addressed in this study.

Acknowledgement: We thank Scott Grayburn for his DNA sequencing skills. This work was supported by funds from the Department of Biological Sciences and the Plant Molecular Biology Center, Northern Illinois University.

1. Baldwin, B.G., Sanderson, M.J., Porter, J.M., Wojciechowski, M.F., Campbell, C.S. and Donoghue, M.J. 1995. Ann. Mo. Bot. Gard. 82: 247-277.
2. Ben Ze'ev, N. and Zohary, D. 1973. Israel J. Bot. 22: 73-91.
3. Hoey, B.K., Crowe, K.R., Jones, V.M. and Polans, N.O. 1996. Theor. Appl. Genet. 92: 92-100.
4. Marx, G.A. 1977. In: Sutcliffe, J.F. and Pate, J.S. (eds.). Physiology of the Garden Pea. Academic Press, New York, pp 21-43.
5. Palmer, J.D., Jorgensen, R.A. and Thompson, W.F. 1985. Genetics 109: 195-213.
6. Saar, D.E. and Polans, N.O. 2000. Pisum Genetics 32: 42-45.
7. Swofford, D.L. 1998. PAUP, Version 4.0b4a. Sinauer Associates, Sunderland, MA.

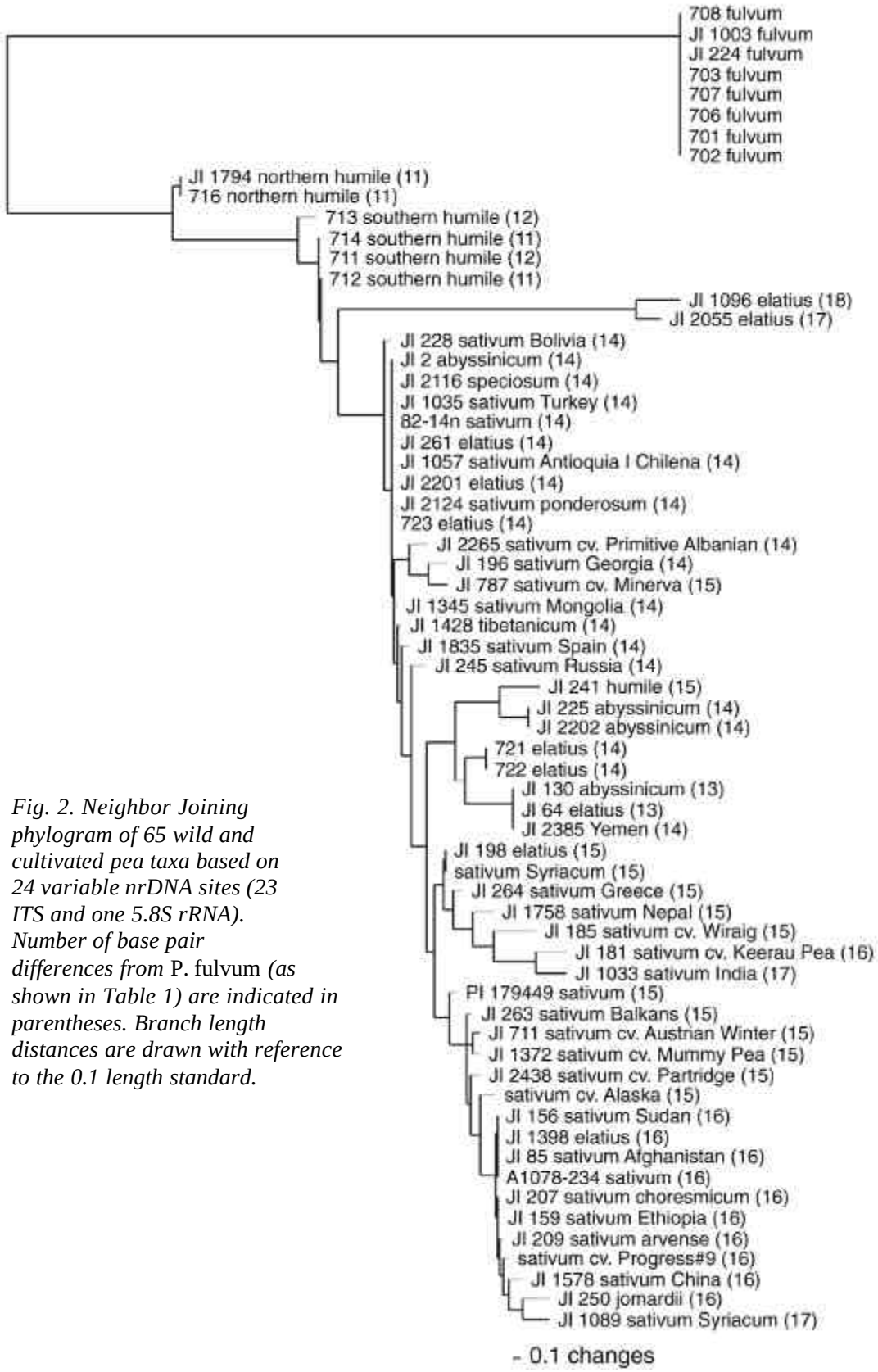

Fig. 2. Neighbor Joining phylogram of 65 wild and cultivated pea taxa based on 24 variable nrDNA sites (23 ITS and one 5.8S rRNA). Number of base pair differences from P. fulvum *(as shown in Table 1) are indicated in parentheses. Branch length distances are drawn with reference to the 0.1 length standard.*