ORIGINAL PAPER

H. Ozkan · A. Brandolini · C. Pozzi · S. Effgen
J. Wunder · F. Salamini

A reconsideration of the domestication geography of tetraploid wheats

Abstract The domestication of tetraploid wheats started from their wild progenitor *Triticum dicoccoides*. In this paper, the geographical distribution of this progenitor is revised to include more sampling locations. The paper is based on a collection of wild and domesticated lines (226 accessions in total) analyzed by AFLP at 169 polymorphic loci. The collection includes the 69 wild lines considered by Mori et al. (2003) in their study on chloroplast DNA haplotypes of *T. dicoccoides*. The goal of the experiment was to reconsider which location thought to have generated the domesticated germplasm has the highest chance of being the actual site from which wild progenitors were sampled during domestication. Phylogenetic analysis of the nuclear AFLP databases indicates that two different genetic taxa of *T. dicoccoides* exist, the western one, colonizing Israel, Syria, Lebanon and Jordan, and the central-eastern one, which has been frequently sampled in Turkey and rarely in Iran and Iraq. It is the central-eastern race that played the role of the progenitor of the domesticated germplasm. This is supported by the cumulative results of the AFLP data from the collections of Ozkan et al. (2002)

and of Mori et al. (2003), which indicate that the Turkish Karacadag population, intermixed with some Iraq-Iran lines, has a tree topology consistent with that of the progenitor of domesticated genotypes. The Turkish Kartal population belongs genetically to the central-eastern *T. dicoccoides* race but at the nuclear DNA level is less related to the domesticated gene pool. A general agreement between published work on tetraploid wheat domestication emerges from these results. A disagreement is nevertheless evident at the local geographical scale; the chloroplast DNA data indicate the Kartal mountains while AFLP fingerprinting points to the Karacadag Range as the putative site of tetraploid wheat domestication.

Introduction

Two lines of evidence support the argument that the Fertile Crescent is the site of origin of Western agriculture. First, the geographical distributions of the wild progenitors of wheat (*Triticum boeticum, T. urartu, T. dicoccoides, Aegilops tauschii*), wild barley (*Hordeum spontaneum*) and wild rye (*Secale vavilovii*) intersect in this region (Nesbitt and Samuel 1996; Zohary and Hopf 2000; Salamini et al. 2002). Second, seeds of the wild species have been found in early archaeological sites of the region, followed in stratigraphic succession by the remains of domesticated forms (Moore et al. 2000; Gopher et al. 2002).

The domestication of the tetraploid wheats started from their wild progenitor *T. dicoccoides*. This species, with its A genome derived from *T. urartu* (Dvorak et al. 1993, 1998), has brittle ears that shatter at maturity. Domesticated emmer wheat, *T. dicoccum*, has hulled seeds and its AABB genome is common to other domesticated tetraploid wheats. Emmer was the most important crop in the Fertile Crescent until the early Bronze Age (Bar-Yosef 1998), and domesticated forms

H. Ozkan
Department of Field Crops, Faculty of Agriculture,
University of Cukurova, 01330 Adana, Turkey

A. Brandolini
Istituto Sperimentale per la Cerealicoltura,
via R. Forlani 3, S. Angelo Lodigiano, Italy

S. Effgen · J. Wunder · F. Salamini
Max-Planck-Institut für Züchtungsforschung,
Carl-von-Linné Weg 10, Köln, Germany

Present address: C. Pozzi · F. Salamini (✉)
Parco Tecnologico Padano, Polo Universitario,
via Einstein 2 Llocalità Cascina Codazza,
26900 Lodi, Italy
E-mail: francesco.salamini@tecnoparco.org
Tel.: +39-0371-442438
Fax: +39-0371-437385

are present at early Neolithic sites (such as Tell Aswad, ~10,800 BP; Van Zeist and Bakker-Heeres 1982). Naked forms, such as the small grain *T. parvicoccum* (Kislev 1980), are also found in early Neolithic locations.

The geographical distribution of *T. dicoccoides* reported by Zohary and Hopf (2000) includes the western Fertile Crescent, the central part of southeastern Turkey and areas in eastern Iran and Iraq. Johnson (1975) reports that from southeastern Turkey to Iran and Iraq the species is progressively substituted by the wild tetraploid wheat *T. araraticum*. *T. dicoccoides* also grows on the basaltic rocky slopes of the Karacadag mountains in southeastern Turkey (Johnson 1975; Harlan and Zohary 1966).

AFLP data have indicated that domesticated tetraploid AABB wheats are most closely related to wild emmer populations from southeastern Turkey (Ozkan et al. 2002). In this study, 99 lines of wild emmer from Turkey, Israel, Jordan, Lebanon, Syria and Iran, as well as 19 wild emmer lines from a Karacadag population (southeastern Turkey), were studied. The genetic profile of 15 of the 19 Karacadag lines was consistent with their close genetic relationship to domesticated emmer. This was interpreted as more than a coincidence, considering that using the same technique the domestication of einkorn was also located to the Karacadag mountains in southeastern Turkey (Heun et al. 1997). Based on these and similar findings this area was later defined as the "core area" of the origin of agriculture (Lev-Yadun et al. 2000).

In the study by Ozkan et al. (2002) evidence was also provided that the hulled emmer and *T. durum* (free-threshing) tetraploid genotypes group separately on a phylogenetic tree, although the two groups merge into a common lineage soon before joining the *T. dicoccoides* line topologies. At least based on principal coordinate analysis, both a monophyletic and a diphyletic domestication of tetraploid wheats are possible (Salamini et al. 2004), with the former hypothesis having a higher probability. In addition, chloroplast DNA fingerprinting of wild and domesticated emmer wheat indicates that two distinct maternal lineages have been involved in the domestication process, suggesting that the event occurred at least twice (Mori et al. 2003). While referring again to southeastern Turkey as the domestication site for emmer wheat, these authors indicate that the Kartal mountains (280 km west of the Karacadag Range) host wild emmer lines having both chloroplast DNA haplotypes prevalent in domesticated emmer.

In this study, we use AFLP fingerprinting to analyze a new set of wild emmer lines, including 69 accessions from the collection of Mori et al. (2003), 7 lines from Dr. B. Gill, and 12 lines from Dr. A. Karagoz, and compared the results with those derived from 92 wild emmer lines previously characterized by Ozkan et al (2002). The reasoning behind this experiment was to reconsider which location of southern Turkey has the highest chance of being the actual site from which the wild progenitors of tetraploid wheats were sampled.

