Theor Appl Genet (1999) 99:405–418

D. Sicard · S.-S. Woo · R. Arroyo-Garcia · O. Ochoa
D. Nguyen · A. Korol · E. Nevo · R. Michelmore

Molecular diversity at the major cluster of disease resistance genes in cultivated and wild *Lactuca* spp.

Received: 7 March 1999 / Accepted: 25 March 1999

Abstract Diversity was analyzed in wild and cultivated *Lactuca* germplasm using molecular markers derived from resistance genes of the NBS-LRR type. Three molecular markers, one microsatellite marker and two SCAR markers that amplified LRR-encoding regions, were developed from sequences of resistance gene homologs at the main resistance gene cluster in lettuce. Variation for these markers were assessed in germplasm including accessions of cultivated lettuce, *Lactuca sativa L.* and three wild *Lactuca* spp., *L. serriola L.*, *L. saligna* and *L. virosa L.* Diversity was also studied within and between natural populations of *L. serriola* from Israel and California; the former is close to the center of diversity for *Lactuca* spp. while the latter is an area of more recent colonization. Large numbers of haplotypes were detected indicating the presence of numerous resistance genes in wild species. The diversity in haplotypes provided evidence for gene duplication and unequal crossing-over during the evolution of this cluster of resistance genes. However, there was no evidence for duplications and deletions within the LRR-encoding regions studied. The three markers were highly correlated with resistance phenotypes in *L. sativa.* They were able to discriminate between accessions that had previously been shown to be resistant to all known isolates of *Bremia lactucae.* Therefore, these markers will be highly informative for the establishment of core collections and marker-aided selection. A hierarchical analysis of the population structure of *L. serriola* showed that countries, as well as locations, were significantly differentiated. These differences may reflect local founder effects and/or divergent selection.

Key words *Lactuca* · Resistance genes · LRR multigene family · Diversity · Microsatellite

Communicated by H.F. Linskens

Major contributions were made by D. Sicard, S.-S. Woo and R. Arroyo-Garcia who should be considered as joint first authors

D. Sicard · S.-S. Woo · R. Arroyo-Garcia · O. Ochoa
D. Nguyen · R.W. Michelmore (✉)
Department of Vegetable Crops, University of California, Davis, CA 95616, USA
e-mail: rwmichelmore@ucdavis.edu
Fax: +1 530 752 9659

A. Korol · E. Nevo
Institute of Evolution, University of Haifa, Haifa 31905, Israel

R. Arroyo-Garcia
Departamento de Mejora Genetica y Biotechnologia, CIT-INIA, Ctra, De la Coruna, Km7, Madrid 28040, Spain

Introduction

One of the major challenges in analyzing wild populations as sources of germplasm is in measuring relevant population diversity (Hawkes 1991). An important use of germplasm in crop improvement is as a source of disease resistance (Leppik 1970; Nevo et al. 1985; Dale 1991; Lenne and Wood 1991). However, measuring diversity in the genetic basis of resistance has until recently been difficult. Typically, diversity of resistance in natural populations or germplasm collections has been analyzed by assessing resistance phenotypes (Jana and Nevo 1991; Burdon 1996, 1997). However, the ability to distinguish between different resistances has been heavily dependent on the range of isolates used to assess resistance. Also, it has been impossible to distinguish between sources of resistance that are effective against all known isolates of a particular pathogen. The recent cloning of resistance genes provides the opportunity to assay variation specifically at resistance loci using molecular markers derived from sequences of the resistance genes.

Classical and molecular genetics have increasingly demonstrated that the resistance genes in diverse plant species are clustered in the genome either as genetically separable loci or as an apparent multiallelic series (Pryor and Ellis 1993; Hulbert 1997; Michelmore and Meyers 1998). On the basis of this clustered distribution, and by interference from other cell-cell recognition systems, resistance genes have been hypothesized to be functionally and evolutionary related (Michelmore et al. 1987; Pryor

1987). Cloned resistance genes to diverse pathogens from a variety of species share common sequence motifs indicative of gene products involved in signal reception and transduction (reviewed in Hammond-Kossack and Jones 1997). The most prevalent class of resistance genes encode a nucleotide-binding site (NBS) and a leucine-rich repeat (LRR) region. Little is known about mechanisms generating variation at resistance loci. Recombination, duplication, divergent selection and transposition have all been implicated in the evolution of new resistance genes (Ellis et al. 1995; Hulbert 1997; Song et al. 1997; Meyers et al. 1998a, b). However, these studies have been based on the analysis of only one or a few haplotypes, usually in experimental or cultivated genotypes. The relative importance of these mechanisms in generating variation in natural populations has not been assessed.

Natural populations of *Lactuca* are widely represented around the world, with at least 100 species described in the genus (Lindqvist 1960a). The high levels of diversity of *Lactuca* species in the Mediterranean area and Southwest Asia indicates that this area may be a center of diversity for *Lactuca* spp. (Zohary 1991; De Vries 1997). Cultivated lettuce, *Lactuca sativa*, is part of a reproductively isolated group that includes the wild species *Lactuca serriola*, *Lactuca saligna*, *Lactuca virosa* (Lindqvist 1960a, b; Whitaker 1969; Ryder 1986). Other species have been described in this group, but they are less widespread and their status as distinct species is unclear (Ferakova 1977; Zohary 1991). *L. sativa* is fully interfertile with *L. serriola*, only partly cross-fertile with *L. saligna* and almost completely infertile with *L. virosa* (Zohary 1991). Cytogenetic and molecular marker data have confirmed the close taxonomic affinity between *L. sativa* and *L. serriola* and their more distant relationship to *L. saligna* and *L. virosa* (Lindqvist 1960b; Kesseli and Michelmore 1986; Cole et al. 1991; Kesseli et al. 1991; Hill et al. 1996; Witsenboer et al. 1998). *L. serriola* and, more rarely, *L. saligna* have been used as sources of resistance genes for introgression into *L. sativa* (Ochoa et al. 1987; Crute 1990). *L. serriola* is a common colonizer of disturbed habitats and has the most widespread global distribution of the *Lactuca* spp. (Ferakova 1977, Zohary 1991); extensive populations can be found on all continents; however, the genetic variation among populations has not yet been studied.

The genetic and molecular bases of disease resistance in lettuce have been investigated with a primary focus on downy mildew caused by the fungus *Bremia lactucae*. Parallel genetic studies on host and pathogen demonstrated that at least 15 dominant genes for resistance (*Dm* genes) in lettuce were matched by avirulence genes in *B. lactucae* in a gene-for-gene interaction (Flor 1956; Crute and Johnson 1976; Hulbert and Michelmore 1985; Farrara et al. 1987). Many other resistant accessions have been identified but few have been characterized genetically (e.g. Farrara and Michelmore 1987; Bonnier et al. 1992, 1994). The *Dm* genes characterized so far are clustered in four linkage groups along with resistance to other pathogens (Kesseli et al. 1993; Maisonneuve et al. 1994; Robbins et al. 1994; Witsenboer et al. 1995). The major cluster determines at least 11 *Dm* specificities including *Dm3* as well as resistance to root aphid. This cluster has been saturated with molecular markers using several approaches (Michelmore et al. 1991; Paran and Michelmore 1993). Recently, resistance-gene candidates (RGCs) encoding NBS and LRR motifs have been identified using PCR with degenerate oligonucleotide primers designed from sequences conserved between resistance genes cloned from other species (Shen et al. 1998). One family of over 24 members, RGC2, is localized in the major cluster of resistance genes and contains the *Dm3* gene (Meyers et al. 1998a; K. Shen et al., unpublished).

In the present study, we used molecular markers to analyze diversity at the major resistance gene cluster in *Lactuca* spp. Molecular markers were developed from the sequences of RGC2 members. Variation was analyzed in a broad collection of *L. sativa* and the three wild relatives, *L. serriola*, *L. saligna* and *L. virosa*, that had previously been characterized for resistance to downy mildew. In addition, a detailed analysis of diversity was conducted within and between wild populations of *L. serriola* from two climatically similar regions, Israel and California, the former being close to the center of diversity for *Lactuca* and the latter an area of more recent colonization (Zohary 1991). This analysis provides the first molecular data on the level and distribution of resistance gene variation within and between natural populations. The large number of haplotypes detected in wild species provides a basis for decisions on the conservation and exploitation of *Lactuca* germplasm as well as the experimental basis for future studies on the evolution of disease resistance genes.

Materials and methods

Plant material

Two sets of materials were examined. The first set included 74 accessions of the cultivated species, *L. sativa*, and 74 accessions of wild *Lactuca* species. The cultivated accessions were selected to represent diversity within the species and had been previously studied with a range of molecular markers (Kesseli et al. 1991; Hill et al. 1996). The cultivated samples included the diversity for downy mildew resistance genes (*Dm* genes; Farrara et al. 1987) as well as genotypes of importance to US agriculture. The wild *Lactuca* spp. comprised a total of 16 accessions of *L. serriola*, 47 accessions of *L. saligna*, eight accessions of *L. virosa*. These wild accessions had previously been shown to be resistant to all isolates of *B. lactucae* tested up to 1996 (O. Ochoa, unpublished). Single accessions of *Lactuca augustana*, *Lactuca indica* and *Lactuca perennis*, *Cichorium endivia* (endive) and *Helianthus annuus* (sunflower), all members of the *Compositae* family, were included as outgroups. Cultivars and wild accessions were obtained from germplasm collections at the Plant Introduction Center. Pulman, Wash., USA, the Centre for Genetic Resources, CPRO-DLO, P.O. Box, 16, 6700 AA Wageningen, The Netherlands, and the Department of Vegetable Crops, UC Davis, USA.

