

Genetic diversity of 38 spinach (*Spinacia oleracea* L.) germplasm accessions and 10 commercial hybrids assessed by TRAP markers

Jinguo Hu · Beiquan Mou · Brady A. Vick

Received: 18 April 2006 / Accepted: 20 November 2006 / Published online: 16 February 2007
© Springer Science+Business Media B.V. 2007

Abstract Target region amplification polymorphism (TRAP) markers were used to assess genetic variability among 38 germplasm accessions and 10 commercial hybrids of spinach (*Spinacia oleracea* L.), an economically important leafy vegetable crop in many countries. Germplasm accessions with different geographic origins and 10 commercial hybrids were examined. For assessing genetic diversity within accessions, DNA was extracted from 12 individual seedlings from six germplasm accessions and two hybrids. A relatively high level of polymorphism was found within accessions based on 59 polymorphic TRAP markers generated from one fixed primer derived from the *Arabidopsis*-like telomere repeat sequence and two arbitrary primers. For evaluating interaccession variability, DNA was

extracted from a bulk of six to 13 seedlings of each accession. Of the 492 fragments amplified by 12 primer combinations, 96 (19.5%) were polymorphic and discriminated the 48 accessions from each other. The average pair-wise genetic similarity coefficient (Dice) was 57.5% with a range from 23.2 to 85.3%. A dendrogram indicated that the genetic relationships among the accessions were not highly associated with the geographic locations in which the germplasms were collected. The seven commercial hybrids were grouped in three separate clusters, suggesting that the phenotype-based breeding activities tended to reduce the genetic variability. This preliminary study demonstrated that TRAP markers are effective for fingerprinting and evaluating genetic variability among spinach germplasms.

Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the US Department of Agriculture.

The U.S. Government's right to retain a non-exclusive, royalty-free license in and to any copyright is acknowledged.

J. Hu (✉) · B. A. Vick
USDA-ARS, Northern Crop Science Laboratory,
1307 18th Street North, Fargo, ND 58105-5677, USA
e-mail: huj@fargo.ars.usda.gov

B. Mou
USDA-ARS, 1636 East Alisal Street, Salinas, CA
93905, USA

Keywords DNA fingerprinting · Genetic diversity · Genetic relationships · *Spinacia oleracea* L. · Target region amplification polymorphism (TRAP)

Introduction

Spinach (*Spinacia oleracea* L.) is an economically important leafy vegetable crop in many countries, and over 800,000 hectares are grown annually worldwide (FAO 2005). In the US, the farm value of the US spinach crop (fresh and processing) averaged \$162 million during 2000–2002, with

fresh market spinach accounting for 91% (Lucier and Plummer 2004). US market spinach is divided into three main types based on leaf morphology: (1) *savoy* (wrinkled leaf) with dark green crinkly, curly leaves; (2) *semi-savoy* (semi-wrinkled) with slightly crinkled leaves; and (3) *smooth* (flat leaf) with unwrinkled, spade-shaped leaves. Savoy and semi-savoy types are popular in the fresh market, while smooth-flat types are better suited for processing (including fresh-cut) because their wrinkle-free leaves are easier to wash (Le Strange et al. 2003). Despite substantial progress in the genetic improvement of spinach, genetic variability of the germplasm at the molecular level has not been investigated.

TRAP is a relatively new marker technique (Hu and Vick 2003) which uses a fixed primer with a known sequence in combination with arbitrary primers to generate polymorphic markers. Studies applying TRAP to lettuce cultivar genotyping (Hu et al. 2005) and sugarcane germplasm diversity assessment (Alwala et al. 2006) suggested that TRAP was capable of identifying DNA polymorphisms for fingerprinting cultivars and germplasm accessions. In this paper, we report the preliminary assessment of genetic diversity among 38 spinach germplasm accessions and 10 commercial hybrids using the TRAP technique.