Materials and methods

Plant material

The nomenclature followed in this paper indicates with *T. dicoccoides* the wild lines of *T. turgidum* ssp. *dicoccoides*, with *T. dicoccum* the domesticated hulled emmer lines (*T. turgidum* ssp. *dicoccum*) and with *T. durum* the domesticated free-threshing hard wheat lines (*T. turgidum* ssp. *durum*) (see also Zohary and Hopf 2000).

A total of 224 wheat accessions were considered, including 179 accessions of *T. dicoccoides*, 22 of *T. durum*, 23 of *T. dicoccum* and 1 line of *T. polonicum*. The geographic origin and the source of the accessions are given in Tables 1 and 2. One hundred and thirty-one lines were those typed by Ozkan et al. (2002), of which 91 were wild *T. dicoccoides* accessions, 18 *T. dicoccum* emmer lines, and 22 *T. durum* hard wheat lines. Of the remaining lines analyzed in this paper, 69 were from the collection studied by Mori et al. (2003) and 7 and 12 lines were received from Dr. B. Gill (University of Kansas State, USA) and Dr. A. Karagoz (Department of Field Crops, Ankara), respectively. In addition, 5 *T. dicoccum* emmer lines were also obtained from Dr. A. Karagoz (Ankara University). The three collections had some lines in common based on their AFLP patterns.

Species identification of the 88 new lines characterized was carried out by field observations. Five plants were grown in pots and transplanted to the field after sampling material for DNA extraction.

Genomic DNA extraction and AFLP analyses

Young leaves were collected, lyophilized and kept at −70°C until use. DNA was extracted using the Qiagen DNA isolation kit (DNeasy Plant Mini, Qiagen, Hilden, Germany). AFLP procedure was performed according to Ozkan et al. (2002).

Data analysis

The AFLP gels were manually scored for the presence or absence of relevant bands. Each band was considered to represent a polymorphic locus. The genetic relationships among the taxa considered were established using the phylogenetic analysis package PHYLIP (Felsenstein 1989) and the multivariate analysis system NTSYS (Rohlf 1998). For the phylogenetic analysis concerning AFLP allele frequencies in populations or groups of lines, ten independent trees were constructed as described (Heun et al. 1997; Badr et al. 2000), using CONTML and distance matrix methods (Fitch and Margoliash 1967; Saitou and Nei 1987), and employing various measures of genetic distance (Cavalli-Sforza and Edwards 1967; Nei 1972; Wright 1978; Reynolds et al. 1983) calculated from AFLP allele frequency. To cluster single accessions (Fig. 1b), the neighbor-joining (NJ)

