



Article

Retrotransposon-Based Genetic Diversity Assessment in Wild Emmer Wheat (*Triticum turgidum* ssp. *dicoccoides*)

Anssi L. Vuorinen ¹, Ruslan Kalendar ^{2,3,*} , Tzion Fahima ⁴, Helena Korpelainen ², Eviatar Nevo ⁴ and Alan H. Schulman ^{5,6,*}

¹ Food Chemistry and Food Development, Department of Biochemistry, University of Turku, Vatselankatu 2, FI-20014 Turku, Finland; anssi.vuorinen@utu.fi

² Department of Agricultural Sciences, University of Helsinki, P.O. Box 27 (Latokartanonkaari 5), FI-00014 Helsinki, Finland; helena.korpelainen@helsinki.fi

³ RSE National Center for Biotechnology, 13/5 Kurgalzhynskoye Road, Astana 010000, Kazakhstan

⁴ Institute of Evolution, University of Haifa, Mt. Carmel, Haifa 3498838, Israel; tfahima@evo.haifa.ac.il (T.F.); nevo@research.haifa.ac.il (E.N.)

⁵ Institute of Biotechnology and Viikki Plant Science Centre, University of Helsinki, P.O. Box 65, FI-00014 Helsinki, Finland

⁶ Natural Resources Institute Finland (Luke), Latokartanonkaari 9, FI-00790 Helsinki, Finland

* Correspondence: ruslan.kalendar@mail.ru (R.K.); alan.schulman@helsinki.fi (A.H.S.); Tel.: +358-91-915-8869 (R.K.); +358-40-768-2242 (A.H.S.)

Received: 28 March 2018; Accepted: 26 June 2018; Published: 29 June 2018



Abstract: Wild emmer wheat (*Triticum turgidum* ssp. *dicoccoides*) is the wild ancestor of all cultivated tetraploid and hexaploid wheats and harbors a large amount of genetic diversity. This diversity is expected to display eco-geographical patterns of variation, conflating gene flow, and local adaptation. As self-replicating entities comprising the bulk of genomic DNA in wheat, retrotransposons are expected to create predominantly neutral variation via their propagation. Here, we have examined the genetic diversity of 1 Turkish and 14 Israeli populations of wild emmer wheat, based on the retrotransposon marker methods IRAP and REMAP. The level of genetic diversity we detected was in agreement with previous studies that were performed with a variety of marker systems assaying genes and other genomic components. The genetic distances failed to correlate with the geographical distances, suggesting local selection on geographically widespread haplotypes ('weak selection'). However, the proportion of polymorphic loci correlated with the population latitude, which may reflect the temperature and water availability cline. Genetic diversity correlated with longitude, the east being more montane. Principal component analyses on the marker data separated most of the populations.

Keywords: *Triticum turgidum* ssp. *dicoccoides*; wild emmer wheat; IRAP; REMAP; genetic diversity

1. Introduction

Wild emmer wheat (WEW), *T. turgidum* ssp. *dicoccoides* (Körn.) Thell., is the tetraploid progenitor of both cultivated tetraploid durum wheat and hexaploid common wheat. Accordingly, it has been used as a material for many genetic, ecological, physiological, and cytogenetic studies investigating the genus. Considerable genetic diversity has been reported among WEW populations, which harbor a rich source of many disease resistance alleles, agronomic traits, and environmental adaptations [1–3]. Wild emmer wheat grows in the Fertile Crescent, and populations are often isolated or semi-isolated from each other. However, those in Israel are in fairly close proximity.

Over its geographical range, the sharp environmental gradients of the Fertile Crescent, even on a local scale, impose differential selective pressures on plant populations. Gene flow and selection, as well as differences in rates of recombination along chromosomes, have created patterns of variation between accessions and populations of WEW [4], for which a reference sequence is now available [5]. Studies with various marker systems have revealed a high level of polymorphism within WEW, with the isolated marginal populations clearly differing from the others and genetic diversity patterns reflecting soil type at each location [1,6,7].

Retrotransposons, or Class I transposable elements [8], are both ubiquitous throughout the plant kingdom and comprise a large proportion of most plant genomes, in WEW reaching 70% [5]. They propagate by a “copy-and-paste” cycle of transcription, reverse transcription, and integration, rather than the “cut-and-paste” of DNA transposons [9]. Thus, retrotransposons remain a part of the chromosome and spread by producing daughter copies, which integrate at new loci, the precise insertion points of which are not likely to be used more than once. These properties make retrotransposons well-suited as genetic markers. The presence or absence of retrotransposon insertions can be surveyed by PCR-based genotyping, such as by the inter-retrotransposon amplified polymorphism (IRAP) and retrotransposon-microsatellite amplified polymorphism (REMAP) methods [10,11]. IRAP targets two retrotransposons sufficiently close to generate a PCR product between them, whereas REMAP produces products between retrotransposons and microsatellites.

The stability of retrotransposon integration sites allows them to be used as molecular markers in genetic map construction [12]. Retrotransposon markers have also been widely used to assess genetic diversity in many species [13–16]. Given that plant retrotransposons are stress activated [13,17,18], their role in generating eco-geographical patterns of genomic diversity is of particular interest. Such patterns have been seen for example in wild barley [19] and wild diploid wheats [20,21] in Israel. In the present study, we used IRAP and REMAP markers to analyze polymorphisms among WEW populations. The main goal was to find useful retrotransposon markers for diversity analyses and to assess the polymorphism of these markers among 150 genotypes originating from 15 populations, which all differ in their ecogeographical background. We also aimed to examine the adaptive nature of the pattern of retrotransposon integrations.

2. Materials and Methods

2.1. Plant Material and DNA Extraction

Plants from 1 Turkish and 14 Israeli populations of WEW, 8–10 genotypes per population, were collected randomly, at least 1 m apart. The populations and collection sites were described in earlier publications [22,23]. The material was stored in the cereal gene bank of the Institute of Evolution, University of Haifa, from where the genotypes for this study were obtained. The locations and ecogeographical backgrounds of the populations are described in Table 1 and Figure 1. The location of the Turkish Diyarbakir population is shown in Figure 1 of Nevo and Beiles [23]. The climate data for Israel is from the Atlas of Israel (1970), which reflect the era of collection, and from publications of the meteorological service of Israel.

Table 1. Geographical and climatic data for the 15 WEW populations from Israel and Turkey that were used in this study.

Population ^a	Ln	Lt	Al	Tm	Ta	Tj	Td	Tdd	Rn	Rd	Hu14	Huan	Dw	Trd	Ev	So	Rv	Rr
1 Mt. Hermon	35.73	33.30	1300	11	21	3	18	6	1400	66	48	60	60	0	150	1	30	20
7 Yehudiyya	35.70	32.93	200	19	27	11	16	12	550	47	42	58	58	100	160	5	38	25
8 Gamla	35.74	32.88	200	19	26	9	17	12	470	50	43	58	58	60	155	5	39	26
9 Rosh Pinna	35.52	32.95	700	18	25	9	16	10	697	50	48	58	50	35	150	1	35	22
11 Tabigha	35.53	32.90	0	24	32	15	17	10	436	45	45	57	58	120	160	5	39	25
16 Mt. Gilboa	35.42	32.50	150	21	28	12	16	12	400	44	43	58	40	160	165	1	34	24
17 Mt. Gerizim	35.28	32.20	800	17	23	8	15	9	700	47	45	60	42	0	155	1	38	25
18 Gitit	35.40	32.10	300	21	29	13	16	12	360	39	39	55	25	100	170	1	38	24

Table 1. Cont.