The second set of genotypes was composed of 505 samples of *L. serriola* collected as individual plants from two regions, Israel and California (Fig. 1). In each region, collections were made

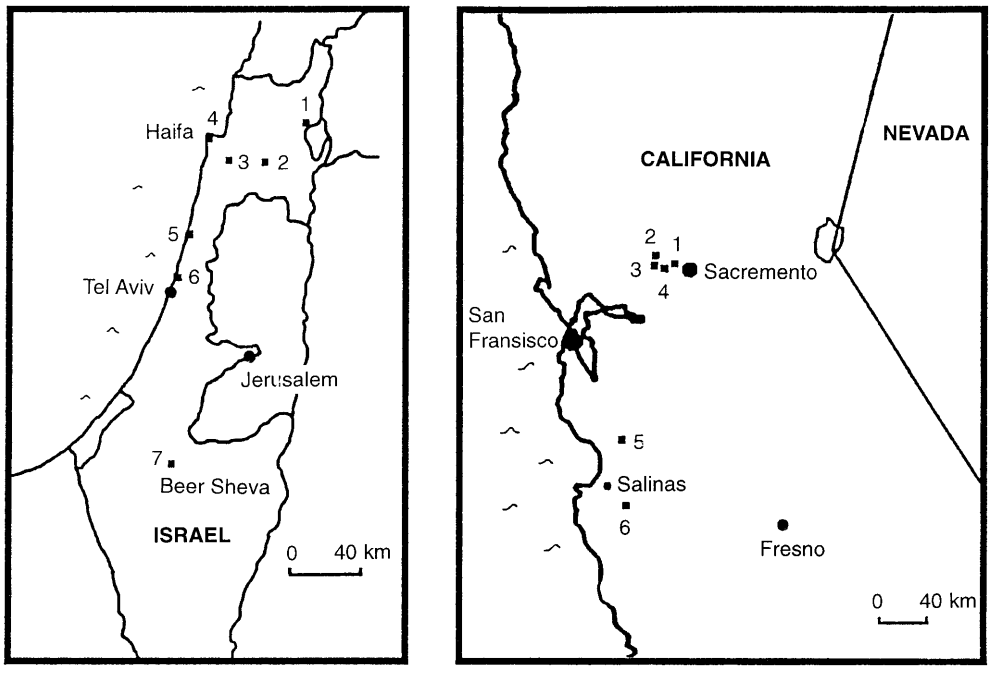

Fig. 1 Origin of natural populations of *L. serriola* collected in Israel and California. Israeli populations. *1*: Tiberias, *2*: Nazareth, *3*: Tel Hanan, *4*: Haifa, *5*: Netanya, *6*: Shefayim, *7*: Beer Sheva. Californian populations: *1*: Yolo Landfill, *2*: Davis CR32, *3*: Davis Putah Creek, *4*: Davis CR29, *5*: Gilroy, *6*: Salinas

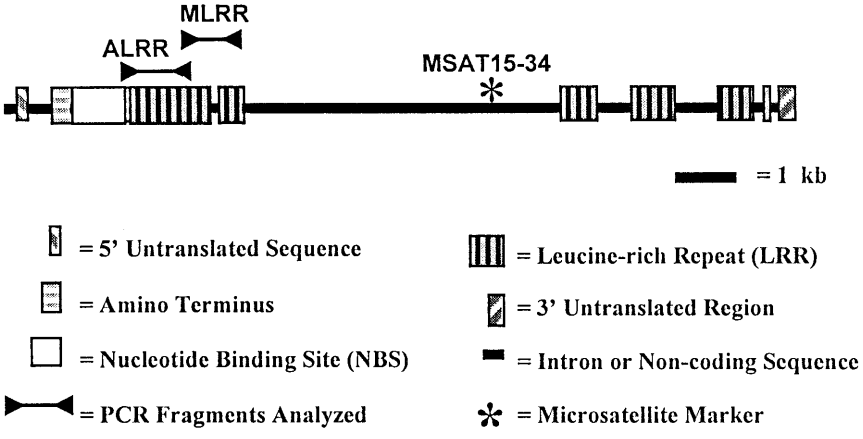

Fig. 2 Position of markers in RGC2 genes

Table 1 Oligonucleotide primers used for PCR-amplification of markers

Markers	Primers	Sequence
MSAT15-34	MSAT15-3	(5′GTATCACATCCCAAACTCTC3′)
	MSAT15-4	(5′GACAACAAAGTTGAACTGCC3′)
ALRR	RLG3R2	(5′GAACGCTCTGCCATCTCATTG3′)
	RLG3F2	(5′GAGAAGCAAGAACCAGGCTCA3′)
MLRR	3RACE3C	(5′GCAAACACTTTGTCAAGACTTGAG5′)
	KSF2	(5′GCACCGACACAATCCAAG3′)

from multiple locations representing different environments. Seeds were harvested from single plants, separated by at least 1 m. In Israel, 196 samples were collected from seven sites in July and August of 1994 and 1995. In California, 309 samples from six sites were obtained in September 1995.

For all accessions, approximately, ten plants were grown for each genotype and their leaves pooled for DNA extraction using a modified CTAB procedure (Bernatzky and Tanksley 1986).

Microsatellite-marker analysis

A microsatellite marker, MSAT15-34, is located in an intron involving at least five members of the *RGC2* gene family in the cultivar Diana (Fig. 2; Okubara et al. 1997; Meyers et al. 1998a). This compound microsatellite sequence contains a complex combination of di- and tetranucleotide repeats; for example in copies RGC2B and RGC2C; the microsatellite sequences are $(CG)_3(CA)_9T(AG)_{15}(AAAG)_3(AG)_5$ and $(CG)_3(CA)_8TAA(AG)_{20}$. The microsatellite marker was amplified using primers MSAT15-3 and MSAT15-4 (Table 1). The primer MSAT15-3 was end-labelled by phosphorylation with γP^{32} ATP us-

ing T4 polynucleotide kinase (New England Biolabs, Beverly, Mass.). The PCR was performed in 20 µl containing 50 ng of template DNA, 200 µM of each primer, 100 µM of dNTP, 1 U of *Taq*-polymerase, and 2 µl of 10× PCR buffer (100 mM Tris-HCl, 500 mM KCl, 15 mM $MgCl_2$). Amplification was conducted in a Perkin-Elmer 9600 thermocycler with 3 min at 94° C initially, followed by 30 cycles of 30 s at 94°C, 30 s at 60° C and 1 min at 72° C, and a final extension step at 72° C for 5 min. Amplification products were separated in a denaturing 8% polyacrylamide gel.

RGC2 SCARs of LRR encoding regions

Primers were used to amplify one of two regions encoding LRRs from multiple *RGC2* family members (Table 1; Fig. 2). The primers, RLG3R2 and RLG3F2, amplify from sequences encoding the anterior region of the LRR (designated ALRR). The primers, 3RACE3C and KSF2, amplify sequences encoding a region in the middle of the LRR region and include a small intron (designated MLRR). PCR amplification was made in 25 µl containing 50 ng of template DNA, 250 µM of each primer, 100 µM of dNTP, 1 U of *Taq*-polymerase and 2.5 µl of 10× PCR buffer (100 mM Tris-HCl, 500 mM KCl, 15 mM $MgCl_2$). The amplification was conducted as described for MSAT15-34. Amplified products were visualized in a 2% agarose gel stained with ethidium bromide. Ten microliters of amplified product were digested in a total volume of 12 µl containing 0.8 U of *Tsp*509I (New England Biolabs, Beveryl, Mass.) and 1× NEB buffer (10 mM Bis Propane-HCl, 10 mM $MgCl_2$, 1 mM DTT) for 3 h at 65° C. The products were separated in a 8% non-denaturing polyacrylamide gel and visualized with Syber green (FMC Bioproducts, Rockland, Me.).

Data analysis

MSAT15-34 is duplicated in the cluster and amplifies from multiple members of the *RGC2* family (Meyers et al. 1998a). SCAR markers, ALRR and MLRR, were also duplicated in the genome. Genotypes were therefore considered as haplotypes because it was not possible to distinguish between alleles of the same gene and different paralogs. MSAT15-34 haplotypes were designated using a two-letter code. Ax: first identified in *L. sativa*; Bx first identified in wild *L. serriola* from Isareal; Cx: first identified in *L. serriola* from California; Dx and Ex: first identified from wild *Lactuca* germplasm. In the case of the same haplotype being detected in multiple species, the designation that was initially assigned was used throughout. SCAR haplotypes were designated using the name of the marker (ALRR or MLRR) and a number assigned randomly. Haplotype diversity was estimated using the Shannon diversity index (Peet 1974). The Shannon diversity indices were compared using a *t*-test (Hutcheson 1970).

The partitioning of the diversity among individuals within location, between locations within a country, and between countries was estimated using the AMOVA function of Arlequin computer software packade (Schneider et al. 1997). Signifance levels for estimates of variance components were computed by non-parametric permutation.

Results

Polymorphism for the microsatellite MSAT15-34 among cultivated and wild accessions of *Lactuca* spp.

MSAT15-34 was assayed on a total of 148 genotypes from seven *Lactuca* spp. to determine the level of polymorphism at the major cluster of resistance genes. The primers allowed successful amplification from all cultivated and wild germplasm accessions except for one sample of *L. saligna* and one of *L. virosa.* This pair of primers was also sufficiently conserved to allow amplification from endive and sunflower. We analyzed the amplification products as combinations of array sizes rather than alleles for two reasons. The *RGC2* family is duplicated to varying extents in different genotypes, therefore the allelic relationships between orthologs could not be determined. In addition, although these species are au-

Fig. 3 Polymorphism detected by microsatellite marker MSAT15-34 in germplasm of *Lactuca* spp.