Table 1 List of the *Triticum diccoccoides* wild lines analyzed

Number	Accession number	Origin	Group	Number	Accession number	Origin	Group
Lines from the collection of Ozkan et al. (2002)							
IS1	17902[a]	Israel, Rosh Pina	IV	TR51	PI 538657[c]	Turkey, 37°48′N, 39°46′E	II
IS2	17901[a]	Israel, Korasim	IV	TR52	PI 538659[c]	Turkey, 37°47′N, 39°46′E	II
TR3	PGR0006[b]	Turkey, Karacadag	II	IS54	PI 538684[c]	Israel, 32°58′N, 35°32′E	IV
IS4	PI 233288[c]	Israel, 31°30′N, 34°45′E	IV	IS56	PI 538690[c]	Israel, 32°58′N, 35°29′E	IV
LB5	PI 352322[c]	Lebanon, 33°25′N, 35°52′E	V	IS57	PI 538699[c]	Israel, 31°50′N, 35°27′E	IV
IS6	PI 414718[c]	Israel, 32°54′N, 35°35′E	IV	LB58	PI 538700[c]	Lebanon, 33°30′N, 35°50′E	V
IS7	PI 414719[c]	Israel, 32°43′N, 35°23′E	IV	LB59	PI 538705[c]	Lebanon, 33°30′N, 35°52′E	V
IS8	PI 414720[c]	Israel, 32°35′N, 35°0′E	IV	LB60	PI 538708[c]	Lebanon, 33°30′N, 35°52′E	V
IS9	PI 414721[c]	Israel, 32°54′N, 35°32′E	IV	LB61	PI 538713[c]	Lebanon, 33°26′N, 35°46′E	V
IS10	PI 414722[c]	Israel, 32°58′N, 35°29′E	IV	TR63	PI 554580[c]	Turkey, 37°47′N, 39°46′E	II
IS11	PI 428013[c]	Israel, 32°58′N, 35°32′E	IV	TR64	PI 554581[c]	Turkey, 37°45′N, 40°6′E	II
IS12	PI 428014[c]	Israel, 31°47′N, 35°14′E	IV	TR65	PI 554582[c]	Turkey, 37°47′N, 39°46′E	II
IS13	PI 428015[c]	Israel, 31°30′N, 34°45′E	IV	TR66	PI 554583[c]	Turkey, 37°47′N, 39°47′E	II
TR14	PI 428017[c]	Turkey, 37°18′N, 36°48′E	II	SY68	PI 487255[c]	Syria, 33°45′N, 36°5′E	V
TR15	PI 428018[c]	Turkey, 37°53′N, 39°52′E	II	TR69	PI 503310[c]	Turkey, 37°53′N, 39°52′E	II
TR17	PI 428054[c]	Turkey, 37°50′N, 39°49′E	II	IS70	PI 503312[c]	Israel, 32°58′N, 35°32′E	IV
TR18	PI 428063[c]	Turkey, 37°48′N, 39°46′E	II	IS71	PI 503314[c]	Israel, 32°58′N, 35°29′E	IV
TR19	PI 428069[c]	Turkey, 37°48′N, 39°46′E	II	IS72	PI 503315[c]	Israel, 32°36′N, 35°5′E	IV
TR20	PI 428077[c]	Turkey, 37°47′N, 39°46′E	II	LB73	PI 503316[c]	Lebanon, 33°30′N, 35°50′E	V
TR22	PI 428092[c]	Turkey, 37°48′N, 39°46′E	II	TR74	IG 116173[d]	Turkey, Gaziantep, Kartal	I
IS23	PI 428093[c]	Israel, 32°36′N, 35°17′E	IV	TR75	IG 116184[d]	Turkey, Gaziantep, Kartal	I
IS24	PI 428097[c]	Israel, 32°54′N, 35°23′E	IV	SY76	IG 110737[d]	Syria, Suweida,	V
IS25	PI 428099[c]	Israel, 32°54′N, 35°23′E	IV	LB77	IG 110815[d]	Lebanon, Biqaa Al Gharbi	V
IS26	PI 428100[c]	Israel, 32°58′N, 35°32′E	IV	LB78	IG 46526[d]	Lebanon, Rachaiya,	V
IS27	PI 428105[c]	Israel, 32°58′N, 35°32′E	IV	SY79	IG 46504[d]	Syria, May Saloun	V
IS28	PI 428119[c]	Israel, 32°58′N, 35°29′E	IV	SY80	IG 46473[d]	Syria, Rawda	V
LB31	PI 428132[c]	Lebanon, 33°31′N, 35°52′E	V	SY81	IG 46466[d]	Syria, Suweida	V
LB32	PI 428135[c]	Lebanon, 33°31′N, 35°52′E	V	SY82	IG 45492[d]	Syria, Dar'a, Nawa	V
LB33	PI 428143[c]	Lebanon, 33°30′N, 35°50′E	V	SY84	IG 45490[d]	Syria, Suweida	V
IS34	PI 466926[c]	Israel, 31°30′N, 34°45′E	IV	SY85	IG 45493[d]	Syria, Dar'a, Nawa	V
IS35	PI 466949[c]	Israel, 32°2′N, 34°53′E	IV	SY86	IG 45494[d]	Syria, Zabadani	V
IS36	PI 466955[c]	Israel, 32°35′N, 35°0′E	IV	JO88	IG 46526[d]	Jordan, Ebbien	V
IS37	PI 466981[c]	Israel, 31°48′N, 35°2′E	IV	SY89	IG 45502[d]	Syria, Idlib	V
IS38	PI 466991[c]	Israel, 33°18′N, 35°48′E	IV	SY90	IG 46397[d]	Syria, Suweida	V
IS39	PI 466995[c]	Israel, 31°30′N, 34°45′E	IV	JO91	IG 45726[d]	Jordan, Irbid	V
IS40	PI 467004[c]	Israel, 32°52′N, 35°32′E	IV	JO92	IG 46323[d]	Jordan, Balqa	V
IS41	PI 470988[c]	Israel, 32°58′N, 35°32′E	IV	JO93	IG 45676[d]	Jordan, Irbid	V
IS42	PI 471016[c]	Israel, 31°48′N, 35°13′E	IV	JO94	IG 46324[d]	Jordan, Amman	V
IS43	PI 471035[c]	Israel, 31°57′N, 35°20′E	IV	JO96	IG 45964[d]	Jordan, Balqa	V
IS44	PI 479780[c]	Israel, 32°32′N, 35°21′E	IV	JO97	IG 46386[d]	Jordan, Amman	V
SY45	PI 487252[c]	Syria, 32°28′N, 36°47′E	V	SY98	IG 46439[d]	Syria, Suweida	V
SY46	PI 487253[c]	Syria, 32°48′N, 36°7′E	V	SY99	IG 46457[d]	Syria, Suweida	V
TR47	PI 538626[c]	Turkey, 37°53′N, 39°52′E	II	SY100	IG 46420[d]	Syria, Hama	V
TR48	PI 538633[c]	Turkey, 37°52′N, 39°53′E	II	IRAN102	PI 428016[c]	Iran, Bakhtaran	III
TR49	PI 538651[c]	Turkey, 37°50′N, 39°49′E	II	IRAQ103	TTD 151[h]	Iraq	III
TR50	PI 538656[c]	Turkey, 37°47′N, 39°46′E	II				
Lines from Mori et al. (2003)							
IRAQ105	8736A[e]	Iraq, SW of Rowanduz	III	TR148	1947[e]	Turkey, SE of Maras, Kartal	I
IRAQ106	8736B[e]	Iraq, SW of Rowanduz	III	TR149	1948[e]	Turkey, SE of Maras, Kartal	I
IRAQ107	8737[e]	Iraq, SW of Rowanduz	III	TR150	1949[e]	Turkey, SE of Maras, Kartal	I
IRAQ108	8804[e]	Iraq, N of Kursi	III'	TR151	1951[e]	Turkey, SE of Maras, Kartal	I
IRAQ109	8805[e]	Iraq, N of Kursi	III'	TR152	1952[e]	Turkey, SE of Maras, Kartal	I
IRAQ110	8806[e]	Iraq, N of Kursi	III'	TR153	1953[e]	Turkey, SE of Maras, Kartal	I
IRAQ111	8807[e]	Iraq, N of Kursi	III'	TR154	1955[e]	Turkey, SE of Maras, Kartal	I
IRAQ112	8808[e]	Iraq, N of Kursi	III'	TR155	1957[e]	Turkey, SE of Maras, Kartal	I
IRAQ113	8809[e]	Iraq, N of Kursi	III'	TR156	1959A[e]	Turkey, SE of Maras, Kartal	I
IRAQ114	8810[e]	Iraq, N of Kursi	III'	TR157	1959B[e]	Turkey, SE of Maras, Kartal	I
IRAQ115	8811[e]	Iraq, N of Kursi	III'	TR158	1972B[e]	Turkey, SE of Maras, Kartal	I
IRAQ116	8812[e]	Iraq, N of Kursi	III'	TR159	1974[e]	Turkey, SE of Maras, Kartal	I
IRAQ118	8815[e]	Iraq, N of Kursi	III'	TR160	1976B[e]	Turkey, SE of Maras, Kartal	I
IRAQ119	8816A[e]	Iraq, N of Kursi	III'	TR161	1978B[e]	Turkey, SE of Maras, Kartal	I
IRAQ120	8816B[e]	Iraq, N of Kursi	III'	TR162	1991[e]	Turkey, SE of Maras, Kartal	I
IRAQ121	8817[e]	Iraq, N of Kursi	III'	IRAQ163	8536[e]	Iraq, S of Sulaymaniyah	III
IRAQ122	8821A[e]	Iraq, NE of Dohuk	III'	IRAQ164	8537[e]	Iraq, S of Sulaymaniyah	III
IRAQ123	8821C[e]	Iraq, NE of Dohuk	III'	IRAQ165	8538[e]	Iraq, S of Sulaymaniyah	III
TR124	8915A[e]	Turkey, E of Silvan, near Karacadag	II	IRAQ166	8539[e]	Iraq, S of Sulaymaniyah	III

Table 1 (Contd.)