Materials and methods

Plant materials

A random sample of 38 accessions obtained from the USDA National Plant Germplasm System (NPGS) was used in the current study. The PI number, names of the cultivars and the geographic origin of these accessions are listed in Table 1. In addition, 10 commercial F₁ hybrids grown in the US and developed by six seed companies were also included. For confidentiality reasons, the names of the commercial F₁ hybrids were coded CH-A through CH-J and the seed companies were coded as A–F.

DNA extraction

DNA was obtained from leaf tissue of 2-week-old seedlings. A DNeasy 96 Plant Kit (QIAGEN,

Table 1 List of the 38 accessions used in the current study

Accession number	Name or ID	Origin
NSL 6082	Wisconsin Blooms	New York, US
NSL 6782	Hollandia	Netherlands
PI 608762	K-17068	Thailand
Ames 20168	Da Yuan Ye Bo Cai	China
PI 608712	SPI 153/89 (S. tetrandra)	Germany
NSL 68263	Indures	New York, US
PI 164966	Harlan 42	Turkey
PI 167434	Cavallius	Belgium
PI 179507	Beledi	Syria
PI 179508	Harlan 9889	Iraq
PI 179595	Victoria	Belgium
PI 193618	Neusseelander	Ethiopia
PI 193619	Victoria	Ethiopia
PI 217425	CGN 14178	South Korea
PI 222270	Esfenaj	Iran
PI 254565	Polack	Afghanistan
PI 261787	Monstrueux de Vi	France
PI 261788	Geant d'Hiver	France
PI 262911	Espinaca nobel	Spain
PI 266926	Universal	Germany
PI 285751	Koda	Warszawa, Poland
PI 296393	CGN 9662	Iran
PI 358248	Ohridski	Yugoslavia
PI 379549	Stipski	Yugoslavia
PI 418978	Chung Pu Lao	Shaanxi, China
PI 445785	Baladi	Syria
PI 449353	Universal	Turkey
PI 499372	Ispolinskij	Former Soviet Union
PI 499373	Godir	Former Soviet Union
PI 531448	Eszkimo	Hungary
PI 531449	Hegykoj	Hungary
PI 531452	Matador Darko	Denmark
PI 604785	SPI 161/86	Mongolia
PI508504	Summer Green	South Korea
PI527332	Da Jian Ye Bo Cai	China
NSL 6088	Blight Resis Savoy	New York
PI358258	Lisicki	Yugoslavia
PI374233	Baladi	Egypt

Valencia, CA) was used to extract DNA following the manufacturer's instructions. To measure genetic diversity within accessions, DNA was isolated from 12 individual seedlings from each of six germplasm accessions and two hybrids. For evaluating interaccession genetic variability, DNA was extracted from a bulk of six to 13 seedlings of each entry. DNA concentrations were determined with a DU7400 spectrophotometer

(Beckman Coulter, Fullerton, CA) and adjusted to 10–30 ng/μl for PCR reactions.

TRAP marker generation

Six fixed primers and eight arbitrary primers designed for other fingerprinting (Hu et al. 2005) or mapping (Hu 2006) projects in the laboratory were used in the current study (Table 2). All of these primers were designed with the web-based PCR primer designing program “Primer 3”, http://www.frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi (Rozen and Skaletsky 2000).

The TRAP protocol of Hu and Vick (2003) was modified as follows: (1) the molar ratio of fixed primer to arbitrary primer was decreased from 30:1 to 10:1, and (2) the annealing temperature of

the first five cycles was increased from 35 to 40°C, and for the last 30 cycles it was increased from 50 to 53°C. After the last cycle, 10 μl of the PCR products were subjected to restriction enzyme digestion to generate additional polymorphism (Zhang et al. 2005). Both digested and undigested PCR products were run in a Li-Cor Global DNA sequencer (Li-Cor Biosciences, Lincoln, NE) using a 6.5% polyacrylamide sequencing gel at 1,500 volts for 3.5 h, and the images were collected by the SAGA software.