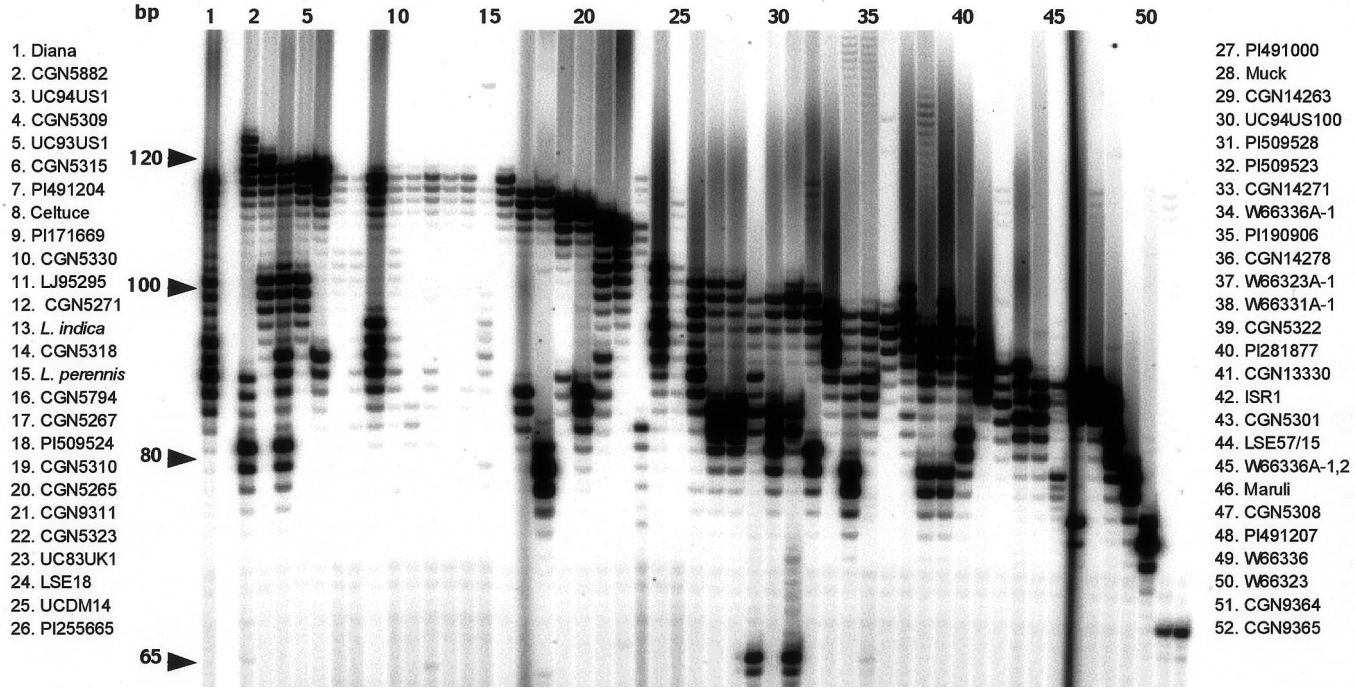

Table 2 Microsatellite array sizes of each haplotype detected by MSAT15-34 in *Lactuca* spp., sunflower, endive and natural populations of *L. serriola*

Haplo[a]	Array sizes[b]	Haplo[a]	Array sizes[b]
AA	89, 93, 101, 105, 117	**DF**	105, 117
AB	91, 99, 105, 117	**DG**	83, 89, 93, 99, 117
AC	77, 83, 87, 93, 99, 103, 117	**DH**	91, 95, 117
AD	89, 93, 99, 117	**DI**	87, 91, 117
AE	91, 115	**DJ**	89, 103, 117
AF	91, 113	**DK**	89, 117
AG	93, 101, 103	**DL**	87, 97, 101, 117
AH	101	**DM**	87, 117
AI	93, 95, 99	**DN**	77, 79, 117
AJ	95, 97	**DO**	117
AK	83, 89, 93, 95	**DP**	87, 115
AL	89, 93, 95	**DQ**	79, 115
AM	83, 93	**DR**	89, 113
AN	93	**DS**	87, 113
AO	87	**DT**	87, 91, 101, 103, 111
AP	85	**DU**	101, 103, 109
AQ	83	**DV**	79, 83, 109
AR	71	**DW**	95, 101, 103
BA	103	**DX**	95, 101, 103, 113
BB	83, 101	**DY**	89, 91, 97, 101
BC	83, 99	**DZ**	83, 85, 101
BD	99	**EA**	85, 101
BE	81, 85, 89, 97	**EB**	85, 89, 99
BF	83, 97	**EC**	81, 85, 99
BG	79, 83, 97	**ED**	85, 99
BH	89, 95	**EE**	81, 99
BI	83, 95	**EF**	93, 97
CA	79, 83, 89, 93, 99, 119	**EG**	79, 89, 97
CB	93, 105	**EH**	89, 97
CC	85, 99, 105	**EI**	97
CD	93, 103	**EJ**	85, 97
CE	93, 101	**EK**	79, 89, 93, 95
CF	93, 99	**EL**	79, 93, 95
CG	79, 93, 99	**EM**	83, 91, 95
CI	79, 97	**EN**	89, 91
CJ	77, 97	**EO**	87, 91
CK	79, 83, 95	**EP**	85, 91
CL	79, 95	**EQ**	87, 89
CM	77, 95	**ER**	79, 89
CN	95	**ES**	73, 89
CO	79, 85, 93	**ET**	89
CP	85, 93	**EU**	83, 87
CQ	79, 93	**EV**	79
DA	81, 87, 89, 119, 123, 125	**EW**	71, 77
DB	101, 119, 121	**EX**	66, 99, 105, 117
DC	81, 91, 101, 119	**EY**	66, 91, 101
DD	101, 119	**FA**	83, 99, 103, 117
DE	91, 119		

[a] Haplo=haplotype. See Materials and methods for origins of haplotype designations
[b] Size given in base pairs

togamous there was a low level of heterozygosity in some wild accessions; therefore it was not possible to distinguish whether arrays of different sizes originated from different alleles of the same gene or from different paralogs.

Both the size and number of fragments amplified with MSAT15-34 primers were highly variable. Twenty seven different fragment sizes ranging from 66 to 125 bp were observed; these formed a continuous ladder differing by 2 bp except that 75- and 107-bp fragments were not observed and the smallest product differed by 5 bp (Fig. 3, Table 2). The complexity of the haplotype (the number of arrays per haplotype) varied from one to seven products (Fig. 4A). A single product was amplified from 42% and two products from 27% of the accessions. More than two products were detected in 31% of the accessions. Accessions possessing a single microsatellite array were the most frequent in all the species except *L. saligna* where the majority of accessions contained two arrays. The most complex pattern of seven amplification products was observed in the cultivar BGBH9A. These species of *Lactuca* are diploid suggesting that MSAT15-34 has been duplicated to varying extents between genotypes. This is consistent with the duplicated nature of the major cluster of resistance genes (Meyers et al. 1998a) and the evolution of this cluster by a birth-and-death process (Michelmore and Meyers 1998).

Twenty four haplotypes were detected in 74 accessions of the cultivated species, *L. sativa.* The frequencies of each haplotype reflect the sampling of genotypes to represent the diversity of resistance genes at the major

Table 3 Origins and diversity of *Lactuca* spp.

A. Cultivated species

Name	Type[a]	Origin	*Dm* or R genes[b]	MSAT 15-34[c]	ALRR[d]	MLRR[d]	Name	Type[a]	Origin	*Dm* or R genes[b]	MSAT 15-34[c]	ALRR[d]	MLRR[d]
Diana	BH	UK	**1**, **3**, 7, 5/8	AA	ALRR 1	MLRR 1	Alpha	I	USA	**1**, 7, 5/8	AQ	ALRR 15	MLRR 15
Mildura	BH	UK	**1**, **3**	AA	ALRR 1	MLRR 1	Autumn Gold	1	USA		AQ	ALRR 15	MLRR 15
BGBH 9B	BH	FRA	4, 5/8, 9, 10, 13, **14**	AB	ALRR 2	MLRR 2	Calmar	I	USA	7, 5/8, 13	AQ	ALRR 15	MLRR 15
BGBH 9A	BH	FRA	4, 5/8, 10, 13, **14**	AC	ALRR 3	MLRR 2	Climax	I	USA		AQ	ALRR 15	MLRR 15
Dandie	BH	NLD	**3**	AD	ALRR 1	MLRR 1	El Toro	I	USA	10	AQ	ALRR 15	MLRR 15
Kordaat	BH	NLD	**1**, **3**, 4	AD	ALRR 1	MLRR 1	Empire	I	USA	13	AQ	ALRR 15	MLRR 15
Solito	BH	NLD	**1**, **3**, 7	AD	ALRR 1	MLRR 1	Marquette	I	USA		AQ	ALRR 15	MLRR 15
Divalo	B	DEU		AD	ALRR 1	MLRR 1	Merit	I	USA		AQ	ALRR 15	MLRR 15
Edgar	BH	NLD	**2**, **3**, 7	AD	ALRR 1	MLRR 1	Mesa659	I	USA	7, 13	AQ	ALRR 15	MLRR 15
Bullseye	I	USA	**1**, **3**, 5/8	AD	ALRR 4	MLRR 1	Pennlake	I	USA	13	AQ	ALRR 15	MLRR 15
Lobjoit Cos	C	NLD	No *Dm*	AE	ALRR 6	MLRR 3	Salinas 88	I	USA	7, 5/8,13	AQ	ALRR 15	MLRR 15
UCDM2	GS	USA	**2**	AF	ALRR6	MLRR3	Sea Green	I	USA	10	AQ	ALRR 15	MLRR 15
Portato	BH	NLD	**1**, **2**, 7	AF	ALRR 6	MLRR 3	Target	I	USA	5/8, 11, 13	AQ	ALRR 15	MLRR 15
Amplus	BH	NLD	**2**, 4	AF	ALRR 6	MLRR 6	Valverde	I	USA	5/8	AQ	ALRR 15	MLRR 15
Fila	BH	NLD	**2**, 11	AF	ALRR 6	MLRR 6	Vampire	I	USA		AQ	ALRR 15	MLRR 15
Liba	BH	NLD	**1**, **2**	AF	ALRR 6	MLRR 6	Vanguard	I	USA	7, 10, 13	AQ	ALRR 15	MLRR 15
MayKing	BH	NLD	**2**, 4	AF	ALRR 6	MLRR 6	Vanguard75	I	USA	7, 10, 13	AQ	ALRR 15	MLRR 15
Saffier	BH	NLD	1, 7, **16**	AG	ALRR 7	MLRR 7	Winterhaven	I	USA	7, 10, 13	AQ	ALRR 15	MLRR 15
Kinemontepas	BH	FRA	10, 13, **16**	AH	ALRR 8	MLRR 8	Premier	I	USA		AQ	ALRR 15	MLRR 15
Lednicky	BH	CZCH	**1**, **14**	AI	ALRR 6	MLRR 4	Salad Crisp	L	USA		AQ	ALRR 15	MLRR 15
Sucrine	LA	FRA	**1**, 5/8, 10	AJ	ALRR 6	MLRR 17	Gallega	LA	ESP		AQ	ALRR 15	MLRR 15
Ithaca	I	USA		AK	ALRR 9	MLRR 18	UCDM10	GS	USA	10	AQ	ALRR 15	MLRR 15
G. Winterkönig	BH	NLD	4, 13, **14**	AL	ALRR 10	MLRR 10	Valmaine	C	USA	5/8	AQ	ALRR 15	MLRR 15
Salad Bowl	L	USA		AL	ALRR 9	MLRR 63	Iceberg	B	USA		AQ	ALRR 16	MLRR 15
Grand Rapids	L	USA		AL	ALRR 9	MLRR 63	Mariska	BH	NLD	**18**	AQ	ALRR 17	MLRR 12
Fulton	I	USA		AL	ALRR 9	MLRR 9	Salinas	I	USA	7, 5/8, 13	AQ	ALRR 18	MLRR 15
Green Lakes	I	USA		AL	ALRR 9	MLRR 9	Ardente	BH	NLD	**1**, **6**, 7	AR	ALRR 20	MLRR 15
Montello	I	USA		AL	ALRR9	MLRR9	Avondefiance	BH	UK	5/8, **6**	AR	ALRR 20	MLRR 15
Oswego	I	USA	7	AL	ALRR 9	MLRR 9	Sabine	BH	NLD	**6**	AR	ALRR 20	MLRR 15
UK60	I	USA		AM	w	No	Avoncrisp	B	UK	5/8, **6,7**	AR	ALRR 20	MLRR 15
Passion Blonde	BH	FRA		AN	ALRR 11	MLRR 11	Lakeland	I	UK		AR	ALRR 20	MLRR 15
Capitan	BH	NLD	11	AO	ALRR 12	MLRR 12	Celtuce	ST	USA		DG	ALRR 24	MLRR 22
Cobham Green	BH	UK	No *Dm*	AO	ALRR 12	MLRR 12	PI171669	L	TUR		DH	ALRR 21	MLRR 21
Hilde	BH	–	12	AO	ALRR 12	MLRR 12	Pirat	L	DEU		DO	ALRR 26	MLRR 27
Mondian	BH	NLD	11	AO	ALRR 12	MLRR 12	UCDM14	GS	USA	**14**	DX	ALRR 24	MLRR 49
R4T47	BH	UK	4	AO	ALRR 12	MLRR 12	Muck	B	DEU		EA	ALRR 20	MLRR 50
Great Lakes	I	USA	7	AP	ALRR 15	MLRR 15	Maruli	L	GRC		ES	ALRR 21	MLRR 60