Number	Accession number	Origin	Group	Number	Accession number	Origin	Group
TR125	8915B[e]	Turkey, E of Silvan, near Karacadag	II	IRAQ167	8541[e]	Iraq, S of Sulaymaniyah	III
TR126	8935[e]	Turkey, SE of Ergani, near Karacadag	II	IS168	14443[e]	Israel, Sanhedriya	IV
TR127	8937B[e]	Turkey, SE of Ergani, near Karacadag	II	IS169	14445[e]	Israel, Sanhedriya	IV
IRAN128	8941[e]	Iran, N of Kermanshah	III	IS170	14451[e]	Israel, Bet Meir	IV
IRAN129	8942[e]	Iran, N of Kermanshah	III	IS171	14453[e]	Israel, Bet Meir	IV
IRAN130	8943[e]	Iran, N of Kermanshah	III	IS172	14462[e]	Israel, Mt. Hermon	IV
IS131	14401[e]	Israel, Katzrin	IV	IS173	14464[e]	Israel, Mt. Hermon	IV
IS132	14403[e]	Israel, Katzrin	IV	IS174	14474[e]	Israel, Tabigha	IV'
IS133	14417[e]	Israel, Yehudiya	IV'	IS175	14476[e]	Israel, Tabigha	IV'
IS134	14419[e]	Israel, Yehudiya	IV'	IS176	14490[e]	Israel, Bat Shlomo	IV
IS135	14427[e]	Israel, Rosh Pina	IV'	IS177	14492[e]	Israel, Bat Shlomo	IV
IS136	14429[e]	Israel, Rosh Pina	IV'	IS178	14505[e]	Israel, Tayiba	IV
SY144	108-2[e]	Syria, NW of Suweida	V	IS179	14507[e]	Israel, Tayiba	IV
SY145	108-3[e]	Syria, NW of Suweida	V	IS180	14517[e]	Israel, Kochav Hashahar	IV
TR146	1921[e]	Turkey, W of Mardin, near Karacadag	II	IS181	14519[e]	Israel, Kochav Hashahar	IV
TR147	1945[e]	Turkey, SE of Maras, Kartal	I				
Lines from other sources							
SY137	TA 122[f]	Syria, Gamla	V	TR186	TR 03371[g]	Turkey, Karacadag	II
IS138	TA 1030[f]	Israel, Rosh Pina	IV	TR187	TR 00842[g]	Sanliurfa, near Karacadag	II
SY139	TA 1058[f]	Syria, Al Qunavtirah	V	TR188	TR 03402[g]	Turkey, Karacadag	II
LB140	TA 1071[f]	Lebanon, 33°33′N, 35°53′E	V	TR189	TR 03362[g]	Turkey, Karacadag	II
TR141	TA 1138[f]	Turkey, 37°40′N, 39°35′E	II	TR190	TR 03399[g]	Turkey, Karacadag	II
JO142	TA 1181[f]	Jordan, Al Samirah	V	TR191	TR 02637[g]	Sanliurfa, near Karacadag	II
IS143	TA 1404[f]	Israel, Nahef	IV	TR192	TR 03388[g]	Turkey, Karacadag	II
TR182	TR 03396[g]	Turkey, Karacadag	II	TR193	TR 03346[g]	Turkey, Karacadag	II
TR183	TR 03369[g]	Turkey, Karacadag	II	TR194	TR 03391[g]	Turkey, Karacadag	II
TR185	TR 03376[g]	Turkey, Karacadag	II				

Sources of the lines used in this study:
[a]Australian Winter Cereals Collection
[b]Canada Cereals Germplasm Collection
[c]National Small Grains Collection, USA
[d]International Center for Agricultural Research in Dry Areas
[e]Kyoto Plant Germplasm Institute, School of Agriculture, Kyoto University
[f]WGRC: Kansas State University (J. Raupp and B.Gill)
[g]Field Crops Central Research Institute, Ankara, Turkey (A.Karagoz)
[h]Weizmann Institute of Science, Israel (Moshe Feldman)

Table 2 List of *T. durum* and *T. dicoccum* domesticated lines considered. Domesticated hulled and free-threshing lines fingerprinted were from Ozkan et al. (2002), except for HW21, HW22, HW23, HW24 and HW25, which were from A. Karagoz (see Materials and methods)