	Population ^a	Ln	Lt	Al	Tm	Ta	Tj	Td	Tdd	Rn	Rd	Hu14	Huan	Dw	Trd	Ev	So	Rv	Rr
19	Kokhav Hashahar	35.34	31.95	600	20	28	12	16	12	400	40	45	59	30	25	165	1	38	22
23	J'aba	35.08	31.67	660	17	25	9	15	9	500	41	49	62	57	30	155	1	35	21
24	Amirim	35.45	32.93	600	15	24	8	16	8	850	61	48	60	53	13	153	1	35	23
28	Bet-Oren	35.03	32.73	400	17	24	11	13	8	700	55	59	69	80	0	142	1	25	19
30	Bat-Shelomo	35.02	32.60	75	20	26	13	13	10	650	55	58	68	77	30	150	2	24	20
33	Givat Koach	34.92	32.03	75	20	26	12	14	12	540	46	50	64	65	105	160	1	32	26
36	W. Diyarbakir	39.63	37.89	850	13	27	2	25	-	546	65	-	46	-	-	5	-	-	-

^a Populations numbered as earlier [20]. Symbols for variables: Geographical: Ln, longitude (decimals); Lt, latitude (decimal); Al, Altitude (m). Temperature: Tm, mean annual temperature; Ta, mean August temperature; Tj, mean January temperature; Td, seasonal temperature difference; Tdd, day-night temperature difference; Trd, mean number of tropical days. Water availability: Rn, mean annual rainfall (mm); Rd, mean number of rainy days; Huan, mean annual humidity; Hu1, mean humidity at 14:00; Dw, mean number of dew nights in summer; Ev, mean annual evaporation; Rv, mean interannual variability of rainfall; Rr, mean relative variability of rainfall. Edaphic: So, soil type; 1, terra rossa; 2, rendzina; 5, basalt.

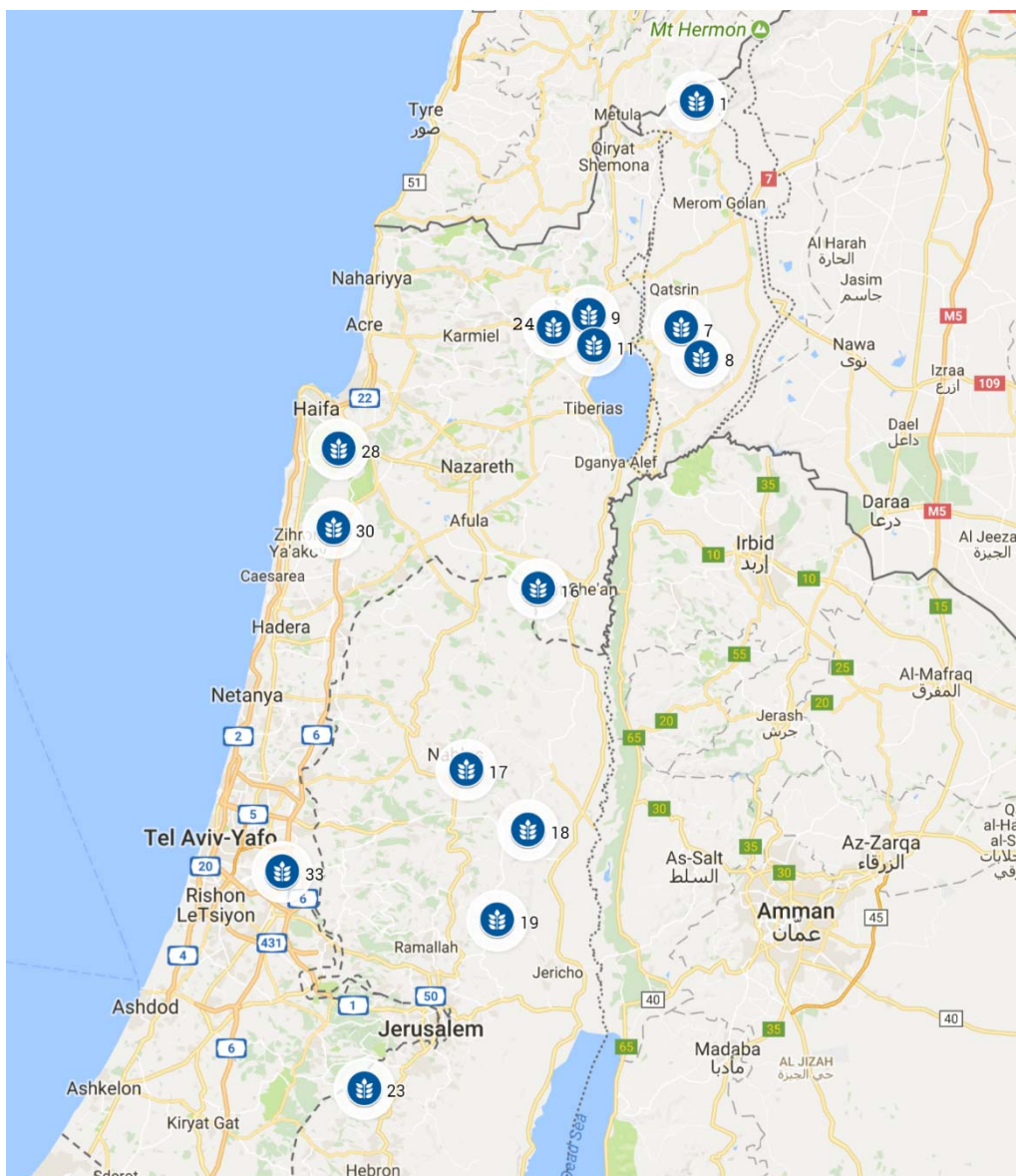


Figure 1. A map showing the geographic distribution of the 14 Israeli WEW populations. For names of the populations, see Table 1. For a map showing the Turkish West Diyarbakir population, see Nevo and Beiles [20].

One seed of each genotype was grown into a seven-day-old seedling, from which leaves were collected for preparation of DNA. DNA was isolated from 200 mg of fresh leaves using the CTAB extraction protocol with modifications (CTAB solution: 1.5% CTAB, 1.5 M NaCl, 20 mM Na₃EDTA, 0.1 M HEPES, pH ~5.3) as described (<http://primerdigital.com/dna.html>), with RNase A treatment. A detailed protocol for DNA isolation was deposited at protocols.io (DOI: 10.17504/protocols.io.mghc3t6). The DNA samples were diluted into 1× TE buffer. The DNA quality checked was electrophoretically, as well as spectrophotometrically with a Nanodrop apparatus (Thermo Fisher Scientific Inc., Waltham, MA, USA).