[a] B=Batavia, BH=Butterhead, C=Cos, I=Iceberg, L=Leaf, LA=Latin, ST=Stem, GS=Genetic stocks

[b] *Dm*/R: resistance to *B. lactucae* (Resistance is designated as an R-factor until a single *Dm* gene is demonstrated). No *Dm*=no known *Dm* genes. *Dm* genes located at the major cluster of resistance genes (*Dm1*, *Dm2*, *Dm3*, *Dm6*, *Dm14*, *Dm15*, *Dm16*, *Dm18*) are designed in bold. Ref: Farrara et al. 1987; Crute 1992; Witsenboer et al. 1995

[c] Haplotypes are designed using a two-letter code. Ax: first identified in *L. sativa*, Dx and Ex: first identified in wild *Lactuca* germplasm. w=Weak amplification of the marker, No=no amplification of the marker

[d] Haplotypes are designed using the name of the SCAR marker (ALRR or MLRR) and numbers assigned randomly. w=weak amplification of the marker, No=no amplification of the marker

B. Wild *Lactuca* spp.

Name	Origin	MSAT 15-34[a]	ALRR[b]	MLRR[b]	Name	Origin	MSAT 15-34[a]	ALRR[b]	MLRR[b]	Name	Origin	MSAT 15-34[a]	ALRR[b]	MLRR[b]
Lactuca saligna														
CGN5314	ISR	AO	ALRR 13	MLRR 49	CGN5330	ISR	DI	ALRR 25	MLRR 25	PI509523	GRC	EE	ALRR 6	MLRR 54
CGN5157	ITA	AO	ALRR 14	MLRR 13	CGN5271		DK	ALRR 24	**MLRR 27**	CGN5322	ISR	EL	ALRR 37	MLRR 37
CGN4662		AO	ALRR 14	MLRR 14	CGN5318	ISR	DM	ALRR 26	**MLRR 27**	CGN13330	TUR	EN	ALRR 39	**MLRR 39**
CGN10883	Portugal	AO	ALRR 14	MLRR 14	CGN5325	ISR	DM	ALRR 26	**MLRR 27**	CGN5317	ISR	EO	ALRR 21	MLRR 58
PI491208	GRC	AP	ALRR 11	MLRR 5	CGN10888		DO	ALRR 24	**MLRR 27**	ISR-1	ISR	EO	ALRR 21	MLRR 59
PI509525	GRC	AP	ALRR 5	MLRR 20	CGN9314	ISR	DO	ALRR 26	No	ISR-2	ISR	EO	ALRR 21	MLRR 59
CGN5882	ISR	DA	ALRR 21	MLRR 21	CGN5267		DP	ALRR 21	MLRR 14	CGN5301	FRA	EP	ALRR 21	MLRR 58
UC94US1		DB	ALRR 22	MLRR 14	PI509524	GRC	DQ	ALRR 33	MLRR 33	CGN5308	ISR	ET	ALRR 41	MLRR 41
UC94US3		DB	ALRR 22	MLRR 14	CGN5310	ISR	DR	ALRR 21	MLRR 14	CGN5895	ISR	ET	ALRR 41	MLRR 41
UC94US6		DB	ALRR 22	MLRR 14	CGN5265		DS	ALRR 21	MLRR 14	CGN9313	ISR	ET	ALRR 41	MLRR 41
UC94US10		DB	ALRR 22	MLRR 14	CGN5796		DS	ALRR 21	MLRR 14	CGN5947	ISR	ET	ALRR 42	MLRR 42
CGN5309	ISR	DC	ALRR 21	MLRR 21	PI261653	Portugal	DS	ALRR 21	MLRR 14	PI491207	GRC	EU	ALRR 33	MLRR 61
UC93US1	USA	DD	ALRR 22	MLRR 14	CGN9311	ISR	DT	ALRR 21	MLRR 21	PI491204	GRC	EU	ALRR 33	MLRR 62
UC94US12		DD	ALRR 23	MLRR 23	CGN5323	ISR	DU	ALRR 22	MLRR 46	UC94US11		EV	ALRR 43	MLRR 43
CGN5315	ISR	DE	ALRR 21	MLRR 14	PI491000	TUR	DZ	ALRR 20	MLRR 50	PI491206	GRC	No	ALRR 26	**MLRR 27**
PI491204	GRC	DF	ALRR 24	MLRR 24	UC94US100	GRC	EC	ALRR 21	MLRR 52					
Lactuca serriola														
93G252		AQ	ALRR 17	MLRR 12	PI509528	GRC	ED	ALRR 22	MLRR 53	LSE57/15	UK	EQ	ALRR 40	MLRR 40
CGN5916	ISR	AQ	ALRR 19	MLRR 16	CGN14271	HUN	EF	ALRR 34	MLRR 34	W6633A-1,2	Russia	ER	w	n
PIVT1309[c]	ISR	DO	ALRR 29	MLRR 29	W66336A-1	Russia	EG	ALRR12	MLRR55	W66336	Russia	EV	ALRR 44	MLRR 44
LSE18	CSK	DW	ALRR 24	MLRR 48	CGN14278	HUN	EI	ALRR 36	MLRR 36	W66336-1	Russia	EV	ALRR 45	MLRR 44
PI255665	AFG	DY	ALRR 21	MLRR 1	W66331A-1	Russia	EK	ALRR 12	MLRR 57					
CGN14263	HUN	EB	ALRR 22	MLRR 51	PI281877	IRK	EM	ALRR 38	MLRR 38					
Lactuca virosa														
GCN5794		DO	ALRR 30	MLRR 30	W66323A-1	ARM	EJ	ALRR21	MLRR 56	PI274375		No	ALRR 31	MLRR 32
UC83UK1	UK	DV	ALRR 26	MLRR 47	W66323	ARM	EW	ALRR 42	MLRR 19	CGN9365	IRN	ÉY	ALRR 46	MLRR 16
LJ95295	UK	DJ	ALRR 26	**MLRR 26**	CGN9364		EX	ALRR 46	MLRR 16					
Lactuca augustana														
PI190906		EH	ALRR 35	MLRR 35										
Latuca perennis		DN	ALRR28	**MLRR 28**										
Lactuca indica		DL	ALRR 27	n										
Helianthus annuus (sunflower)		DO	n	n										
Cichorium endiva		DO	n	n										

[a] Haplotypes are designed using a two-letter code. Ax: first identified in *L. sativa*, Dx and Ex: first identified in wild *Lactuca* germplasm. w=weak amplification of the marker, n=no amplification of the marker

[b] Haplotypes are designed using the name of the SCAR marker (ALRR or MLRR) and numbers assigned randomly. w=weak amplification of the marker, n=no amplification of the marker; bold haplotypes are the ones for which a second less-intense band was observed

[c] PIVT1309 is now in the CGN collection as CGN5099

cluster as well as morphological types (Table 3). Molecular-marker haplotypes were highly correlated with the *Dm* genes detected phenotypically at the major cluster but not with *Dm* genes in other clusters. The most-common haplotype, AQ, had a frequency of 0.35. Twentyone of the twenty-six AQ-containing lines are closely related US crisphead types that lack any known *Dm* in this cluster. Five haplotypes had frequencies between 0.06 and 0.09 and were detected either in closely related groups of similar plant type (AL and AO) or in lines sharing the same *Dm* gene at this cluster (AF, AD and AR). The remaining haplotypes were present in only one or two lines that represented infrequently sampled *Dm* phenotypes.