| Number	Accession number or name	Origin	Group	Number	Accession number or name	Origin	Group
T. durum (free-threshing)				*T. dicoccum* (hulled)			
DW1	Aristan	France	II	HW1	Farro garfagnana	Siena, Italy	II
DW2	Aziziah	Italian, from a Palestinian landrace	II	HW2	10	Italy	II
DW3	Baio	Italy	II	HW3	Farro da Trivento	Trivento, Italy	II
DW4	Capeiti 8	Italy	II	HW4	Leonessa 2	Leonessa, Italy	II
DW5	Cappelli	Italian, from a Tunisian landrace	II	HW5	Leonessa 5	Leonessa, Italy	II
DW7	Coll. Jordan	Landrace, Jordan	II	HW6	33	Montemonaco, Italy	II
DW8	Muri S 50 3	Cyprus	II	HW7	Agnone invernale	Agnone, Italy	II
DW9	Ofanto	Italy	II	HW8	Agnone primaverile	Agnone, Italy	II
DW10	Razzak	Tunisia	II	HW9	Monteleone	Monteleone S, Italy	II
DW12	Sabil 1	Syria	II	HW10	Potenza	Potenza, Italy	II
DW13	Santa	Greece	II	HW11	Riofreddo	Riofreddo, Italy	II
DW14	Simeto	Italy	II	HW12	VIR 54	Unknown	II
DW15	Taganrog	Landrace, Ukraine	II	HW14	BGRC 11025	India	II
DW16	Triminia	Italy, from landrace	II	HW15	BGRC 11053	Unknown	II
DW17	Tripolino	Italy, from landrace	II	HW16	BGRC 21074	Iran	II
DW18	Vatan	Tajikistan	II	HW17	BGRC 43443	Unknown	II
DW19	Villemur	France	II	HW19	ID 1120	Lebanon	II
DW20	Waha	Mexico	II	HW20	Lonigo 19	Italy	II
DW21	Bufala	Italy	II	HW21	TUR 03558	Turkey	II
DW22	Pavone	Italy	II	HW22	TUR 02440	Turkey	II
DW23	Russello	Italy	II	HW23	TUR 03560	Turkey	II
DW24	Timilia	Italy	II	HW24	TUR 03562	Turkey	II
				HW25	TUR 02456	Turkey	II
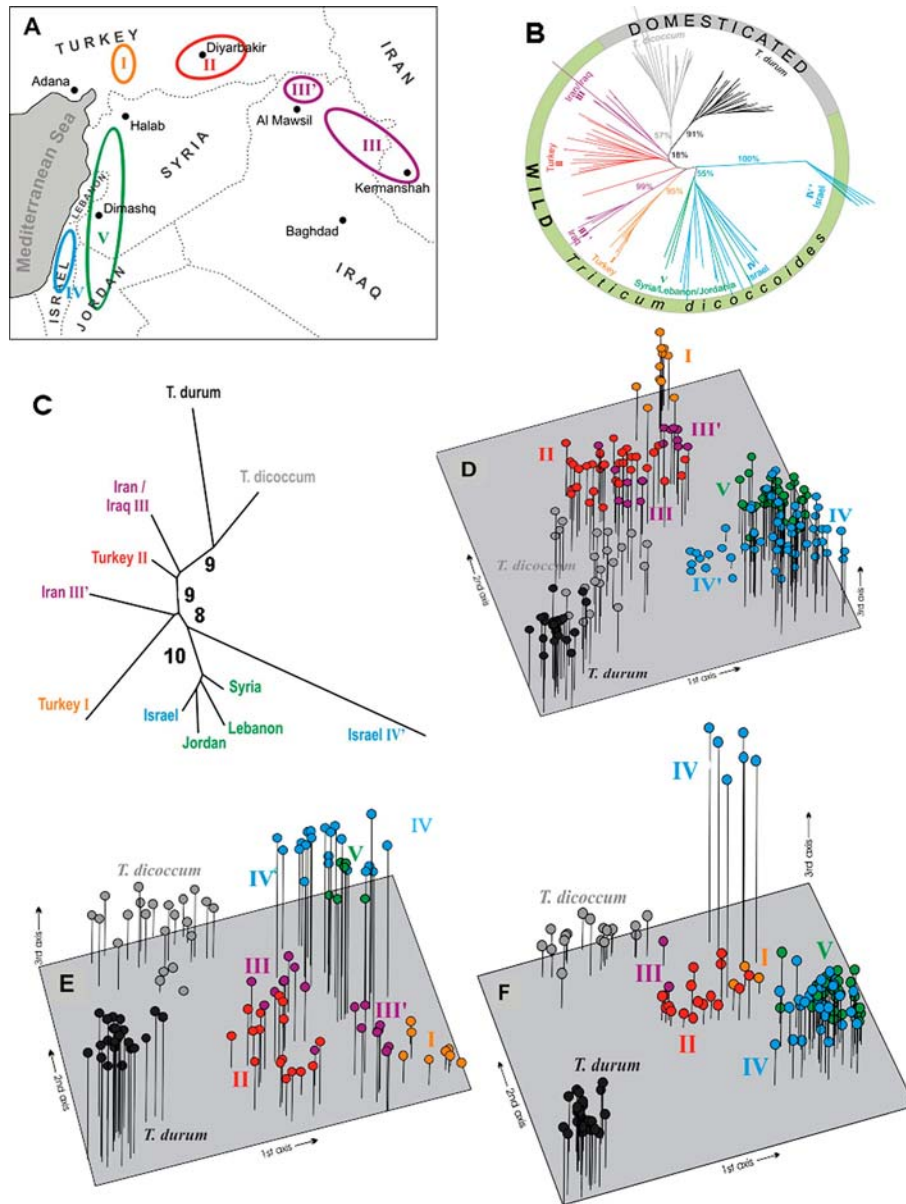
Fig. 1 a Sampling locations of *Triticum dicoccoides* wild lines in the Fertile Crescent. **b** Genetic relationship among *T. dicoccoides* (wild emmer), *T. dicoccum* (hulled emmer) and *T. durum* (free-threshing hard wheat). AFLP phylogeny of 45 domesticated tetraploid wheats (23 *T. dicoccum*, 22 free-threshing *T. durum* varieties) and 88 wild emmer lines (69 from Mori et al. 2003 and 19 from A. Karagoz and B. Gill) with different sampling locations (Turkey I, Kartal populations; Turkey II, mainly Karakadag populations; Iraq-Iran III and III', as in **a**; Israel IV as in **a** and IV' as in Table 1; Syrian, Jordanian and Lebanese lines are included in group V). The neighbour-joining tree (Saitou and Nei, 1987) of Jaccard (1908) genetic distances is shown. Domesticated *T. dicoccum* lines were from Turkey, Rumania, Iran, India, German, and Italy. The hard wheat (*T. durum*) lines were from France, Palestine, Italy, Tunisia, Jordan, Cyprus, Spain, Syria, Greece, Ukraine, Tajikistan, and Mexico. **c** Genetic relationships of domesticated tetraploid wheats with wild *T. dicoccoides* groups sampled in different regions of the Fertile Crescent with origins already specified in **b**. Using programs of the PHYLIP package (1993), 10 independent trees were constructed as described (Heun et al. 1997; Badr et al. 2000) using CONTML (topology shown; Felsenstein 1981) and distance matrix methods (Fitch and Margoliash 1967; Saitou and Nei 1987) employing various measures of genetic distance (Cavalli-Sforza and Edwards 1967; Nei

1972; Wright 1978; Reynolds et al. 1983) calculated from AFLP allele frequencies between groups of lines from the geographic regions indicated. The number of tree-buiding methods generating the same topologies is indicated at branches. **d, e** Results of the principal coordinates analysis of AFLP data used for the phylogenetic reconstructions of domesticated to wild lines relationships in tetraploid wheats. The analyses were carried out using the computer package NTSYSpc (Rohlf 1998). The pairwise genetic distance were calculated according to the algorithm of Jaccard (1908).**d** Principal coordinates analysis of the data obtained by merging the AFLP fingerprints of the lines of Ozkan et al. (2002), of the Mori et al. (2003) and of the Gill and Karagoz collections. The first 3 axes explain 22.2% of the total variability. **e** The AFLP database concerns the 55 domesticated lines specified in **b**, the 69 wild lines studied by Mori et al. (2003) and the 7 and 12 lines obtained from Dr. B. Gill and Dr. A. Karagoz (some lines were duplicated in the 3 collections, therefore the actual results of the analysis reports the positions of 72 lines).The first three axes explain 28.5% of the total variability. **f** A total of 131 domesticated and wild accessions considered by Ozkan et al. (2002) were the basis for this figure. Group III' and IV' wild lines were not represented in this set of materials. The first three axes explain 24.8% of the total variability

method based on Jaccard distances was used as well as several other methods of tree building and genetic distance calculations (citations above).