2.2. TE Sequence Sources and PCR Primer Design

To determine the inter-population diversity of seven LTR retrotransposons (*BARE1/WIS2*, *Wilma/Bagy2*, *Cereba*, *Nikita*, *Sukkula*, *Sabrina*, and *Daniela*), IRAP analyses of the populations was performed and the results compared. The retrotransposon sequences were taken from the TREP database (Table 2; <http://botserv2.uzh.ch/kelldata/trep-db/>) [8]. For a given family of retrotransposons, their LTRs show some sequence variability, but certain domains are relatively conserved. For each family, the sequence accessions were aligned and conservation assessed with the multiple alignment procedure of MULTALIN [24]. The conserved segments of the LTR of the retrotransposons were used for the design of PCR primers, which was carried out with the program FastPCR (<http://primerdigital.com/fastpcr.html>) [25–27]. We designed several primers for the LTRs of each retrotransposon to compare the efficiency and reproducibility of amplification. The sequences of the primers are shown in Table 2. None of the primers chosen form dimers and all showed high PCR efficiency and were very effective for IRAP fingerprinting. The chosen primers match motifs sufficiently conserved in the retrotransposons to allow amplification of almost all targets in the genome.

Table 2. Primers used for IRAP and REMAP.

Primer ID	Sequence (5'-3')	Retrotransposon LTR	T _m ^a , °C	CG ^b (%)	Linguistic Complexity (%) [26]
432	GATAGGGTCGCATCTTGGGCGTGAC	<i>Sukkula</i>	64.1	60.0	93
443	ACACACACACACACACT	(AC) ₉ T microsatellite	54.3	47.4	24
455	TTGAATTCTGCTACGTTCCCC	<i>WIS2/BARE1</i>	55.8	45.5	88
554	CCAACCTAGAGGCTTGCTAGGGAC	<i>WIS2/BARE1</i>	60.0	56.5	80
639	ACACACAAAGCATTCTCCGG	<i>Sabrina</i>	58.4	52.4	79
678	AAAGTTGTATCCGGGCGTTAC	<i>Sukkula</i>	57.8	50.0	88
679	GGGTCGCATATTGGGCGTGAC	<i>Sukkula</i>	62.0	61.9	87
692	GCGATTGCTAAGGCGCAACG	<i>Cereba</i>	61.1	60.0	89
734	TTCCCATGCGACGTTCCCCAAC	<i>WIS2/BARE1</i>	62.6	59.1	78
738	AATTTCTGCTACGTTCCCTAC	<i>WIS2/BARE1</i>	55.0	45.5	80
773	CCCTCTAGGCGACATCCACG	<i>Nikita</i>	60.6	65.0	89
2105	ACTCCATAGATGGATCTTGGTGA	<i>WIS2/BARE1</i>	54.6	43.5	88
2106	TAATTTCTGCAACGTTCCCCAACA	<i>WIS2/BARE1</i>	57.1	41.7	83
2107	AGCATGATGAAAATGGACGTATCA	<i>Wilma/Bagy2</i>	56.8	40.0	84
2108	AGAGCCTTCTGCTCCTCGTTGGGT	<i>Wilma/Bagy2</i>	63.4	58.3	83
2110	TCGCTGCGACTGCCCGTGACACA	<i>Daniela</i>	67.3	68.2	78
2113	TACGCATCCGTGCGCCCCGAAC	<i>Daniela</i>	66.6	68.2	90

^a T_m melting temperature, calculated according to [26]. ^b CG% percentage of C and G bases.

2.3. IRAP and REMAP Analysis

IRAP analysis was conducted according to Kalendar and Schulman [10,11] using 17 primers from the LTRs of the seven retrotransposons. The PCRs were performed in 25 µL reaction mixtures containing 25 ng DNA, 1× DreamTaq buffer, 200 mM dNTP, 400 nM primer and 1 U DreamTaq DNA Polymerase (Thermo Fisher Scientific). The amplifications were performed in the MasterCycler Gradient (Eppendorf AG, Hamburg, Germany). The PCR reaction program consisted of: 1 cycle at 95 °C, 3 min; 30 cycles of 95 °C for 20 s, 60 °C for 20 s, 72 °C for 90 s; a final elongation at 72 °C for 2 min.

For REMAP, we used one primer from the WIS2 LTR (738) combined with an ISSR primer (443), which is also shown in Table 2. Each LTR primer or primer combination was tested in PCR reactions using a genomic DNA mixture composed of equal amounts from all the accessions. The PCR products were separated by electrophoresis at 70 V for 8–12 h in a 1.3% agarose gel (Wide Range, SERVA Electrophoresis GmbH) with $0.5 \times$ TBE electrophoresis buffer. Gels were stained with EtBr and scanned using a FLA-5100 imaging system (Fuji Photo Film GmbH) with a resolution of 50 μm .

2.4. Data Analysis

From the IRAP and REMAP profiles, all distinct bands were scored as present (1) or absent (0) at each band position for each primer in all samples. Each PCR band was treated as a single locus. The presence or absence of a fragment of a given length was recorded in binary code. The sets that contained missing values were removed from the raw scored data sets. Monomorphic bands were scored and removed from the dataset before analysis for genetic distances.

The gels were scored for the presence or absence of totally 224 polymorphic bands for the samples. Based on the primary data, the level of genetic diversity as defined by Nei [28] was determined using the Arlequin software [29]. The primary genetic data were bootstrapped with SEQBOOT, after which the pairwise genetic distances were calculated using GENDIST (<http://www.bablokb.de/gendist/>). Both programs are from the PHYLIP software package (<http://evolution.gs.washington.edu/phylip.html>). The ability of IRAP and REMAP markers to reveal genetic relationships among all WEW accessions was evaluated phylogenetically by neighbor-joining (NJ), for which an algorithm was constructed using PAUP software [30]. Support for the tree was determined by performing 1000 bootstrap operations on the data set generated by distance analysis. To study the partition of IRAP genetic variation into among- and within-population variance components, the analysis of molecular variance (AMOVA) was conducted with the program GenAlex 6.5 [31]. The genetic distances among groups were analyzed with the Phi statistic (Φ_{st}). The number of permutations was set at 999 for AMOVA for a test of significance of the genetic distance among groups. Principal component analyses for the data matrix were run with the statistical software SAS (Cary, NC, USA) and the output was visualized on SigmaPlot (Systat Software Inc., San Jose, CA, USA). SAS was used to conduct a stepwise multiple regression with all eco-geographical variables (Table 1) as independent variables to find the best predictors of the proportion of polymorphic loci (P) and genetic diversity (H_e).

3. Results

Screening for primers resulted in the selection of 15 LTR primer pairs for IRAP and one primer pair for REMAP (LTR primer 738 with ISSR primer 443) (Figures 2 and 3). All selected LTR primers yielded 10 to 40 scorable bands, of which 10 to 20 were easily scorable. Altogether 224 polymorphic bands were scored.