Within the major cluster (*Dm1*, *Dm2*, *Dm3*, *Dm6*, *Dm14*, *Dm15*, *Dm16*, *Dm18*; Farrara et al. 1987; Maisonneuve et al. 1994), there were correlations between array sizes and known *Dm* genes. All accessions with *Dm3* had haplotype AA (89, 93, 101, 105, 117 bp) or AD (89, 93, 99, 117 bp). The gene encoding *Dm3* is known to include the 117-bp array of MSAT15-34 (Meyers et al. 1998a; K. Shen et al., unpublished). The array size of 117 bp was diagnostic for accessions known to have *Dm3* except that BGBH9A and BGBH9B also had the 117-bp array; the resistance phenotype of these lines is complex and unclear, although they do not appear to carry *Dm3* (Farrara et al. 1987). *Dm2* was closely but not absolutely correlated with haplotype AF (91 bp, 113 bp). The 113-bp array was observed only in accessions carrying *Dm2* except for cultivar UCDM14 that has the 113-bp array but not *Dm2*; interestingly this line was derived from a line carrying *Dm2* (S. Hulbert and R. Michelmore, unpublished). Also, in the cultivar Edgar that carries *Dm2*, the presence or absence of the 113-bp array could not be determined due the strong amplification of the 117-bp array. *Dm6* was correlated with the 71-bp array (haplotype AR). Cultivar Grand Rapids that was originally thought to contain *Dm6* did not have the 71-bp array. The lack of correlation with *Dm1* which is in the major cluster reflects the genetic distance between *Dm1* and *RGC2* (10 cM between *Dm1* and *Dm3*). The genetic position of *Dm14* within the cluster is poorly defined; no correlation was observed between *RGC2*-derived markers and *Dm14*, possibly indicating that *Dm14* is not encoded by sequences closely related to MSAT15-34-carrying *RGC2* sequences (Meyers et al. 1998a). There were insufficient samples to identify correlations with *Dm15*, *Dm16* or *Dm18*.

Fifty one different haplotypes were detected in the 74 accessions of wild *Lactuca* germplasm. The majority accessions had unique haplotypes. The most common haplotype had a frequency of less than 0.08. Therefore, there was very little redundancy amongst the lines identified as being resistant to all known isolates of *B. lactucae*. In six cases, accessions with identical haplotypes had the same geographical origin; these may be redundant collections of the same genotype. Alternatively, they may carry the same resistance genes at this locus.

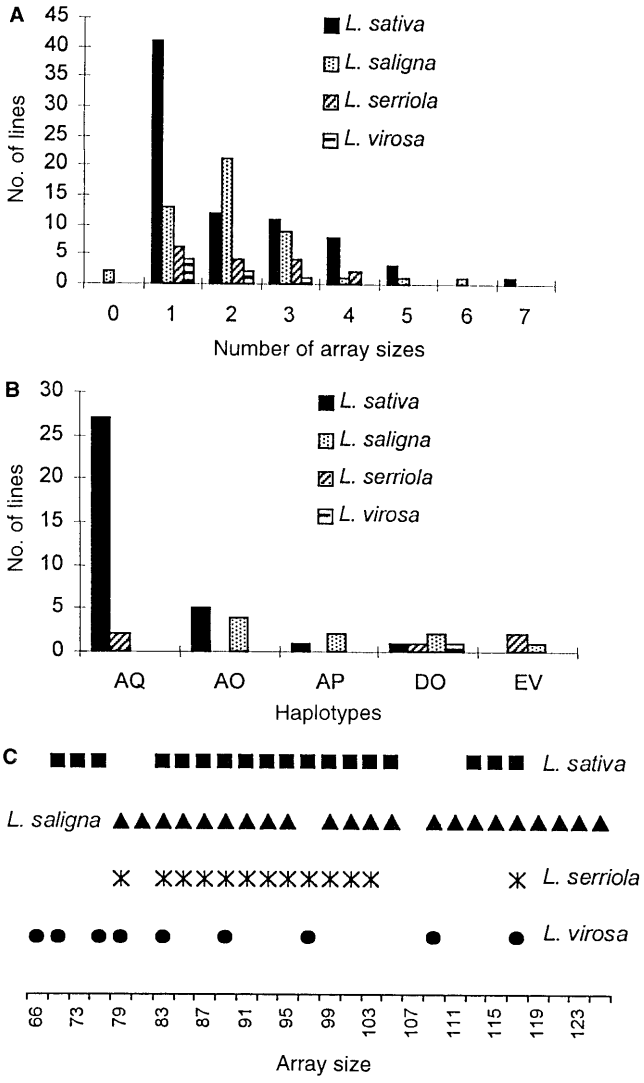

Fig. 4A, B, C Diversity of microsatellite marker MSAT15-34 among four *Lactuca* spp. **A** Complexity of haplotypes in *Lactuca* spp. **B** Prevalence of haplotypes present in more than one species. The haplotype designations are detailed in Table 3. **C** Array sizes detected in *Lactuca* spp.

The distribution of haplotypes was compared between the cultivated and wild germplasm. Most of the haplotypes were found only in one species. Only five haplotypes were present in more than one species (Fig. 4B). Four haplotypes were observed in *L. sativa* and at least one wild species. Haplotype DO was found at low frequency in all of the four well-represented *Lactuca* species as well as the single sunflower accession. AQ, the prevalent *L. sativa* haplotype, was also present in two accessions of *L. serriola.* Two other haplotypes were found in both *L. sativa* and *L. saligna.* One haplotype was detected in *L. serriola* and *L. saligna* but no other species. The lack of overlap between species indicates that the wild species represent a largely untapped resource of diversity for cultivated lettuce.

In contrast with the haplotypes, most of the array sizes were detected in multiple species (Fig. 4C). Nineteen

Table 4 Diversity and distribution of MSAT15-34 haplotypes in natural populations of *L. serriola* from California and Israel. Populations are arranged from North to South in both tables

A. Diversity of haplotypes in each population

Region	Population	N	No. of haplotypes	No. of private haplotypes	Shannon's Index[a]
California	Yolo Landfill (YOL)	53	10	2	1.852
	Davis, CR32 (D32)	49	7	1	0.982
	Davis, Putah creek (DPU)	41	8	0	1.102
	Davis, CR29 (D29)	56	8	0	1.646
	Gilroy (GIL)	47	12	6	1.904
	Salinas (SAL)	63	7	1	1.143
	Total	309	22	16	1.996
Israel	Tiberias (TIB)	4	3	0	–
	Nazareth (NAZ)	47	8	2	1.639
	Tel Hanan (TEL)	19	7	1	1.383
	Haifa (HAI)	54	9	2	1.439
	Netanya (NET)	11	4	1	1.12
	Shefayim (SHEF)	22	1	0	0
	Beer Sheva (BEER)	39	7	1	1.364
	Total	196	17	8	1.97

[a] Shannon's index: $h=\Sigma p_i \ln p_i$, where pi is the frequency of haplotype i. Shannon's index was not calculated for Tiberias because of the small population size

B Distribution of haplotypes in each population

Haplo[a]	California							**Haplo**[a]	Israel							
	Total	YOL	D32	DPU	D29	GIL	SAL		Total	TIB	NAZ	TEL	HAI	NET	SHE	BEE
EV	151	19	36	29	23	9	35	**EV**	2		2					
EI	13	1	1	1	10			**EI**	26	1	15		3			7
BD	21					1	20	**BD**	23	1	4	3	12	3		
AN	5				2		3	**AN**	6		1		4			1
ET	1						1	**ET**	2		1		1			
No	11				3	7	1	**No**	2			1		1		
CJ	24	4	4	5	9		2	**AQ**	73	2	10	11	29			21
CD	19					18	1	**AH**	31			1	2	6	22	
CI	16	3	5	1	7			**BF**	15		13	1				1
CN	16	10	1	2	1	2		**BC**	7				1			6
CL	11	9		1		1		**BI**	2			1				1
CM	5	2	1	1	1			**ED**	2							2
CQ	4	3		1				**BG**	1		1					
CF	3					3		**BH**	1			1				
CP	2					2		**BB**	1				1			
CA	1	1						**BE**	1				1			
CK	1	1						**BA**	1					1		
CC	1		1													
CB	1					1										
CE	1					1										
CG	1					1										
CO	1					1										

[a] Haplo=haplotype; haplotypes are designed using a two-letter code. Ax: first identified in *L. sativa*, Bx: first identified in wild *L. serriola* from Israel, Cx: first identified in *L. serriola* from California, Ex: first identified in wild *Lactuca* germplasm. No=no amplification of MSAT15-34

array sizes out of the 27 observed, were detected in at least two species. All of the array sizes, except one (79 bp) that was observed in *L. serriola*, were also observed in *L. sativa*. All array sizes found in *L. sativa* and *L. serriola*, except one that was unique to *L. sativa* (73 bp), were also detected in *L. saligna* or *L. virosa*. Of the six array sizes unique to *L. saligna*, four of them were the largest identified (119–125 bp). The one that was specific to *L. virosa* was the smallest detected (66 bp). The overlap in array sizes, but not haplotypes, between species could be either due to convergence reflecting the limited number of possible 2-bp changes in array size or to the different species sharing RGC2 genes that are identical by descent.