Data analysis

In TRAP, each primer combination amplified scorable fragments ranging from 40 to 900 base pairs (bp). Only high quality polymorphic frag-

Table 2 Primer sequences and primer combinations used for TRAP analysis

Primer name	Sequences (5′–3′)	Homology by BLAST search
A: Fixed primers		
QHF6H21L	ACAGGAAAAGCCTGTAC	gi 30682785 ref NM_112219.2 <i>Arabidopsis thaliana</i> ATP binding protein-tyrosine kinase AT3G13690
Ms9	AAGAAAAGGAAGATGATTT	gi 2746338 gb AF037589.1 AF037589 An <i>Arabidopsis thaliana</i> anther-specific gene
TeloRA	CCCTAAACCCTAAACCCTAAAA	<i>Arabidopsis thaliana</i> L. telomere repeat sequence
TeloRG	CCCTAAACCCTAAACCCTAAAG	<i>Arabidopsis thaliana</i> L. telomere repeat sequence
TeloRT	CCCTAAACCCTAAACCCTAAAT	<i>Arabidopsis thaliana</i> L. telomere repeat sequence
TeloTRG	CCCAAACCCAAAACCCAAAAG	<i>Arabidopsis thaliana</i> L. telomere repeat sequence
B: Arbitrary primers		
TRAP03-700	CGTAGCGCGTCAATTATG	700 Infrared dye
TRAP04-700	CGTAGTGATCGAATTCTG	700 Infrared dye
Odd26-700	CTATCTCTCGGGACCAAAC	700 Infrared dye
Sa12-700	TTCTAGGTAATCCAACAACA	700 Infrared dye
TRAP13-800	GCGCGATGATAAATTATC	800 Infrared dye
TRAP14-800	GTCGTACGTAGAATTCT	800 Infrared dye
Odd3-800	CCAAAACCTAAAACCAGGA	800 Infrared dye
Ga5-800	GGAACCAAACACATGAAGA	800 Infrared dye
C: Primer combinations		
S01	TeloRA + Sa12-700 + Ga5-800	
S02	Ms9 + TRAP04-700 + TRAP14-800	
S03	TeloRT + TRAP03-700 + TRAP13-800	
S04	QHF6H21L + Odd26-700 + Odd3-800	
S05	TeloRG + Sa12-700 + Ga5-800	
S06	TeloTRG + Sa12-700 + Ga5-800	

ments were manually scored across all genotypes to form a binary data matrix (“1” for present, “0” for absence and “-” for missing).

Intraaccession variation was evaluated by allele frequency (%) of each polymorphic fragment. Since TRAP markers are dominant in nature, the allele frequency for each fragment was calculated with the formula $p = 1 - \text{SQRT}[1 - (\text{Number of seedlings with the fragment} / \text{Total number of seedlings tested})]$, assuming Hardy–Weinberg equilibrium.

For variation among the accessions, the data were analyzed with the computer software package NTSYSpc, Numerical Taxonomy and Multivariate Analysis System, version 2.11 (Rohlf 2002). A pairwise genetic similarity coefficient matrix was generated using the SIMQUAL procedure with the formula, $2a/(2a + b + c)$, as proposed by Dice (1945) and Nei and Li (1979). For each pair of accessions, a is the number of fragments present in both accessions, b and c are the numbers of fragments that are present in either accession, respectively. The genetic similarity coefficient matrix was then used to construct the dendrogram with the UPGMA (unweighted pair-group method using arithmetic averages) algorithm, employing the SAHN (Sequential, Agglomerative, Hierarchical, and Nested) clustering procedure.

Results

TRAP patterns in spinach

The modified TRAP protocol produced satisfactory results with all the DNA samples. The

amplified fragments ranged from 40 to 900 base pairs (bp) and the number of scorable fragments amplified by each primer combination ranged from 28 to 57. A specific primer combination amplified the same number of fragments from DNA samples isolated from either single or multiple seedlings of a given entry (Fig. 1).

High level of genetic variability within and among accessions

Although spinach displays a range of monoecious types, it generally is considered to be dioecious (Janick and Stevenson 1954). Thus, genetic variation within each accession is expected. Twelve seedlings from each of the eight entries (six accessions and two commercial hybrids) were used to assess genetic variation within accessions.