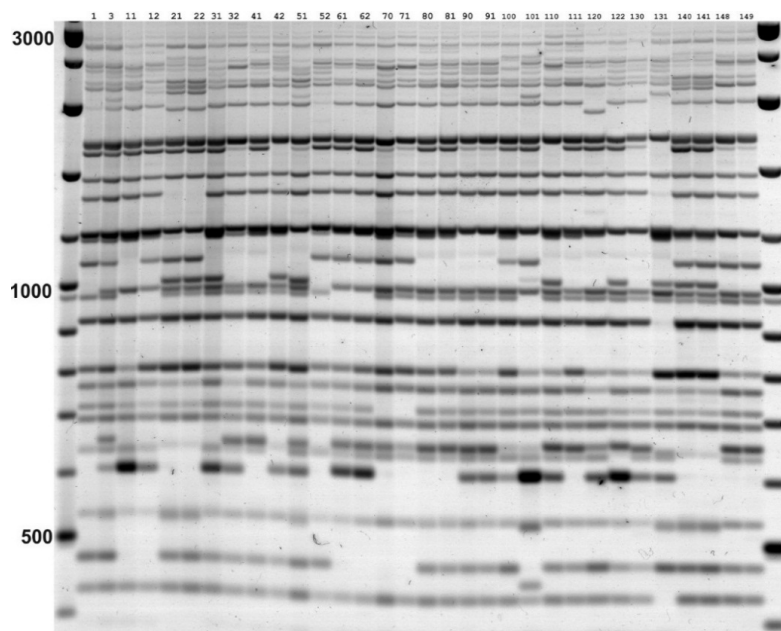


Figure 2. An IRAP gel produced with primer 679 (*Sukkula* LTR). DNA samples of the WEW populations: Mt. Hermon (1,3); Rosh-Pinna (11,12); Tabigha (21,22); Bat-Shelomo (31,32); Mt. Gilboa (41,42); Mt. Gerizim (51,52); Kokhav Hashahar (61,62); Amirim (70,71); Bet-Oren (80,81); Givat Koach (90,91); Gitit (100,101); J'aba (110,111); Gamla (120,122); West Diyarbakir (130,131); Yehudiyya (140,141,148,149). Size marker, both sides: Thermo Fisher Scientific GeneRuler DNA Ladder Mix (100 bp to 3000 bp), marked on the left in bp.

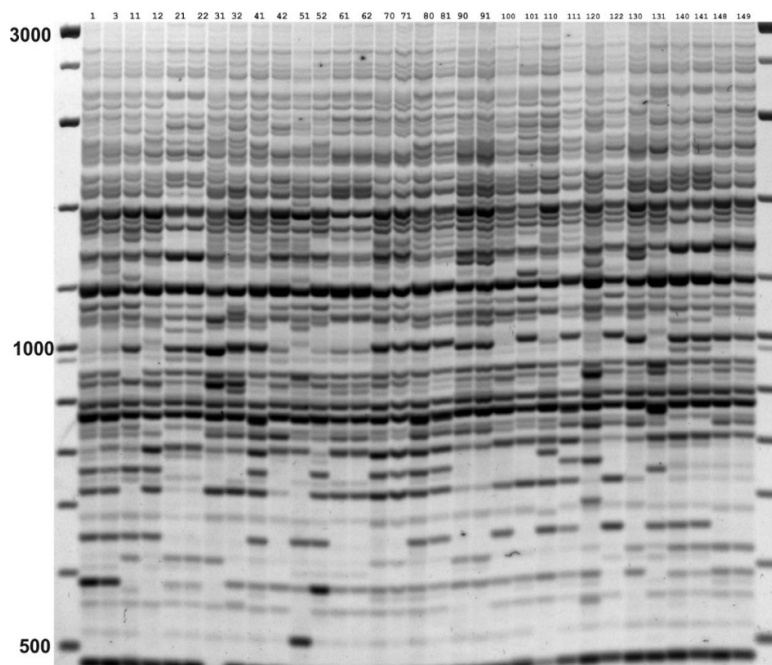


Figure 3. An IRAP gel produced with primer 2108 (*Wilma-Bagy2* LTR). DNA samples of the WEW populations: Mt. Hermon (1,3); Rosh-Pinna (11,12); Tabigha (21,22); Bat-Shelomo (31,32); Mt. Gilboa (41,42); Mt. Gerizim (51,52); Kokhav Hashahar (61,62); Amirim (70,71); Bet-Oren (80,81); Givat Koach (90,91); Gitit (100,101); J'aba (110,111); Gamla (120,122); West Diyarbakir (130,131); Yehudiyya (140,141,148,149). Size marker on both sides, Thermo Fisher Scientific GeneRuler DNA Ladder Mix (100 to 3000 bp), marked on the left in bp.

3.1. Genetic Diversity

The results concerning genetic diversity including the proportion of polymorphic loci ($P-5\%$) and genetic diversity (H_e , [26]) are shown in Table 3. The mean proportion of polymorphic loci and degree of genetic diversity respectively equaled 0.360 and 0.138 ± 0.063 (standard deviation), respectively. The H_e values varied between 0.002–0.204. The lowest H_e value (0.002 ± 0.002) was obtained in the small, isolated population of Bet-Oren and the highest value (0.204 ± 0.114) in Mt. Hermon, which belongs to the marginal steppe populations. Other high values were detected in the populations of Mt. Gilboa (0.201 ± 0.108), and Rosh Pinna (0.195 ± 0.105).

Table 3. Number of polymorphic loci and average gene diversities over the loci.

No. ^a	Population	All ^b	Polymorphic Loci	Fraction Loci Polymorphic	Diversity \pm SD
1	Mt. Hermon	197	93	0.47	0.204 ± 0.114
7	Yehudiyya	210	20	0.10	0.052 ± 0.030
8	Gamla	192	77	0.40	0.118 ± 0.065
9	Rosh Pinna	215	93	0.43	0.195 ± 0.105
11	Tabigha	221	48	0.22	0.109 ± 0.050
16	Mt. Gilboa	211	114	0.54	0.201 ± 0.108
17	Mt. Gerizim	207	104	0.50	0.186 ± 0.100
18	Gitit	214	102	0.48	0.189 ± 0.102
19	Kokhav Hashahar	201	78	0.39	0.157 ± 0.086
23	J'aba	207	96	0.46	0.182 ± 0.098
24	Amirim	217	58	0.27	0.112 ± 0.062
28	Bet-Oren	215	2	0.01	0.002 ± 0.002
30	Bat-Shelomo	175	82	0.47	0.186 ± 0.100
33	Givat Koach	204	48	0.24	0.063 ± 0.035
36	W. Diyarbakir, 22 km	202	77	0.38	0.118 ± 0.064
	Mean	205.9	72.8	0.36	0.138 ± 0.063

^a Population numbers are according to Nevo & Beiles [4]. ^b The number of loci included in the analyses of each population depends on the amount of missing data (loci included have less than 5% missing data).