Dates reported in the Introduction and the Discussion of this paper are calibrated years before present (cal BP). They refer to 14C dates that were transformed into calendar years of the absolute dendrochronological record using the data provided by Zohary and Hopf (2000) and by Moore at al. (2000) and were cross-checked for consistency with the data of Gopher et al. (2002) and Maier (1996).

Results

An attempt to assemble a large *T. dicoccoides* collection

Precisely where wild *T. dicoccoides* grows and how complete are the existing collections is still a matter of discussion. The distribution of *T. dicoccoides* (Zohary and Hopf 2000) includes the western Fertile Crescent, southeastern Turkey, and areas in eastern Iran and Iraq. From southeastern Turkey to Iran and Iraq the species is progressively substituted by the wild tetraploid wheat *T. araraticum* having the AAGG genome (Johnson 1975). *T. dicoccoides* accessions have been only rarely collected in Iran or Iraq, thus supporting Johnson's conclusion. Stands of *T. dicoccoides* grow on the rocky slopes of the Karacadag Mountains (Harlan and Zohary 1966; Johnson 1975).

The study of Ozkan et al. (2002) has shown that in spite of their morphological similarity, *T. dicoccoides* and *T. araraticum* are easily distinguished at the molecular level. In the current study, the materials analyzed by Ozkan et al. (2002) were integrated with the accessions analyzed by Mori et al. (2003). Figure 1a summarizes, in a graphical representation, the *T. dicoccoides* geographical distribution based on the sampling locations of the wild lines analyzed for their species assignment using molecular techniques (Ozkan et al. 2002; Mori et al. 2003). Wild lines are included in groups I to V (including groups III' and IV', therefore defined also according to the phylogenies described in Fig. 1, but mainly based on geographical criteria). The figure highlights two aspects: (1) a number of wild lines have also been collected in the eastern part of the Fertile Crescent, and (2) the Kartal Mountains in southeast Turkey host rich natural stands of the species.

AFLP fingerprinting of the wild *T. dicoccoides* accessions studied by Mori et al. (2003)

The collection of wild *T. dicoccoides* lines (69 accessions) of Mori et al. (2003) was considered in this study, and 19 additional lines were obtained from Dr. B. Gill and Dr. A. Karagoz. Molecular fingerprinting was based on the same AFLP primer combinations as those used by Ozkan et al. (2002). In total, 169 polymorphic AFLP loci were scored for the presence or absence of amplified DNA fragments. The database was integrated with the AFLP data of the domesticated *T. dicoccum* and *T. durum* lines studied by Ozkan et al. (2002). The data were analyzed with several methods for genetic distance calculation and phylogenetic tree building. The topologies of the trees obtained with the different methods were consistently similar. A representative NJ tree is reported in Fig. 1b. As expected from previous experiments, the domesticated accessions of both *T. dicoccum* and *T. durum* have a lineage in common, but the two taxa diverge quite early after splitting from the wild lineage. Turkish wild accessions of group II (Karacadag Range) are, on average, highly related to the domesticated landraces. An interesting, and in part new, finding is that a small group of 8 lines from Iran and Iraq also appear to be genetically very similar to the domesticated gene pool. Turkish lines from the Kartal Mountain (group I), although more related to *T. dicoccum* and *T. durum* than the wild lines of the western Fertile Crescent, are not as highly related to the domesticated gene pool as the wild Karacadag and the Iraq-Iran lines. The Israeli, Jordanian, Lebanese and Syrian wild lines (groups IV, IV' and V) are clearly separated from the other wild *T. dicoccoides* accessions and are genetically far less related to the domesticated lines (Fig. 1b). As in Ozkan et al. (2002), the Israeli *T. dicoccoides* lines are split in two groups (IV and IV' in the figure).

The same conclusions can be drawn from Fig. 1c, which shows a consensus tree based on 10 different phylogenetic trees considering the relative frequency of the AFLP marker alleles in groups of wild lines (including Turkey I and II, Iran-Iraq III, Iran III', the group IV Israeli lines, Israel IV', the Jordanian, Syrian and Lebanese lines of group V) and in the two groups of domesticated genotypes.

Combining AFLP data from the two collections

Figures 1d, 1e and 1f reconsider the results of the same AFLP database used in Fig. 1b and 1c to illustrate the wild to domesticated genetic relationships. A different type of approach, the method known as principal coordinate (PC) analysis (Rohlf 1998), was used. The existence of different degrees of correlation among the allelic frequencies between pairs of genetic populations allows the extraction of successive components of total variability as axes along which the populations distribute. The similarity between the results of a phylogenetic tree and of a PC analysis should be close (Cavalli-Sforza and Piazza 1975). In PC analysis the first principal components frequently reflect the geographic distribution of the sample sites, therefore with this method genetic and geographical distances should be correlated and allow the alignment of the PC map with a geographic map. This is why PC analysis has been extensively used to follow the geographical distribution of human genes (Cavalli-Sforza et al. 1994). With the visual

limitations imposed by the need to consider several two-dimensional maps to correlate all PC components extracted from a set of data, the analysis is useful to assign a geographical origin to crops, their mode of domestication included. PC maps supplement and uphold phylogenetic trees by presenting data in a different form; thus, when PC analysis produces the same information as phylogenetic trees, the overall result is highly reliable (Cavalli-Sforza et al. 1994). This is the case of the tetraploid wheat domestication data (Fig. 1d, 1e, 1f); the PC results support the findings reported in Fig. 1b and 1c. The field defined by the first three principal components suggests four major observations: (1) the relative topologies of domesticated lines and of their related groups of wild lines match those reported in the two trees of Fig. 1b and c, (2) the domesticated groups consisting of *T. dicoccum* and *T. durum* lines occupy distinct positions in the principal coordinate space, (3) the *T. dicoccum* landraces seem more related to the wild lines than those of *T. durum*, and (4) the wild lines of *T. dicoccoides* most related to both groups of domesticated tetraploid wheats were sampled in southeastern Turkey and in Iran-Iraq. All analyses carried out on the AFLP database of tetraploid wheats, while supportive of a common geographical origin of the hulled and free-threshing domesticated forms, do not exclude the possibility of a diphyletic domestication.

When the old set of AFLP data from Ozkan et al. (2002) was analyzed by the PC method (Fig. 1f) and compared to the results of Fig. 1e concerning the wild lines from Mori et al. (2003) and from B. Gill and A. Karagoz (see Materials and Methods), a very similar picture emerged, a finding strongly supporting the per se value of each of the two datasets. In Fig. 1f the Israel lines occupy two different topologies, a situation already evident from the work of Ozkan et al. (2002).