Of the 111 fragments amplified from the 96 DNA samples with two primer combinations (TeloRA + Sa12-700 and TeloRA + Ga5-800), 35 were polymorphic in at least one accession. Digestion of the PCR products with restriction enzyme *Hae* III generated an additional 24 polymorphic fragments. The number of polymorphic fragments within individual accessions ranged from 37 to 41. The wide distribution of allele frequency within individual accessions suggests that there is a high level of polymorphism within accessions (Fig. 2). Since the data were generated from only 12 seedlings from each accession, comparisons of allele frequency among accessions are not feasible.

Polymorphism was reduced substantially when DNA samples from several seedlings of an accession were used in the TRAP amplification.

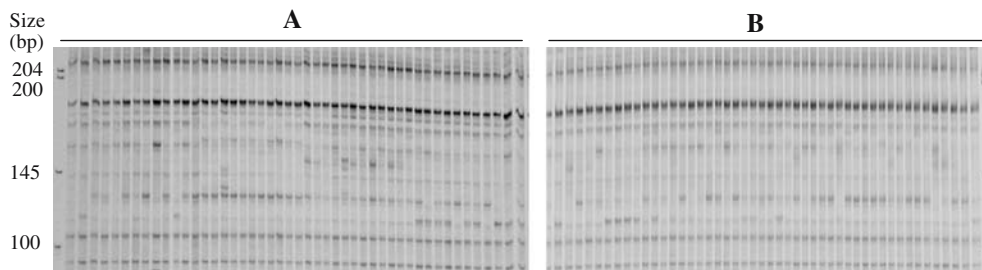
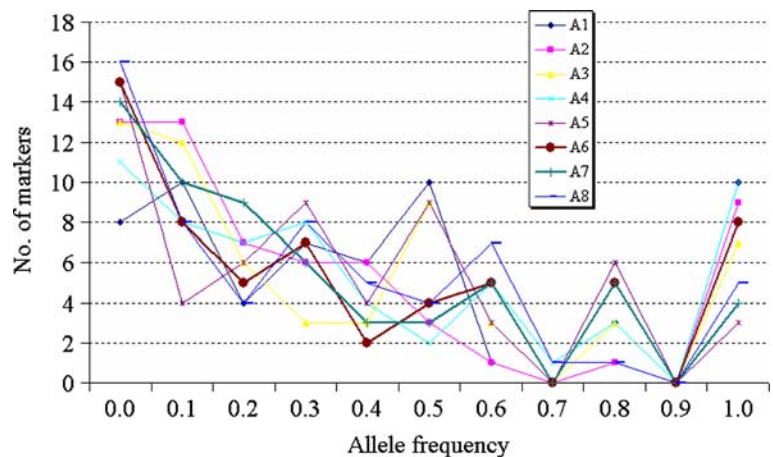


Fig. 1 TRAP fragments amplified by primer combination TeloRA + Sa12-700 from DNA from individual seedlings of four accessions (**A**) and six to 13 seedlings from each of 48 accessions (**B**)

Fig. 2 Distribution of allele frequency of 59 polymorphic fragments within eight individual accessions. A6 and A7 are two commercial F₁ hybrids



The same two primer combinations (TeloRA + Sa12-700 and TeloRA + Ga5-800) detected only 23 polymorphic fragments. Digestion of the PCR products with the same restriction enzyme revealed five additional polymorphic fragments. A total of 96 polymorphic fragments (75 were amplified products and 21 were generated by restriction enzyme digestion of the PCR products) was obtained by amplification with 12 primer combinations. Each of the 48 entries could be differentiated with these polymorphic markers.

Only 30 of the 96 (31.3%) polymorphic markers were present in the commercial hybrid CH-C, the lowest number of fragments among the 48 entries. In contrast, 60 (62.6%) markers were present in cultivar “Victoria”. Two accession-specific fragments, each only present in one accession were identified; S018213, a 213-bp fragment amplified from “Chung Pu Lao” by the primer combination of TeloRA and Ga5-800, and S067320, a 320-bp fragment amplified from “Stipski” by the primer combination of TeloTRG and SA12-700. The most frequent polymorphic marker, S068796, a 796-bp fragment amplified by the primer combination of TeloTRG and Ga5-800, was shared by 45 accessions.