Polymorphism for MSAT15-34 among populations of *L. serriola* from Israel and California

To investigate the distribution of polymorphism in the major cluster in natural populations, MSAT15-34 was analyzed in individuals collected from 13 populations in Israel and California. Haplotypes were determined successfully for 492 out of 505 samples. Fourteen array sizes were observed in these populations, all except three had previously been observed in the resistant *L. serriola* germplasm (Table 2 and Fig. 4C). In contrast, out of a total of 33 haplotypes identified, only eight had been observed previously in the resistant germplasm. Haplotypes containing a single array were always the most frequent within a country, as well as within a population. Six arrays were detected in one accession indicating that as much duplication of RGC2 genes can occur in natural populations of *L. serriola* as in *L. sativa* (Meyers et al. 1998a).

Levels of diversity were compared between Israel and California. In Israel 17 different haplotypes were detected, and in California 22 (Table 4A). In each country, one haplotype was prevalent and all the others had frequencies less than 0.15 (Table 4B). Only six haplotypes (EV, EI, BD, AN, ET and No) were common to samples from both Israel and California (Table 4B). Moreover the most frequent Israeli haplotype (AQ) was never observed in California and the most frequent haplotype in California (EV) was only represented by two Israeli individuals in a single population. The diversity within each country was estimated using the Shannon index. This index takes into account the number and frequency of haplotypes. The California population was as polymorphic as the Israeli population (Table 4A). Therefore, the data provided no evidence of a founder effect reflecting colonization of California by a limited number of genotypes. The distribution of array sizes in both countries was bimodal; however, both peaks in the distribution for California are smaller than those from Israel (Fig. 5). The majority of individuals (371) had a haplotype comprising only a single array; however, when two arrays were present, one from each size class was usually present (90 out of 114).

The Shannon index was also used to estimate the level of haplotype polymorphism within each population in Israel and California (Table 4A). The intra-population diversity depended highly on the location. In Israel, the population from Shefayim was monomorphic and the population from Netanya exhibited significantly less polymorphism than the one from North Nazareth. In California, the populations from Yolo-Landfill, Davis CR29 and Gilroy were significantly more polymorphic than the three other populations. The specific haplotypes also varied between locations (Table 4B). Out of 17 haplotypes in Israel, seven were found only in single populations. Similarly, ten of the haplotypes in California were found only in unique populations. However, these uncommon haplotypes were always present at a frequency below 2%.

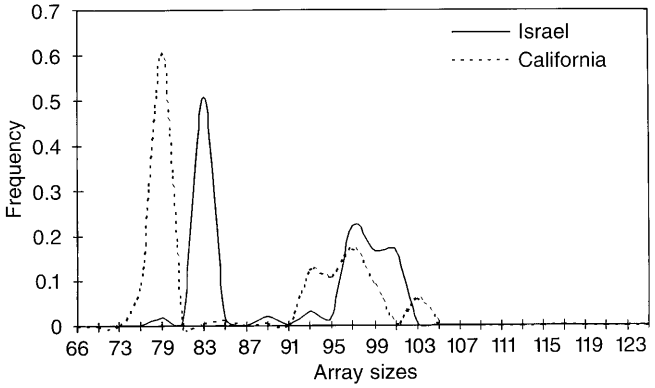

Fig. 5 Frequencies of array sizes for microsatellite marker, MSAT15-34, detected in Israel and California

The partitioning of the diversity among countries and locations was also studied using a molecular variance analysis (Schneider et al. 1997). A total of 66% of the haplotype diversity was within location, 16% among locations within a country and 18% between countries. Statistical analysis revealed significant differences between countries ($P<0.01$) and between locations ($P<0.01$). Pair-wise distances between populations were always significant except between Davis CR32 – Davis Putah creek, Haifa – Tel Hanan, and Beer Sheva – Tel Hanan. High levels of diversity within populations make it possible to study the dynamics of resistance-gene evolution within a single population; however, each population only represents a sub-sample of the variation within each country.

Fingerprints of the LRR-encoding regions and correlation between markers

To provide additional information on polymorphism in RGC genes, two pairs of primers were designed to amplify from multiple *RGC2* sequences encoding two different regions of the LRR (Fig. 2). All samples of cultivated and wild germplasm were analyzed (Table 2). Approximately ten samples in each population of *L. serriola* representing the diversity of MSAT15-34 haplotypes were also studied (Table 5). A product of 703-bp was amplified from sequences encoding the anterior LRR region (ALRR) and a product of 661-bp from the sequences encoding the mild-LRR region (MLRR). Surprisingly, no length polymorphism was detected with either set of primers from 150 germplasm and 137 *L. serriola* samples (Fig. 6A), except that under some conditions a less intense second band was observed in eight accessions for the MLRR region (Table 2B). Therefore, there is minimal variation in the numbers of LRRs in these regions for multiple copies of the gene across diverse germplasm.

Polymorphism was detected after digestion of the amplification products with *Tsp*509I, an endonuclease with a 4-bp recognition site. After digestion of the ALRR

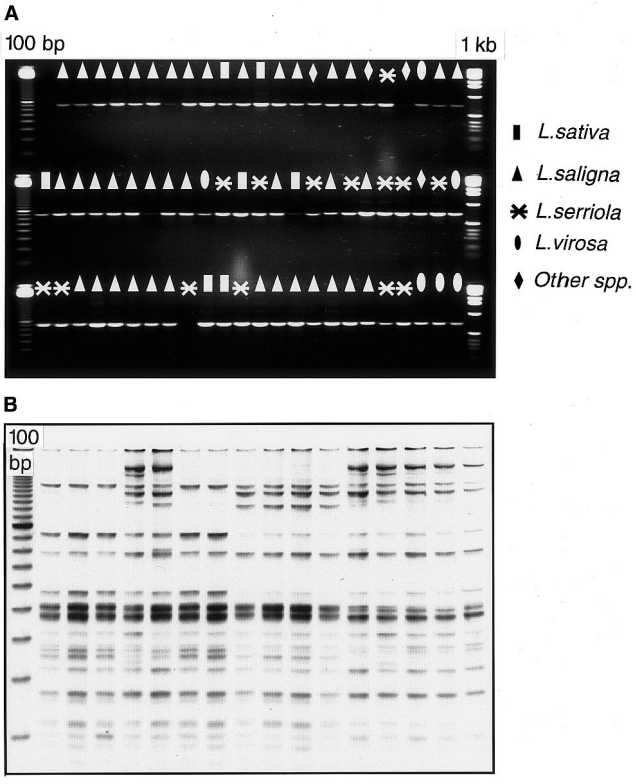

Fig. 6A, B *RGC2* fingerprints of *Lactuca* spp. **A** The 700-bp segment encoding the anterior LRR region (ALRR) was amplified from 71 accessions of four *Lactuca* spp. The left-hand peripheral lanes contain the 100-bp ladder (Pharmacia, Alameda, Calif., USA) and the right-hand lanes contain the 1-kb ladder (Life Technologies, Grand Island, N.Y., USA) as size markers. No length variation was apparent. **B** The 700-bp segment encoding the ALRR was amplified from 16 samples of *L. serriola* collected from California and digested with *Tsp*509I to reveal polymorphisms. The left-hand peripheral lane contains the 100-bp ladder

product, 46 different haplotypes were obtained in the germplasm collection. After digestion of the MLRR product, 63 haplotypes were detected. Sixteen and fifteen haplotypes were observed in natural populations of *L. serriola* in California and Israel respectively (Table 5). As excepted, the sum of fragment sizes observed exceeded the size of the undigested band, reflecting amplification from multiple *RGC2* genes. There were often fragments in common between accessions. For example, within *L. sativa* 60% of the bands were present in at least half the accessions; similarly, 72% of the bands were present in over half of the *L. serriola* samples from California. The diversity of *RGC2* haplotypes correlated well with the MSAT15-34 haplotypes within each species (Table 3). The levels of diversity detected by the three markers were not significantly different using the Shannon index (Table 5). The level of resolution with the *RGC2* haplotypes was slightly lower than with MSAT 15-34, i.e. some groups with identical *RGC2* haplotypes could be split on the basis of MSAT15-34 data. Less frequently, MSAT15-34 haplotypes could be split on the basis of RGC2 SCARs markers. Therefore, MSAT15-34 was a valid marker for diversity in *RGC2* sequences.

Discussion

Sequences related to resistance genes were amplified from all genotypes by at least two of the three markers. There was no evidence for a genotype that lacked homologs related to *RGC2*, even in lines that express no known resistance genes at this locus. The high level of haplotype diversity observed in wild species suggests that the diversity of resistance phenotypes previously identified within *L. sativa* is not an artifact of selection for resistance by plant breeders. The large number of different haplotypes detected in wild germplasm indicated that there is little redundancy and that these accessions will be a rich source of new resistance genes.

We developed highly informative molecular markers for assessing resistance gene diversity from the sequences of resistance gene homologs. These are reliable markers for resistance because recombination between the marker and the determinants of the resistance phenotype will be extremely rare. Haplotypes identified with all three markers were highly correlated, as expected for sequences derived from the same gene. MSAT15-34 was slightly more polymorphic and was therefore used for the majority of the analyses.

MSAT15-34 is a highly polymorphic marker. This may in part be due to the compound nature of this large array $[(CA)_wT(AG)_x(AAG)_y(AG)_z]$ as well as amplification from multiple duplicated genes. A total of 27 array

Table 5 Comparisons of diversity detected with different markers derived from *RGC2*

Lactuca spp.	N[a]	No. of Haplotype			Shannon's index		
		MSAT	ALRR	MLRR	MSAT	ALRR	MLRR
Germplasm accessions							
L. sativa	73	23	19	21	2.41	2.29	2.22
L. saligna	45	28	18	24	3.17	2.42	2.73
L. serriola	15	13	13	14	2.52	2.52	2.62
L. virosa	7	7	5	6	–	–	–
Natural populations							
L. serriola,Israel	69	17	13	15	2.35	2.02	2.16
L. serriola, California	68	22	16	16	2.71	2.0	2.04

[a] This analysis takes into account only individuals for which amplifications were obtained with the three markers

sizes was detected; this is greater than the 21 or 17 alleles reported at the most-polymorphic microsatellite loci in *Glycine* spp. (Maughan et al. 1995) or *Arabidopsis thaliana* (Innan et al. 1997). Although microsatellites have been characterized in other multigene families, they do not exhibit the variation observed for MSAT15-34 (Saha et al. 1993; Maughan et al. 1995; Nenoi et al. 1998).