More information can, nevertheless, be extracted from the two AFLP databases when they are considered together and analyzed by the PC procedure (Fig. 1d). The first two principal components are extremely effective in dividing the *T. dicoccoides* lines of the western Fertile Crescent (groups IV, IV' and V) from those of the central-eastern part (groups I, II, III and III'). Moreover, within the large central-eastern group (left part of the figure), the close relationships between the Turkish Karacadag lines (group II) and the domesticated accessions is evident. A looser but still significant genetic relatedness appears to characterize the genetic relationships among Turkish Kartal lines (group I) and the domesticated varieties.

Discussion

The publication of the paper by Heun et al. (1997) stimulated a discussion of southeastern Turkey as the place of origin of Western agriculture (Nesbitt and Samuel 1998). Later Lev-Yadun et al. (2000) summa-

rized the evidence that the ranges of wild progenitors of cereals and legumes intersect in southeastern Turkey, circumscribing a core area that includes Karacadag, considered the cradle of agricultural innovation, a conclusion supported by the excavation of Neolithic sites in the core region at 10,500 cal BP (Lev-Yadun et al. 2000). The domestication of *T. dicoccoides* was traced back to the same core area by Ozkan et al. (2002) and was reported as monophyletic. In a recent comment by Salamini et al. (2004), the origin of domesticated tetraploid genotypes has been also referred to as possibly diphyletic, with two major phylogenetic tree topologies occupied by *T. durum* (free-threshing) and *T. dicoccum* (hulled) lines.

A diphyletic domestication was the conclusion made by Mori et al. (2003) who, based on chloroplast DNA (cpDNA) fingerprinting, have found two major organelle DNA lineages of domesticated emmer. When comparing the molecular phylogenies obtained from sets of polymorphisms recorded for different subcellular genomes, conclusions may differ due to variations in the rate of genome evolution. For plants, this is the case for comparisons involving cpDNA data and those obtained from nuclear markers. On average, the rate of variation observed in nuclear genomes is three times higher than for chloroplast genomes (Wolfe et al. 1987; Gaut et al. 1996; McDade et al. 2000).

Despite this, it was nevertheless interesting that the extensive investigation of Mori et al. (2003) did not reveal more that two cpDNA haplotypes which were selected during tetraploid wheat domestication. The first cpDNA lineage of Mori et al. (2003), consisting of a large group of haplotypes slightly different at some of several DNA sites, has counterparts in wild populations of the Kartal Dagi site, which is located at the border of the core area of Lev-Yadun et al. (2000). The most common group I haplotype variant (haplotype 10) is present in 39.6% of domesticated and in 4.0% of wild *T. dicoccoides* lines. The authors have, in addition, established that in the wild *T. dicoccoides* populations the second major chloroplast lineage of domesticated tetraploids is represented only by closely related haplotypes which are scattered at low frequency along the Fertile Crescent. This second major lineage is present only in 7.3% of the domesticated lines studied.

The research presented in this paper was mainly stimulated by the desire to compare the phylogenetic information from different sources. This is the reason why we used nuclear AFLP markers to analyze the same wild populations considered by Mori et al. (2003) and compared the results to those of Ozkan et al. (2002). The data reported in Fig. 1b justify this analysis; the two nuclear AFLP databases establish that two very different genetic taxa of *T. dicoccoides* exist, the western one colonizing primary habitats in Israel, Syria, Lebanon and Jordan, and the central-eastern race, frequently sampled in Turkey, and rarely in Iran and Iraq. Several phytogeographers have already reported the existence of two different *T. dicoccoides* races (Sachs

1953; Harlan and Zohary 1966; Rao and Smith 1968). However, only the central-eastern race played a role as progenitor of the domesticated germplasm (Fig 1c; Mori et al. 2003; Ozkan et al. 2002; Salamini et al. 2004). Moreover, the cumulative results from the two collections suggest that the Turkish Karacadag population, intermixed with some Iraq-Iran lines, has a topology consistent with that of the putative progenitor of domesticated genotypes. The Turkish Kartal population, belonging genetically to the central-eastern T. dicoccoides race, seems less related at the nuclear DNA level to the domesticated gene pool.

A general agreement on tetraploid wheat domestication emerges from the results presented in this paper, as well as from those reported by Mori et al. (2003) and Ozkan et al. (2002). A discrepancy is nevertheless evident on the local geographical scale, reflecting the difficulty of studying events which took place in the Neolithic; the chloroplast DNA data indicate the Kartal mountains while AFLP fingerprinting points to the Karacadag Range as the putative site of tetraploid wheat domestication. Similar discrepancies are usually noted when taxa are compared based on different single gene trees. In a well-presented review dealing with soybean interspecific differences, Doyle et al. (2003) introduce such a problem for cpDNA and nuclear DNA phylogenies, explaining that the incongruence of results may either depend on ancient introgressions or may be due to stochastic sorting of ancestral lineages. In our case, similar events have a certain probability of having played such a role since T. dicoccoides, although mainly self-fertilizing, also admits intra-specific crosses (Zohary and Hopf 2000). Thus, post-domestication introgression of wild germplasm into the domesticated gene pool, mediated by crosses in which the wild lines have accepted foreign pollen, may have resulted in the spreading of wild cpDNA haplotypes in domesticated lines. However, a diphyletic domestication is also supported by the data of Mori et al. (2003) and not only by the AFLP results; this should exclude a significant wild gene pool introgression into domesticated varieties.The detection of the opposite event, an introgression of domesticated germplasm into wild populations, was outside of the scope of this paper. However, the Mori et al. (2003) cpDNA haplotype 10 is a candidate for such a possibility, being very frequent in domesticated lines and rare in the wild gene pool.

An attempt to reconcile the T. dicoccoides nuclear and chloroplast DNA data may also be that the Karacadag region hosted representatives of the chloroplast group I haplotypes. As a matter of fact, one T. dicoccoides accession was indentified as carrying a type I chloroplast haplotype molecularly close to the haplotype 10 of Mori et al. (2003), which is largely present in domesticated accessions. However, the consideration of additional 48 wild lines from the Karacadag Mountains led the same authors to conclude in favour of the Kartal Mountains as source of the domesticated germplasm.