The variability among the accessions at the DNA level was estimated by computing the pairwise genetic similarity coefficients. Among the 1,128 pairwise comparisons, genetic similarity coefficients based on the 96 polymorphic markers ranged from 23.2% (between “CH-J”

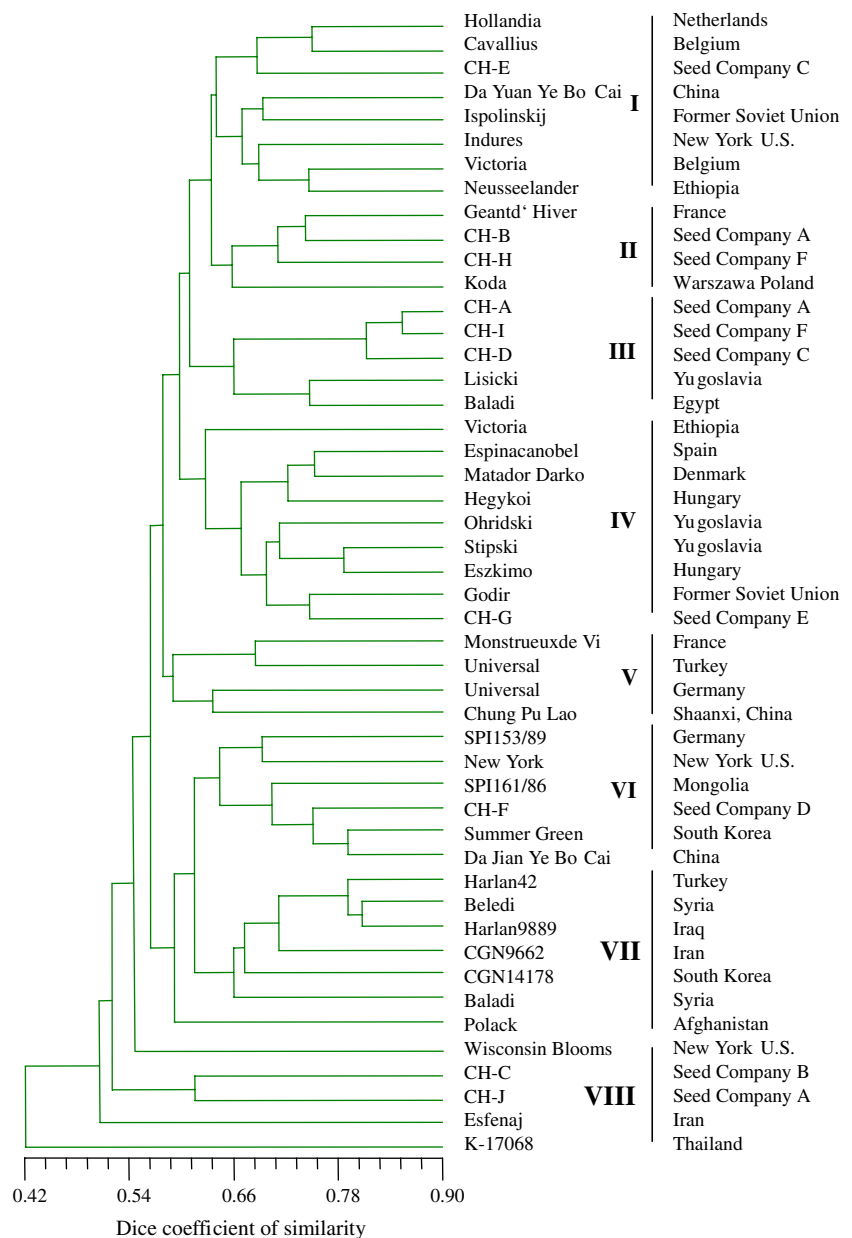
and “K-17068”) to 85.3% (between “CH-A” and “CH-I”) with an average of 57.5%. The mean value of the similarity coefficients of an accession with each of the remaining 47 accessions ranged from 42.0% for “K-17068” to 64.4% for “Victoria”. These statistics suggested substantial genetic variability among the accessions examined.