Variable levels of duplication were observed in addition to variation in array size. More than two products were amplified from MSAT15-34 in different accessions of all the *Lactuca* spp. studied as well as in natural populations of *L. serriola*. Up to seven arrays were observed from a single genotype. MSAT15-34 amplifies from 5 out of 24+ members of the *RGC2* family of sequences clustered at this locus in the cultivar Diana (Meyers et al. 1998a). The complex haplotypes found in the wild species demonstrate that duplication events at resistance gene clusters are not a consequence of selection during breeding. The bimodal distribution of array sizes in Israel and California may be indicative of two allelic lineages (unless the compound nature of the microsatellite tends to generate some allelic forms more than others). Interestingly, when two arrays were amplified, one from each size class was present in most genotypes.

The same array sizes were often present in multiple species; however, haplotypes were often unique to a species. This could be due to several reasons. MSAT15-34 may evolve following a stepwise mutation model (Kimura and Ohta 1978) in each species with the allelic state changing back and forth, giving rise to the same array size in different species by separate mutation events. However, the tight correlation between the microsatellite and the ALRR and MLRR haplotypes within species suggest that this is not the case. Alternatively, the diversity in array sizes may pre-date speciation, and the high intraspecific haplotype diversity has been generated by recombination events producing new combinations of *RGC2* paralogs.

Although a large number of array sizes were observed, the high correlation between MSAT15-34 and the other two markers suggests that none of them are evolving rapidly enough to breakdown the correlation. It also indicates that although unequal crossing-over is occurring to change copy number (as evidenced by the changes in numbers of arrays), the cross-over events are not occurring frequently within the genes. This is consistent with the divergent selection and birth-and-death model proposed for the evolution of disease resistance genes (Michelmore and Meyers 1998). These markers are amplified from paralogs distributed across the whole haplotype that spans several megabases (Meyers et al. 1998a); the consistent correlation between markers may indicate that selection is occurring to maintain the correlation through hitchhiking of linked polymorphisms.

Very little length variation was observed in either the ALRR and MLRR markers amplified from multiple paralogs in diverse germplasm spanning four species. These were of the 703- and 661-bp regions encoding part of the LRR region. This was surprising as changes in the number of LRRs has been proposed to be involved in the evolution of new resistance specifities (Hammond Kosack and Jones 1997; Jones and Jones 1997). In contrast, restriction-site polymorphisms were detected suggesting that substitution or small indels are more frequent than large insertions or deletions. This finding extends data obtained from the sequencing of nine RGC2 paralogs from cv Diana; there was little length variation in the seven exons (Meyer et al. 1998b). However, within each LRR, one motif was hyper-polymorphic suggesting that point mutation was probably an important mechanism of variation at resistance gene clusters. Together these results suggest that point mutations are more important than variation in the number of LRRs in this region for the generation of diversity within and between species.

All of these three markers will be highly informative in breeding programs and conservation purposes. There were strong correlations between these markers and the characterized *Dm* genes of the main cluster within the cultivated *Lactuca* studied. Therefore MSAT15-34 could be used as a genetic marker to identify germplasm carrying diverse alleles as well as in subsequent marker-aided selection. In the analysis of wild *Lactuca* spp., MSAT15-34 could distinguish between accessions that were monomorphic with respect to their resistance against most of the known pathotypes of *B. lactucae*. These molecular markers will allow the identification of recombinant haplotypes at this locus and different resistances effective against all known isolates to be combined. Also, neutral markers have been used to assess diversity and guide the establishment of core collections; it is unclear if this results in representative sampling of adaptive variation as well (Nevo et al. 1986; Bataillon et al. 1996). Molecular markers developed from agriculturally or ecologically significant loci such as those developed in this study for resistance will be helpful in choosing lines for inclusion in core collections.

High levels of diversity were detected in both Israel and California and there was no evidence for limited variation due to a founder effect in California. There was only limited overlap in the haplotypes observed from the two countries. Moreover, two different classes of MSAT15-34 array sizes were observed in these two countries. These results suggest that the Californian populations did not originate from Israel. More extensive sampling from Israel and other parts of the Mediterranean region and the analysis of neutral molecular markers is required to determine the origin of the Californian populations as well as the relationship of Israel to the center of diversity for *Lactuca* spp.

The variation within and between populations reflects the biology of *Lactuca* spp. Like *A. thaliana* (Innan et al. 1997) and *Linum marginale* (Lawrence and Burdon 1989), *L. serriola* is an autogamous species that invades disturbed habitats and may undergo rapid extinctions and re-colonizations. In the present study, locations differed significantly for both the haplotypes present and the lev-

el of polymorphism within the population. This is comparable to data obtained for other autogamous wild plants studied for neutral molecular markers as well as resistance phenotype (Heywood 1991; Bonnin et al. 1996; Cattan-Toupance et al. 1998). Selfing results in homozygosity and tends to increase the diversity between populations. This will be accentuated by local founding effects. The population at Shefayim could clearly have resulted from a recent colonization event; this population was monomorphic for a haplotype found in two nearby coastal populations. Migration was not sufficient to homogenize the populations of *L. serriola* within either Israel or California.

There is little data on the diversity of resistance in natural populations. Levels of polymorphism for resistance have been correlated with ecological conditions or with pathogen population structure indicative of the amount of heterogeneity in selection pressures (Whal 1970; Burdon et al. 1983; Moseman et al. 1984; Parker 1985; Lawrence and Burdon 1989). Intrapopulation polymorphism for resistance in plant natural populations has been observed within a population for the allogamous species *Senecio vulgaris* as well as for the autogamous species *Amphicarpea bracteata* (Clarke 1997; Parker 1988). Our data provides the first molecular evidence for resistance gene diversity in wild populations. The populations were not sampled with reference to the levels of any disease but rather from an ecogeographical perspective. The major cluster of resistance genes in lettuce contains over 20 resistance gene homologs and this locus is known to encode resistance against several races of *B. lactucae* as well as root aphid. This locus probably determines as yet uncharacterized resistance to additional pathogens. Therefore the selection pressure exerted on this cluster may be highly complex and diverse, explaining why locations as well as countries are differentiated.

Future studies will focus on determining the relative rates of duplication, unequal crossing-over and divergent selection of resistance genes in natural populations of *L. serriola.* High levels of polymorphism were detected within some populations in the current study. Therefore, such investigations of resistance gene evolution should be at the level of local populations rather than across large geographic regions. Intensive sampling of large numbers of individuals from these populations will allow us to identify rare novel haplotypes and indicate the progenitor alleles. Subsequent molecular characterization should reveal the genetic mechanisms generating the diversity.

Acknowledgements We thank Rebecca Silady for technical assistance. The research was supported by the US – Israel BARD fund #US-2547-95. R.A.-G. was supported by a fellowship from the Ministry of Agriculture, Espana. The experiments comply with the current laws of the USA in which the experiments were performed. E.N. was also supported by the Israeli Discount Bank Chair of Evolutionary Biology and the Ancell-Teicher Research Foundation of Genetics and Molecular Evolution.

References

Bataillon TM, David JL, Shoen DJ (1996) Neutral genetic markers and conservation genetics: simulated germplasm collections. Genetics 144:409–417

Bernatzky R, Tanksley SD (1986) Genetics of actin-related sequences in tomato. Theor Appl Genet 72:314–321

Bonnier FJM, Reinink K, Groenwald R (1992) A search for new sources of major gene resistance in *Lactuca* to *Bremia lactucae* Regel. Euphytica 61:203–211

Bonnier FJM, Reinink K, Groenwald R (1994) Genetic analysis of *Lactuca* accessions with new major gene resistance to lettuce downy mildew. Phytopathology 84:462–468

Bonnin I, Huguet T, Gherardi M, Prosperi JM, Olivieri I (1996) High level of polymorphism and spatial structure in a selfing plant species, *Medicago truncatula* (Leguminosae), shown using RAPD markers. Am J Bot 30:843–855

Burdon JJ (1996) The role of race-specific resistance in natural plant populations. Oikos 76:411–416

Burdon JJ (1997) The evolution of gene-for-gene interactions in natural pathosystems. In: Crute IR, Holub EB, Burdon JJ (eds) The gene-for-gene relationship in plant-parasite interactions. CAB international, Oxon, UK, pp 245–262

Burdon JJ, Oates JD, Marshall DR (1983) Interaction between *Avena* and *Puccinia* species. I. The wild hosts: *Avena barbata* Pott ex link, *A. fatua* L., *A. ludoviviana* Durieu. J Appl Ecol 20:571–584

Cattan-Toupance I, Michalakis Y, Neema C (1998) Genetic structure of wild bean populations in their South-Andeau centre of origin. Theor Appl Genet 96:844–851

Clarke DD (1997) The genetic structure of natural pathosystems. In: Crute IR, Holub EB, Burdon JJ (eds) The gene-for-gene relationship in plant-parasite interactions. CAB international, Oxon, UK, pp 231-243

Cole RA, Sutherland RA, Riggall WE (1991) The use of polyacrylamide gradient gel electrophoresis to identify variation in isozymes as markers for *Lactuca* species and resistance to the lettuce roon aphid *Pemphigus bursarius*. Euphytica 56:237–42

Crute IR (1990) Resistance to *Bremia lactucae* (downy mildew) in British populations of *Lactuca serriola* (prickly lettuce). In: Burdon JJ, Leather SR (eds) Pests, pathogens and plant communities. Blackwell Scientific Publications, Oxford, UK, pp 203–217

Crute IR (1992) The role of resistance breeding in the integrated control of downy mildew (*Bremia lactucae*) in protected lettuce. Euphytica 63:95–102