3.2. Genetic Distance

Table 4 shows the genetic distances calculated as pair-wise comparisons for the averages of all populations. The greatest distance (0.9421) was found between Bet-Oren and Yehudiyya and the shortest (0.2185) between Mt. Gilboa and Mt. Gerizim. These results also show that Yehudiyya (distance ≥ 0.6391), Tabigha (distance ≥ 0.6068), and Bet-Oren (distance ≥ 0.5994) are all highly separated from the others as well as each other. The geographical distribution of the samples does not explain the genetic distance. For example, the genetic distance between West Diyarbakir in Turkey and many Israeli populations is shorter than the distance between some Israeli populations (Figure 4). This excludes the isolation-by-distance model.

Table 4. Genetic distances as pair-wise comparisons between the 15 WEW populations.

Population ^a	1	7	8	9	11	16	17	18	19	23	24	28	30	33	36
1 Mt. Hermon	0.0000														
7 Yehudiyya	0.6849	0.0000													
8 Gamla	0.4590	0.7727	0.0000												
9 Rosh Pinna	0.2902	0.6738	0.4605	0.0000											
11 Tabigha	0.6284	0.6430	0.7464	0.6514	0.0000										
16 Mt. Gilboa	0.2525	0.6391	0.4650	0.3549	0.6068	0.0000									
17 Mt. Gerizim	0.2593	0.6786	0.4660	0.3430	0.6443	0.2185	0.0000								
18 Gitit	0.3406	0.6641	0.5077	0.4063	0.6460	0.3132	0.3255	0.0000							
19 Kokhav H.	0.3724	0.7430	0.5438	0.3786	0.7187	0.3692	0.3726	0.4259	0.0000						

Table 4. Cont.

Population ^a	1	7	8	9	11	16	17	18	19	23	24	28	30	33	36
23 J'aba	0.3597	0.6724	0.4880	0.3772	0.6427	0.2721	0.2891	0.2749	0.3666	0.0000					
24 Amirim	0.4738	0.8041	0.5914	0.4443	0.7674	0.4472	0.5383	0.5010	0.4311	0.4948	0.0000				
28 Bet-Oren	0.6773	0.9421	0.8171	0.6909	0.8995	0.6215	0.6756	0.6961	0.7205	0.5994	0.8070	0.0000			
30 Bat-Shelomo	0.2992	0.6730	0.4367	0.3466	0.6384	0.2305	0.3114	0.3506	0.3832	0.3180	0.4143	0.6596	0.0000		
33 Givat Koach	0.5879	0.8388	0.7156	0.5573	0.8010	0.4794	0.5343	0.5499	0.5835	0.4604	0.6757	0.8894	0.5121	0.0000	
36 W. Diyarbakir	0.4226	0.7530	0.5860	0.5090	0.7254	0.4626	0.4720	0.5076	0.5534	0.4941	0.6163	0.8092	0.4620	0.7030	0.0000

^a Population numbers as earlier [20].

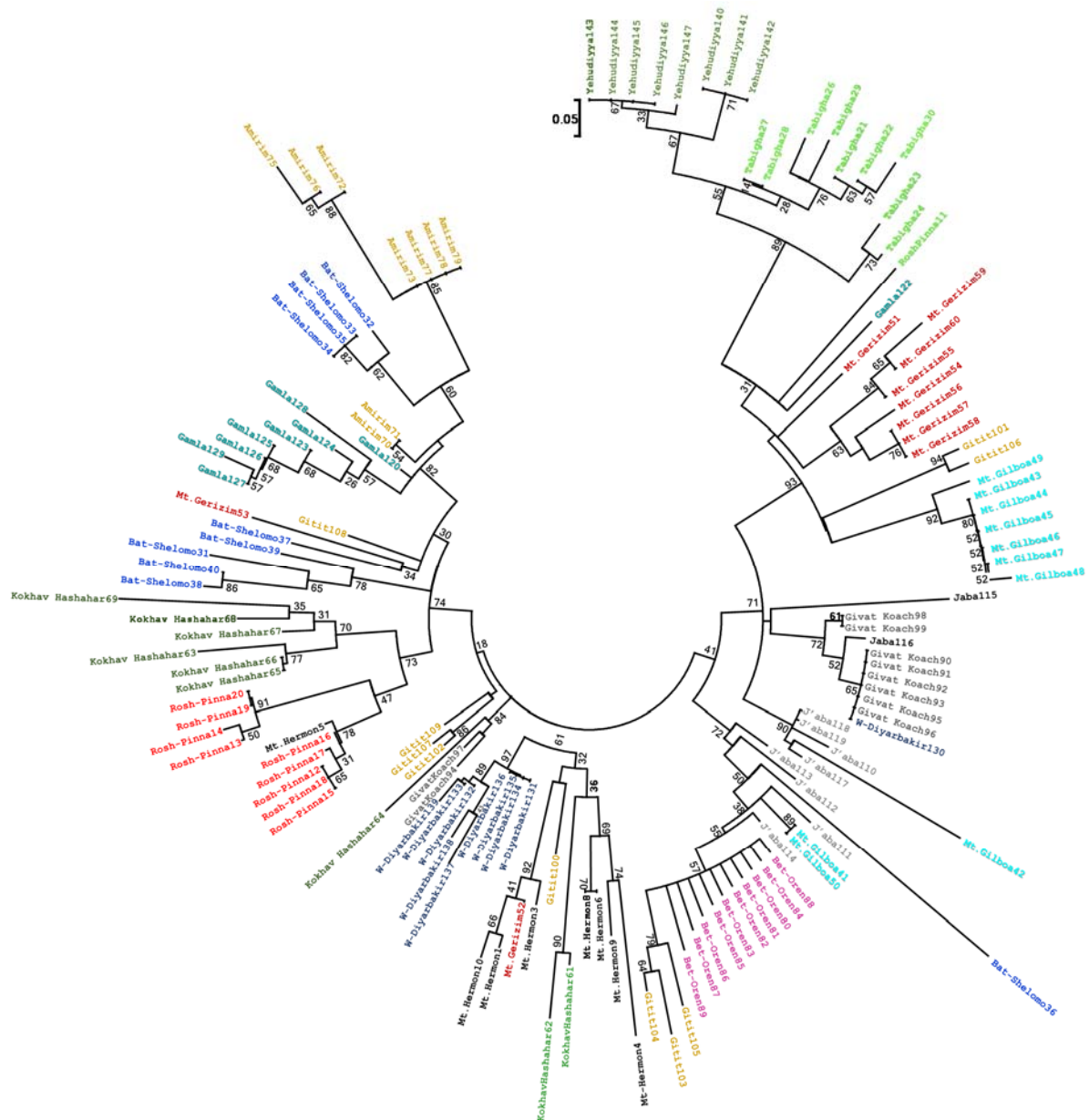


Figure 4. A phenogram obtained for the WEW populations. Bootstrap percentages are shown adjacent to the nodes. The population names appear at the branch ends, colored consistently with those in Figure 5. Populations are described in Table 1.