Phylogenetic relationship among accessions

Based on the similarity coefficient matrix, a dendrogram was constructed (Fig. 3). Forty-three entries formed seven clusters (I–VII) with four to nine entries each. The remaining five entries seemed more different from each other and did not form an apparent cluster (VIII). Cluster IV consisted mainly of accessions with European origin and Cluster VII was comprised of accessions primarily from the Middle East (except one from South Korea). But the remaining five clusters contained accessions from different continents. Three accessions from China (“Da Yuan Ye Bo Cai”, “Chung Pu Lao”, and “Da Jian Ye Bo Cai”) fell into separate clusters. Thus, the genetic relationships among the accessions were not highly correlated with the geographic regions in which they were collected.

The distribution of the 10 commercial hybrids in the dendrogram revealed that phenotype-based breeding activities have reduced genetic variability. The widespread variability of these hybrids suggested that breeders bred the varieties from

Fig. 3 Dendrogram showing relationships among 48 entries, cultivar names and geographic regions where accessions were collected



diverse germplasm sources. However, five of the 10 hybrids were grouped in two clusters (CH-B and CH-H in cluster II, and CH-A, CH-I and CH-D in cluster III). Although these hybrids in the same cluster were from different companies, the marker data indicated that they are genetically similar. The similarity coefficient between CH-B and CH-H is 71%, and between CH-A and CH-D is 83.3%. The similarity coefficient of 85.3% between hybrids CH-A and CH-I made them the closest pair among

the 48 entries. In addition, hybrids CH-C and CH-J have a similarity coefficient of 61.5% and formed a pair in cluster VIII (Fig. 3).

Discussion

It is a common belief that spinach was first cultivated in Iran (Persia). The earliest record of spinach in China occurred in the seventh century

and was referred to as “bo cai”, meaning “Persian vegetable”. It was introduced into China from Nepal in A. D. 647 and reached Spain around A. D. 1100. Spinach was grown in Germany in the 13th century, and by the 14th century it was common in European monastery gardens. It was brought to the Americas by the early colonists (Pandey and Kalloo 1993). By 1806, spinach had become a popular vegetable and was listed in American seed catalogs (Ryder 1979). In the current study, the 38 germplasm accessions were collected from different regions of the world. This small collection possessed considerable genetic variability. However, a close association between geographical origin and TRAP patterns was not apparent; several accessions from the same geographical origin were grouped into different clusters (Fig. 3). This could have resulted from a similar cultivar being introduced into different geographical regions, intensive selection of the introduced population, or genetic drift.

Our study revealed a relatively high level of polymorphisms at the DNA level both within and among accessions. Two factors are responsible for the observed variability. The first is the open pollination nature of the spinach breeding system. In general, spinach is dioecious (Adriance and Brison 1955), that is, individual plants bear either male (staminate) or female (pistillate) flowers, although some plants may be monoecious with varying proportions of male and female flowers on the same plant. Pollen from the male plant is carried by wind to fertilize the flowers on the female. This reproduction system maintains genetic variation. The second factor is the diverse origin of the entries in the current study. The 38 accessions were collected in Africa, Asia, Europe, and North America. It is possible that the preservation of unique alleles in different regions contributed to the genetic variation.

Among the eight entries used for studying genetic variability within accessions, two were commercial hybrids. There was not much difference in allele frequency between the hybrids and the open-pollinated accessions in regard to within-accession variability (Fig. 2). This can be explained by the following. The spinach hybrids are produced by crossing elite lines containing

both male and female plants. For hybrid seed production, the female parent is rogued prior to and during the flowering period to remove male plants (Thomas 2005). In some cases, the female parent was bred to have a high proportion of female plants. In other cases, a highly female monoecious inbred is used as the seed parent and a highly male monoecious inbred is used as the pollen parent. TRAP amplifies alleles contributed by the two parental populations in which the allele frequency is different and results in a high variability at the DNA level within the hybrid.