When tested with nuclear DNA markers, the same T. dicoccoides accessions were less related to the cultivated gene pool than the Karacadag lines. This may result from incomplete information concerning the chloroplast DNA data; in the analysis only one free-threshing domesticated genotype was tested. If a larger collection of T. durum was to be analyzed, the results may provide evidence for the existence of more free-threshing specific chloroplast DNA haplotypes. Moreover, in the PC analysis presented in Fig. 1d, both domesticated sets of lines are more related to Karacadag than Kartal populations and thus we return to the belief that while the core area of Lev-Yadun et al. (2000) has played a clear role in the domestication of tetraploid wheat, the identification of a more precise location(s) within this area is still difficult. Factors contributing to keep this problem open are discussed in Salamini et al. (2002) and Mori et al. (2003).

References

Badr A, Muller K, Schafer-Pregl R, El Rabey H, Effgen S, Ibrahim HH, Pozzi C, Rohde W, Salamini F (2000) On the origin and domestication history of barley. Mol Biol Evol 17:499–510

Bar-Yosef O (1998) The natufian culture in the Levant, threshold of the origin of agriculture. Evol Anthropol 6:159–177

Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis: models and estimation procedures. Evolution 32:550–570

Cavalli-Sforza LL, Piazza A (1975) Analysis of evolution: evolutionary rates, independence and treeness. Theor Popul Biol 8:127–165

Cavalli-Sforza LL, Menozzi P, Piazza A (1994) The history and geography of human genes. Princeton University Press, Princeton

Doyle JJ, Lucknow M (2003) The rest of the iceberg: legume diversity and evolution in a phylogenetic context. Plant Physiol 131:900–910

Dvorak J, Di Terlizzi P, Zhang HB, Resta P (1993) The evolution of polyploid wheats: identification of the A genome donor species. Genome 36:21–31

Dvorak J, Luo MC, Yang ZL, Zhang HB (1998) The structure of the Aegilops tauschii gene pool and the evolution of hexaploid wheat. Theor Appl Genet 67:657–670

Felsenstein J (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. J Mol Evol 17:368–371

Felsenstein J (1989) PHYLIP - Phylogeny Inference Package (v3.2). Cladistics 5:164–166

Fitch WM, Margoliash E (1967) Construction of phylogenetic trees. Science 155:279–284

Gaut BS, Morton BR, McCaig BC, Clegg MT (1996) Substitution rate comparisons between grasses and palms: synonymous rate differences at the nuclear gene Adh parallel rate differences at the plastid gene rbcL. Proc Natl Acad Sci USA 93:10274–10279

Gopher A, Abbo S, Lev-Yadun S (2002) The "when", the "where" and the "why" of the Neolithic revolution in the Levant. Doc Praehist 28:1–14

Harlan JR, Zohary D (1966) Distribution of wild wheats and barley. Science 153:1074–1080

Heun M, Schaefer-Pregl R, Klawan D, Castagna R, Accerbi M, Borghi B, Salamini F (1997) Site of Einkorn wheat domestication identified by DNA fingerprinting. Science 278:1312–1314

Jaccard P (1908) Nouvelle recherches sur la distribution florale. Bull Soc Vaud Sci Nat 44:223–270

Johnson BL (1975) Identification of the apparent B-genome donor of wheat. Can J Genet Cytol 17:21–39

Kislev ME (1980) *Triticum parvicoccum* sp. nov., the oldest naked wheat. Israel J Bot 28:95–107

Lev-Yadun S, Gopher A, Abbo S (2000) The cradle of agriculture. Science 288:1602–2603

Maier U (1996) Morphological studies of free-threshing wheat ears from a Neolithic site in southwest Germany, and the history of naked wheats. Veget Hist Archaeobot 5:39–55

McDade LA, Masta SE, Moody ML, Waters E (2000) Phylogenetic relationship among Acanthaceae: evidence from two genomes. Syst Bot 25:106–121

Moore AMT, Hillman GC, Legge AJ (2000) Village on the Euphrates. Oxford University Press, Oxford

Mori N, Ishii T, Ishido T, Hirosawa S, Watatani H, Kawahara T, Nesbitt M, Belay G, Takumi S, Ogihara Y, Nakamura C (2003) Origin of domesticated emmer and common wheat inferred from chloroplast DNA fingerprinting. 10th International Wheat Genetics Symposium, 1–6 September 2003, Paestum, Italy, pp 25–28

Nei M (1972) Genetic distance between populations. Am Naturalist 106:283–292

Nesbitt M, Samuel D (1996) From staple crop to extinction? The archaeology and history of the hulled wheats. In: Padulosi S, Hammer K, Heller J (eds) Hulled wheats. Proceedings of the 1st international workshop on hulled wheats, pp 41–100

Nesbitt M, Samuel D (1998) Wheat domestication: archaeological evidence. Science 279:1433

Ozkan H, Brandolini A, Schafer-Pregl R, Salamini F (2002) AFLP analysis of a collection of tetraploid wheats indicates the origin of emmer and hard wheat domestication in southeast Turkey. Mol Biol Evol 19:1797–1801

Rao PS, Smith LE (1968) Studies with Israel and Turkish accessions of *Triticum* L. emend. Var. *dicoccoides* (KORN) BOWDEN. Wheat Inform Serv 26:6–7

Reynolds JB, Weir BS, Cockerham CC (1983) Estimation of the coancestry coefficient: basis for a short-term genetic distance. Genetics 105:767–779

Rohlf FJ (1998) NTSYS-pc. Numerical taxonomy and multivariate analysis system. Applied Biostatistics, New York

Sachs L (1953) Chromosome behaviour in species hybrids with *Triticum timopheevii*. Heredity 7:49–58

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Salamini F, Ozkan H, Brandolini A, Schafer-Pregl R, Martin W (2002) Genetics and geography of wild cereal domestication in the near east. Nat Rev Genet 3:429–441

Salamini F, Heun M, Ozkan H, Wunder J (2004) On DNA markers, phylogenetic trees and mode of origin of crops. Genome 46:448–453

Van Zeist W, Bakker-Heeres JAH (1982) Archaeobotanical studies in the Levant. 1. Neolithic sites in the Damascus basin: Aswad, Ghoraifé, Ramad. Palaeohistoria 24:165–256

Wolfe KH, Li W-H, Sharp PM (1987) Rates of nucleotide substitution very greatly amomg plant mitochondrial, chloroplast, and nuclear DNAs. Proc Natl Acad Sci USA 84:9054–9058

Wright S (1978) Evolution and the genetics of populations. Variability within and among natural populations 4. University of Chicago Press, Chicago

Zohary D, Hopf M (2000) Domestication of plants in the old world. Oxford University Press, Oxford