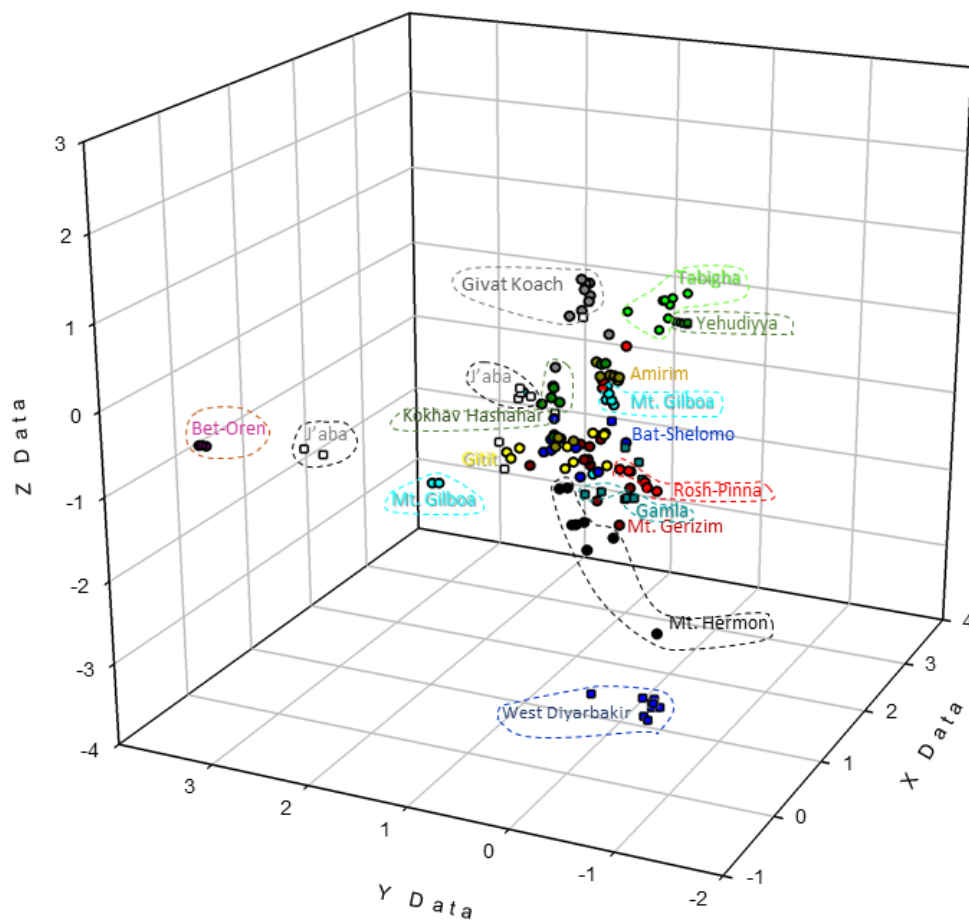


Figure 5. Principal component plot for the 15 WEW populations. The populations are labeled as follows: dark yellow circles, Amirim; dark blue circles, Bat-Shelomo; dark pink circles, Bet-Oren; red circles, Rosh-Pinna; light green circles, Tabigha; dark green boxes, Yehudiyya; dark cyan boxes, Gamla; black circles, Mt. Hermon; white boxes, J'aba; dark green circles, Kokhav Hashahar; cyan circles, Mt. Gilboa; dark red circles, Mt. Gerizim; light yellow circles, Gitit; grey circles, Givat Koach; dark blue boxes, West Diyarbakir. Accessions from individual populations that were clustered by the first three principal components are circled and labeled.

Principal component analysis (Figure 5) managed to separate most of the populations, with the relative positions consistent with the cladistics analysis (Figure 4). Yehudiyya and Tabigha form one group, Bet-Oren is a monomorphic group, West Diyarbakir separates as a group of its own, J'aba accessions are clustered, Mt. Gilboa and Mt. Gerizim form one group, and Amirim with Kokhav Hashahar separates clearly. The other accessions fail to separate clearly from the others. Multiple regression analysis showed a significant negative effect for latitude: $R^2 = 0.844$ ($p < 0.0001$) based on the proportion of polymorphic loci. Furthermore, longitude had a significant positive effect: $R^2 = 0.918$ ($p < 0.0001$) on genetic diversity. No other ecogeographical variables (Table 1) had a significant correlation with the genetic diversity data.

4. Discussion

Retrotransposon marker techniques, here IRAP and REMAP, shed light on the genetic differentiation in wild emmer wheat, *Triticum turgidum* ssp. *dicoccoides*. Retrotransposon families span boundaries between genera, so primers for retrotransposons originally isolated from barley, e.g., *BARE1*, work well as retrotransposon markers in wheat [21,32,33]. Retrotransposon polymorphism

captures the history of heritable integration events, which are driven by retrotransposon activation, and their subsequent fate in plant populations.

The same 15 populations that were used in this study had been used earlier in allozyme [22], RAPD [34], and microsatellite studies [5]. The genetic diversity of the marginal population of Bet-Oren was almost zero, as observed in earlier studies also [7,34], likely because of a small founding population. The level of observed diversity was similar to that observed using RAPDs, although the correlation was lowest of the marker methods. No relationship was found between the genetic distance (D) and the geographical distance for these populations, which may result from the high degree of polymorphism found within and between them. The same result was obtained using all other marker methods. Northern Israel is considered the center of diversity for wild emmer wheat [22,35], associated with the expectation of a wide range of different genotypes.

Interestingly, the genetic diversity found here using retrotransposon markers correlated well with allozyme data [22], which are derived from the coding region. One explanation for this observation is that the scored retrotransposon polymorphisms are directly selective. However, the sizes of the IRAP and REMAP amplification products (<3 kb) makes it unlikely that they contain genes. Moreover, retrotransposons in the Triticeae are usually clustered and most are far from genes, so that most but not all are likely to be neutral [36,37]; the chance of a marker band being a selective insertion is small. A second non-exclusive possibility is that the majority of retrotransposon polymorphisms are “passengers” associated with allelic diversity by linkage disequilibrium. However, data for WEW [38] indicates that linkage disequilibrium decays to background levels in about 1 cM, equivalent to about 1 Mbp in common wheat [39].

A third interpretation of the parallel between allozyme and retrotransposon diversity in WEW is that the environments promoting allelic diversification also activate retrotransposons, leading to new insertional polymorphisms. The genetic diversity detected using retrotransposon markers could not be predicted significantly by any variable other than latitude and longitude. The region has, however, a steep temperature and water availability cline with latitude, the south being progressively hotter and drier. The east of Israel is more montane; diversity is weakly (Pearson $R = 0.3557$) correlated with altitude. Moreover, when using the allele frequencies of all alleles in a population separately, several alleles were found to correlate with environmental variables (data not shown). Abiotic stresses, including UV and drought, are well known to activate retrotransposons [13,17,18]. Higher numbers of retrotransposons in wild barley at Evolution Canyon in Israel were seen at higher elevations, which are drier and receive more solar radiation [19].