In summary, our results demonstrate that TRAP markers are suitable for assessing genetic diversity of spinach germplasm. One implication of this diversity is the need for a relatively large number of plants to maintain the genetic variation within accessions when regenerating spinach germplasm.

Acknowledgements The authors thank David Brenner for providing the germplasm accessions for this study and Angelia Hogness and Chia-Hao Chang for their technical assistance. We also thank Larry Campbell and Larry Charlet for their comments to an earlier version of this report. This work was supported by USDA-ARS CRIS 5442-21000-027-00D and USDA-ARS CRIS 5305-21000-011-00.

References

- Adriance GW, Brison FR (1955) Inflorescences, flowers, fruits, and seed. In: Adriance GW, Brison FR (eds) Propagation of horticultural plants. McGraw-Hill Book Company, Inc., NY Toronto London, pp 26–45
- Alwala S, Suman AJ, Arro A, Veremis JC, Kimbeng CA (2006) Target region amplification polymorphism (TRAP) for assessing genetic diversity in sugarcane germplasm collections. *Crop Sci* 46:448–455
- Dice LR (1945) Measures of the amount of ecologic association between species. *Ecology* 26:297–302
- FAO (Food and Agriculture Organization of the United Nations) (2005) FAOSTAT data. <http://www.fao-stat.fao.org/faostat/>. Accessed March 31, 2006
- Hu J (2006) Defining the sunflower (*Helianthus annuus* L.) linkage group ends with the *Arabidopsis*-type telomere sequence repeat-derived markers. *Chromosome Res* 14:535–548
- Hu J, Vick BA (2003) Target region amplification polymorphism: a novel marker technique for plant genotyping. *Plant Mol Biol Repr* 21:289–294
- Hu J, Ochoa OE, Truco MJ, Vick BA (2005) Application of the TRAP technique to lettuce (*Lactuca sativa* L.) genotyping. *Euphytica* 144:225–235

- Janick J, Stevenson EC (1954) Genetics of the monoecious character in spinach. *Genetics* 40:429–437
- Le Strange M, Koike S, Valencia J, Chaney W (2003) Spinach production in California. Publication 7212. Vegetable Research and Information Center, Division of Agriculture and Natural Resources, University of California, Davis. <http://www.anrcatalog.ucdavis.edu/pdf/7212.pdf>. Accessed March 21, 2006
- Lucier G, Plummer C (2004) Vegetable consumption expected to rise in 2004. Vegetables and melons outlook, Electronic Outlook Report from the Economic Research Service, United States Department of Agriculture VGS-302 April 21, 2004. <http://www.ers.usda.gov/publications/vgs/Apr04/vgs302.pdf>. Accessed March 25, 2006
- Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc Natl Acad Sci USA* 76:5269–5273
- Pandey SC, Kalloo G (1993) Spinach. In: Kalloo G, Bergh BO (eds) Genetic improvement of vegetable crops. Pergamon Press, Oxford New York, pp 325–338
- Rohlf FJ (2002) NTSYSpc: numerical taxonomy system, version 2.1. Exeter Publishing, Ltd., Setauket NY
- Rozen S, Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) Bioinformatics methods and protocols: methods in molecular biology. Humana Press, Totowa NJ, pp 365–386
- Ryder EJ (1979) Leafy salad vegetables. Avi Publishing Company, Inc., Westport CT
- Thomas JM (2005) Crop profile for spinach seed in Washington. <http://www.ipmcenters.org/cropprofiles/docs/waspinachseed.html>. Accessed March 21, 2006
- Zhang J, Lu Y, Yu S (2005) Cleaved AFLP (cAFLP), a modified amplified fragment length polymorphism analysis for cotton. *Theor Appl Genet* 111:1385–1395