Crute IR, Johnson AG (1976) The genetic relationship between races of *Bremia lactucae* and cultivars of *Lactuca sativa.* Ann Appl Biol 83:125–137

Dale MFB (1991) The use of genetic resources in breeding and breeding research. In: Hawkes JG (ed) Genetic conservation of world crop plants. Academic Press, London, UK, pp 3–10

De Vries IM (1997) Origin and domestication of *Lactuca sativa* L. Genet Resour Crop Evol 44:165–174

Ellis JG, Lawrence GJ, Finnegan EJ, Anderson PA (1995) Contrasting complexity of two rust resistance loci in flax. Proc Natl Acad Sci USA 92:4185–4188

Farrara B, Michelmore RW (1987) Identification of new sources of resistance in *Lactuca* spp. Hort Science 22:647–649

Farrara B, Illot TW, Michelmore RW (1987) Genetic analysis of factors for resistance to downy mildew (*Bremia lactucae*) in lettuce (*Lactuca sativa*). Plant Pathol 36:499–514

Ferakova V (1977) The genus *Lactuca* L. in Europe. Univerzita Komenskeho, Bratislava, Poland

Flor HH (1956) The complementary systems in flax and flax rust. Adv Genet 8:29–54

Hammond-Kosack K, Jones JDG (1997) Plant disease resistance genes. Annu Rev Plant Physiol Plant Mol Biol 48:575–607

Hawkes JG (1991) The importance of genetic resources in plant breeding. In: Hawkes JG (ed) Genetic conservation of world crop plants. Academic Press, London, UK, pp 3–10

Heywood JS (1991) Spatial analysis of genetic variation in plant populations. Annu Rev Ecol Syst 22:335–355
Hill M, Witsenboer H, Zabeau M, Vos P, Kesseli R, Michelmore RW (1996) AFLP fingerprinting as a tool for studying genetic relationships in *Lactuca* spp. Theor Appl Genet 93:1202–1210
Hulbert SH (1997) Structure and evolution of the *Rp1* complex conferring rust resistance in maize. Annu Rev Phytopathol 35:293–310
Hulbert SH, Michelmore RW (1985) Linkage analysis of genes for resistance to downy mildew (*Bremia lactucae*) is lettuce (*Lactuca sativa*). Theor Appl Genet 70:520–528
Hutcheson K (1970) A test for comparing diversities based on the Shannon formula. J Theor Biol 29:151–154
Innan H, Terauchi R, Miyashita T (1997) Microsatellite polymorphism in natural populations of wild plant *Arabidopsis thaliana*. Genetics 146:1441–1452
Jana S, Nevo E (1991) Variation in response to infection with *Erysiphe graminis hordei* and *Puccinia hordei* in some wild barley populations in a centre of diversity. Euphytica 57:130–140
Jones DA, Jones JDG (1997) The roles of leucine-rich proteins in plant defences. Adv Plant Pathol 24:89–167
Kesseli RV, Michelmore RW (1986) Genetic variation and phylogenies detected from isozyme markers in species of *Lactuca*. J Hered 77:324–331
Kesseli RV, Ochoa O, Michelmore RW (1991) Variation at RFLP loci in *Lactuca* spp. and origin of cultivated lettuce. Genome 34:430–436
Kesseli RV, Witsenboer H, Stanghellini M, Vandermark G, Michelmore RW (1993) Recessive resistance to *Plasmopara lactucae-radicis* maps by bulked segregant analysis of dominant disease resistance genes in lettuce. Mol Plant-Microbe Interact 6:722–728
Kimura M, Ohta T (1978) Stepwise mutation model and distribution of allelic frequencies in a finite population. Proc Natl Acad Sci USA 75:2868–2872
Lawrence GJ, Burdon JJ (1989) Flax rust from *Linum marginale*: variation in a natural host-population interaction. Can J Bot 67:3192–3198
Lenne JM, Wood D (1991) Plant diseases and the use of wild germplasm. Annu Rev Phtopathol 29:35–63
Leppik EE (1970) Gene centers of plants as sources of disease resistance. Annu Rev Phytopathol 8:323–344
Lindqvist K (1960a) On the origin of cultivated lettuce. Hereditas 46:319–350
Lindqvist K (1960b) Cytogenetic studies in the *serriola* group of *Lactuca*. Hereditas 46:75–151
Maisonneuve B, Bellec Y, Anderson P, Michelmore RW (1994) Rapid mapping of two genes for resistance to downy mildew from *Lactuca serriola* to existing clusters of resistance genes. Theor Appl Genet 89:96–104
Maughan PJ, Saghai Maroof MA, Buss GR (1995) Microsatellite and amplified sequence length polymorphisms in cultivated and wild soybean. Genome 38:715–723
Meyers BC, Chin DB, Shen KA, Sivaramakrisman S, Lavelle DO, Zhang Z, Michelmore RW (1998a) The major resistance gene cluster in lettuce is highly duplicated and spans several megabases. Plant Cell 10:1817–1832
Meyers BC, Shen KA, Rohani P, Gaut BS, Michelmore RW (1998b) Receptor-like genes in the major resistance locus of lettuce are subject to divergent selection. Plant Cell 10:1833–1846
Michelmore RW, Meyers BC (1998) Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. Genome Res 8:1113–1130
Michelmore RW, Hulbert SH, Landry BS, Leung H (1987) Towards a molecular understanding of lettuce downy mildew. In: Day PR, Jellis GJ (ed) Genetics and plant pathogenesis. Blackwell Scientific Publications, Oxford, UK, pp 221–231
Michelmore RW, Paran I, Kesseli RV (1991) Identification of markers linked to disease resistance genes by bulked segregant analysis: a rapid method to detect markers in specific regions by using segregating populations. Proc Natl Acad Sci USA 88:9828–9832
Moseman JGE, Nevo E, El Morshidy MA, Zohary D (1984) Resistance of *Triticum dicoccoides* to infection with *Erysiphe graminis tritici*. Euphytica 33:41–47
Nenoi M, Mita K, Ichimura S, Kawano A (1998) Higher frequency of concerted evolutionary events in rodents than in man at the polyubiquitin gene VNTR locus. Genetics 148:867–876
Nevo E, Moseman JG, Beiles A, Zohari D (1985) Patterns of resistance of Israeli wild emmer wheat to pathogens: predictive method by ecology and allozyme genotypes for powdery mildew and leaf rust. Genetica 67:209–222
Nevo E, Bailes A, Kaplan D, Golenberg EM, Olsvig-Whittaker LS, Naveh Z (1986) Natural selection of allozyme polymorphisms: a microsite test revealing ecological genetic differentiation. Evolution 40:13–20
Ochoa O, Delp B, Michelmore RW (1987) Resistance in *Lactuca* spp. to *Microdochium panattoniana* (lettuce anthracnose). Euphytica 39:609–614
Okubara PA, Arroyo-Garcia R, Shen KA, Mazier M, Meyers BC, Ochoa OE, Kim S, Yang C-H, Michelmore RW (1997) A transgenic mutant of *Lactuca sativa* (Lettuce) with a T-DNA tightly linked to loss of downy mildew resistance. Mol Plant-Microbe Interact 10:970–977
Paran I, Michelmore RW (1993) Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. Theor Appl Genet 85:985–993
Parker MA (1985) Local population differentiation for compatibility in an annual legume and its host-specific pathogen. Evolution 39:713–723
Parker MA (1988) Polymorphism for disease resistance in the annual legume *Amphicarpaea bracteata*. Heredity 60:27–31
Peet RK (1974) The measurement of species diversity. Annu Rev Ecol system 5:285–307
Pryor T (1987) Origin and structure of disease resistance genes in plants. Trends Genet 3:157–161
Pryor T, Ellis J (1993) The genetic complexity of fungal resistance genes in plants. Adv Plant Pathol 10:281–305
Robbins MA, Witsenboer H, Michelmore RW, Laliberte JF, Fortin M (1994) Genetic mapping of turnip mosaic virus resistance in *Lactuca sativa*. Theor Appl Genet 89:583–589
Ryder EJ (1986) Lettuce breeding. In: Bassett MJ (ed) Breeding vegetable crops. AVI Publishing Company, Westport, pp 433–474
Saha BK, Shields JJ, Miller RD, Hansen TD, Schreffler DC (1993) A highly polymorphic microsatellite in the class-II *Eb* gene allows tracing of major histocompatibility complex evolution in mouse. Proc Natl Acad Sci USA 90:5312–5316
Schneider S, Kueffer J-M, Roessli D, Exdoffier L (1997) Arlequin: a software for population genetic data analysis. Ver 1.1. Genetics and Biometry Laboratory, Department of Anthropology, University of Geneva
Shen KA, Meyers BC, Islam Faridi MN, Chin DB, Stelly DM, Michelmore RW (1998) Resistance gene candidates identified by PCR with degenerate oligonucleotide primers map to clusters of resistance genes in lettuce. Mol Plant-Microbe Interact 11:815–823
Song WY, Pi L-Y, Wang G-L, Gardner J, Holsten T, Ronald PC (1997) Evolution of the rice *Xa21* disease resistance gene family. Plant Cell 9:1279–1287
Wahl I (1970) Prevalence and geographic distribution of resistance to crown rust in *Avena sterilis*. Phytopathology 60:746–749
Whitaker TW (1969) Salads for everyone – a look at the lettuce plant. Econ Bot 23:261–264
Witsenboer H, Kesseli RV, Fortin M, Stanghellini M, Michelmore RW (1995) Sources and genetic structure of a cluster of genes for resistance to three pathogens in lettuce. Theor Appl Genet 91:78–188
Witsenboer H, Vogel J, Michelmore RW (1998) Identification, genetic localization, and allelic diversity of selectively amplified microsatellite polymorphic loci in lettuce and wild relatives (*Lactuca* spp.) Genome 40:923–936
Zohary D (1991) The wild genetic resources of cultivated lettuce (*Lactuca sativa* L.). Euphytica 53:31–35