A recent study looked at copy numbers of a TRIM, a SINE, and two superfamily *Gypsy* retrotransposons in WEW over a similar geographic range [40]. TRIMs are non-autonomous LTR retrotransposons driven by other (often unknown) autonomous elements, whereas SINEs are non-autonomous non-LTR retrotransposons, which propagate by a mechanism different from that of the LTR retrotransposons [8]. The retrotransposons in that study [40] differed among them as to which site showed the highest polymorphism and copy number. The *Latidu* element showed high numbers at Tabigha, consistent with the distinct position of that locale based on insertional polymorphisms from our results with seven different LTR retrotransposons (Figures 4 and 5). Both sets of data together emphasize the distinct dynamics of not only Class I, but also Class II, transposable elements in driving WEW genome dynamics in the Fertile Crescent. Retrotransposon insertions allowed the wild barley accessions at Evolution Canyon to be fully separated by canyon site [19]. Like wild barley, WEW is self-pollinating and can form several separate genotypes on a small area. Hence, retrotransposon markers both enable geographic resolution of WEW and other germplasm and give insight into local factors that may be driving genome evolution.

Author Contributions: Conceptualization, E.N. and A.H.S.; Methodology, R.K.; Software, A.L.V., A.H.S.; Validation, A.L.V., E.N. and A.H.S.; Formal Analysis, A.L.V., R.K.; Investigation, A.L.V., R.K.; Resources, T.F., E.N.; Writing-Original Draft Preparation, A.L.V., R.K., A.H.S.; Visualization, A.L.V., A.H.S.; Supervision, A.H.S., H.K., R.K.; Project Administration, A.H.S.; Funding Acquisition, A.H.S., R.K. and H.K.

Funding: This research was funded by Academy of Finland grant 48527, and the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan in the framework of program funding for research (grant AP05130266 and BR05236574).

Acknowledgments: This work was supported by Boreal Plant Breeding Ltd. The authors wish to thank Ursula Lönnqvist and Eeva-Marja Turkki for technical assistance. EN thanks the Ancell-Teicher Research Foundation for Genetics and Molecular Evolution for financial support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Brunazzi, A.; Scaglione, D.; Talini, R.F.; Miculan, M.; Magni, F.; Poland, J.; Enrico Pè, M.; Brandolini, A.; Dell'Acqua, M. Molecular diversity and landscape genomics of the crop wild relative *Triticum urartu* across the Fertile Crescent. *Plant J.* **2018**, *94*, 670–684. [[CrossRef](#)] [[PubMed](#)]
2. Merchuk-Ovnat, L.; Fahima, T.; Krugman, T.; Saranga, Y. Ancestral QTL alleles from wild emmer wheat improve grain yield, biomass and photosynthesis across environments in modern wheat. *Plant Sci.* **2016**, *251*, 23–34. [[CrossRef](#)] [[PubMed](#)]
3. Sela, H.; Loutre, C.; Keller, B.; Schulman, A.; Nevo, E.; Korol, A.; Fahima, T. Rapid linkage disequilibrium decay in the *Lr10* gene in wild emmer wheat (*Triticum dicoccoides*) populations. *Theor. Appl. Genet.* **2011**, *122*, 175–187. [[CrossRef](#)] [[PubMed](#)]
4. Akhunov, E.D.; Akhunova, A.R.; Anderson, O.D.; Anderson, J.A.; Blake, N.; Clegg, M.T.; Coleman-Derr, D.; Conley, E.J.; Crossman, C.C.; Deal, K.R.; et al. Nucleotide diversity maps reveal variation in diversity among wheat genomes and chromosomes. *BMC Genom.* **2010**, *11*, 702. [[CrossRef](#)] [[PubMed](#)]
5. Awni, R.; Nave, M.; Barad, O.; Baruch, K.; Twardziok, S.O.; Gundlach, H.; Hale, I.; Mascher, M.; Spannagl, M.; Wiebe, K.; et al. Wild emmer genome architecture and diversity elucidate wheat evolution and domestication. *Science* **2017**, *357*, 93–97. [[CrossRef](#)] [[PubMed](#)]
6. Volis, S.; Ormanbekova, D.; Yermekbayev, K.; Song, M.; Shulgina, I. Multi-approaches analysis reveals local adaptation in the emmer wheat (*Triticum dicoccoides*) at macro- but not micro-geographical scale. *PLoS ONE* **2015**, *10*, e0121153. [[CrossRef](#)] [[PubMed](#)]
7. Fahima, T.; Roder, M.S.; Wendehake, K.; Kirzhner, V.M.; Nevo, E. Microsatellite polymorphism in natural populations of wild emmer wheat, *Triticum dicoccoides*, in Israel. *Theor. Appl. Genet.* **2002**, *104*, 17–29. [[CrossRef](#)] [[PubMed](#)]
8. Wicker, T.; Sabot, F.; Hua-Van, A.; Bennetzen, J.L.; Capy, P.; Chalhou, B.; Flavell, A.; Leroy, P.; Morgante, M.; Panaud, O.; et al. A unified classification system for eukaryotic transposable elements. *Nat. Rev. Genet.* **2007**, *8*, 973–982. [[CrossRef](#)] [[PubMed](#)]
9. Schulman, A.H. Retrotransposon replication in plants. *Curr. Opin. Virol.* **2013**, *3*, 604–614. [[CrossRef](#)] [[PubMed](#)]
10. Kalendar, R.; Schulman, A.H. Transposon-based tagging: IRAP, REMAP, and iPBS. *Methods Mol. Biol.* **2014**, *1115*, 233–255. [[PubMed](#)]
11. Kalendar, R.; Schulman, A.H. IRAP and REMAP for retrotransposon-based genotyping and fingerprinting. *Nat. Protoc.* **2006**, *1*, 2478–2484. [[CrossRef](#)] [[PubMed](#)]
12. Boyko, E.; Kalendar, R.; Korzun, V.; Fellers, J.; Korol, A.; Schulman, A.H.; Gill, B.S. A high-density cytogenetic map of the *Aegilops tauschii* genome incorporating retrotransposons and defense-related genes: Insights into cereal chromosome structure and function. *Plant Mol. Biol.* **2002**, *48*, 767–789. [[CrossRef](#)] [[PubMed](#)]
13. Galindo-Gonzalez, L.; Mhiri, C.; Deyholos, M.K.; Grandbastien, M.A. LTR-retrotransposons in plants: Engines of evolution. *Gene* **2017**, *626*, 14–25. [[CrossRef](#)] [[PubMed](#)]
14. Kalendar, R.; Flavell, A.J.; Ellis, T.H.; Sjakste, T.; Moisy, C.; Schulman, A.H. Analysis of plant diversity with retrotransposon-based molecular markers. *Heredity* **2011**, *106*, 520–530. [[CrossRef](#)] [[PubMed](#)]
15. Kalendar, R.N.; Aizharkyn, K.S.; Khapilina, O.N.; Amenov, A.A.; Tagimanova, D.S. Plant diversity and transcriptional variability assessed by retrotransposon-based molecular markers. *Vavilov J. Genet. Breed.* **2017**, *21*, 128–134. [[CrossRef](#)]
16. Mandoulakani, B.A.; Yaniv, E.; Kalendar, R.; Raats, D.; Bariana, H.S.; Bihanta, M.R.; Schulman, A.H. Development of IRAP- and REMAP-derived SCAR markers for marker-assisted selection of the stripe rust

- resistance gene *Yr15* derived from wild emmer wheat. *Theor. Appl. Genet.* **2015**, *128*, 211–219. [[CrossRef](#)] [[PubMed](#)]
17. Grandbastien, M.A.; Audeon, C.; Bonnard, E.; Casacuberta, J.M.; Chalhouh, B.; Costa, A.P.; Le, Q.H.; Melayah, D.; Petit, M.; Poncet, C.; et al. Stress activation and genomic impact of *Tnt1* retrotransposons in Solanaceae. *Cytogenet. Genome Res.* **2005**, *110*, 229–241. [[CrossRef](#)] [[PubMed](#)]
 18. Jaaskelainen, M.; Chang, W.; Moisy, C.; Schulman, A.H. Retrotransposon BARE displays strong tissue-specific differences in expression. *New Phytol.* **2013**, *200*, 1000–1008. [[CrossRef](#)] [[PubMed](#)]
 19. Kalendar, R.; Tanskanen, J.; Immonen, S.; Nevo, E.; Schulman, A. Genome evolution of wild barley (*Hordeum spontaneum*) by BARE-1 retrotransposon dynamics in response to sharp microclimatic divergence. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 6603–6607. [[CrossRef](#)] [[PubMed](#)]
 20. Hosid, E.; Brodsky, L.; Kalendar, R.; Raskina, O.; Belyayev, A. Diversity of long terminal repeat retrotransposon genome distribution in natural populations of the wild diploid wheat *Aegilops speltoides*. *Genetics* **2012**, *190*, U263–U412. [[CrossRef](#)] [[PubMed](#)]
 21. Belyayev, A.; Kalendar, R.; Brodsky, L.; Nevo, E.; Schulman, A.H.; Raskina, O. Transposable elements in a marginal plant population: Temporal fluctuations provide new insights into genome evolution of wild diploid wheat. *Mob. DNA* **2010**, *1*. [[CrossRef](#)] [[PubMed](#)]
 22. Nevo, E.; Beiles, A. Genetic diversity of wild emmer wheat in Israel and Turkey : Structure, evolution, and application in breeding. *Theor. Appl. Genet.* **1989**, *77*, 421–455. [[CrossRef](#)] [[PubMed](#)]
 23. Nevo, E.; Beiles, A. Genetic Parallelism of Protein Polymorphism in Nature—Ecological test of the neutral theory of molecular evolution. *Biol. J. Linn. Soc.* **1988**, *35*, 229–245. [[CrossRef](#)]
 24. Corpet, F. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* **1988**, *16*, 10881–10890. [[CrossRef](#)] [[PubMed](#)]
 25. Kalendar, R.; Khassenov, B.; Ramanculov, E.; Samuilova, O.; Ivanov, K.I. FastPCR: An *in silico* tool for fast primer and probe design and advanced sequence analysis. *Genomics* **2017**, *109*, 312–319. [[CrossRef](#)] [[PubMed](#)]
 26. Kalendar, R.; Tselikh, T.V.; Khassenov, B.; Ramanculov, E.M. Introduction on using the FastPCR software and the related Java web tools for PCR and oligonucleotide assembly and analysis. *Methods Mol. Biol.* **2017**, *1620*, 33–64. [[PubMed](#)]
 27. Kalendar, R.; Muterko, A.; Shamekova, M.; Zhambakin, K. In Silico PCR tools for a fast primer, probe, and advanced searching. *Methods Mol. Biol.* **2017**, *1620*, 1–31. [[PubMed](#)]
 28. Marks, J. Molecular evolutionary genetics. *Am. J. Phys. Anthropol.* **1988**, *75*, 428–429. [[CrossRef](#)]
 29. Excoffier, L.; Laval, G.; Schneider, S. Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evol. Bioinform. Online* **2005**, *1*, 47–50. [[CrossRef](#)]
 30. Swofford, D.L. *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods)*; Version 4; Sinauer Associates: Sunderland, MA, USA, 1998.
 31. Peakall, R.; Smouse, P.E. GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research—An update. *Bioinformatics* **2012**, *28*, 2537–2539. [[CrossRef](#)] [[PubMed](#)]
 32. Vicent, C.; Jaaskelainen, M.; Kalendar, R.; Schulman, A. Active retrotransposons are a common feature of grass genomes. *Plant Physiol.* **2001**, *125*, 1283–1292. [[CrossRef](#)] [[PubMed](#)]
 33. Queen, R.A.; Gribbon, B.M.; James, C.; Jack, P.; Flavell, A.J. Retrotransposon-based molecular markers for linkage and genetic diversity analysis in wheat. *Mol. Genet. Genom.* **2004**, *271*, 91–97. [[CrossRef](#)] [[PubMed](#)]
 34. Fahima, T.; Sun, G.L.; Beharav, A.; Krugman, T.; Beiles, A.; Nevo, E. RAPD polymorphism of wild emmer wheat populations, *Triticum dicoccoides*, in Israel. *Theor. Appl. Genet.* **1999**, *98*, 434–447. [[CrossRef](#)]
 35. Ren, J.; Chen, L.; Sun, D.; You, F.M.; Wang, J.; Peng, Y.; Nevo, E.; Beiles, A.; Sun, D.; Luo, M.C.; et al. SNP-revealed genetic diversity in wild emmer wheat correlates with ecological factors. *BMC Evol. Biol.* **2013**, *13*, 169. [[CrossRef](#)] [[PubMed](#)]
 36. Anderson, O.D.; Rausch, C.; Moullet, O.; Lagudah, E.S. The wheat D-genome HMW-glutenin locus: BAC sequencing, gene distribution, and retrotransposon clusters. *Funct. Integr. Genom.* **2003**, *3*, 56–68.
 37. Wicker, T.; Schulman, A.H.; Tanskanen, J.; Spannagl, M.; Twardziok, S.; Mascher, M.; Springer, N.M.; Li, Q.; Waugh, R.; Li, C.; et al. The repetitive landscape of the 5100 Mbp barley genome. *Mob. DNA* **2017**, *8*, 22. [[CrossRef](#)] [[PubMed](#)]
 38. Sela, H.; Ezrati, S.; Ben-Yehuda, P.; Manisterski, J.; Akhunov, E.; Dvorak, J.; Breiman, A.; Korol, A. Linkage disequilibrium and association analysis of stripe rust resistance in wild emmer wheat (*Triticum turgidum* ssp. *dicoccoides*) population in Israel. *Theor. Appl. Genet.* **2014**, *127*, 2453–2463. [[CrossRef](#)] [[PubMed](#)]

39. Stein, N.; Feuillet, C.; Wicker, T.; Schlagenhauf, E.; Keller, B. Subgenome chromosome walking in wheat: A 450-kb physical contig in *Triticum monococcum* L. spans the *Lr10* resistance locus in hexaploid wheat (*Triticum aestivum* L.). *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 13436–13441. [[CrossRef](#)] [[PubMed](#)]
40. Domb, K.; Keidar, D.; Yaakov, B.; Khasdan, V.; Kashkush, K. Transposable elements generate population-specific insertional patterns and allelic variation in genes of wild emmer wheat (*Triticum turgidum* ssp. *dicoccoides*). *BMC Plant Biol.* **2017**, *17*, 175. [[CrossRef](#)] [[PubMed](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).