Research Article

Received: 15 February 2019 Accepted article published: 29 May 2019 Published online in Wiley Online Library:

(wileyonlinelibrary.com) DOI 10.1002/jsfa.9822

The Israeli–Palestinian wheat landraces collection: restoration and characterization of lost genetic diversity

Sivan Frankin,[a,b] Srinivas Kunta,[a,b] Shahal Abbo,[b] Hanan Sela,[c] Ben Z Goldberg,[d] David J Bonfil,[e] Avraham A Levy,[f] Naomi Avivi-Ragolsky,[f] Kamal Nashef,[a] Rajib Roychowdhury,[a] Tomer Faraj,[g] Dikla Lifshitz,[g] Einav Mayzlish-Gati[g] and Roi Ben-David[a]*

Abstract

BACKGROUND: For over a century, genetic diversity of wheat worldwide was eroded by continual selection for high yields and industrial demands. Wheat landraces cultivated in Israel and Palestine demonstrate high genetic diversity and a potentially wide repertoire of adaptive alleles. While most Israeli-Palestinian wheat landraces were lost in the transition to 'Green Revolution' semi-dwarf varieties, some germplasm collections made at the beginning of the 20th century survived in gene banks and private collections worldwide. However, fragmentation and poor conservation place this unique genetic resource at a high risk of genetic erosion. Herein, we describe a long-term initiative to restore, conserve, and characterize a collection of Israeli and Palestinian wheat landraces (IPLR).

RESULTS: We report on (i) the IPLR construction ($n = 932$), (ii) the historical and agronomic context to this collection, (iii) the characterization and assessment of the IPLR's genetic diversity, and (iv) a data comparison from two distinct subcollections within IPLR: a collection made by N. Vavilov in 1926 (IPLR-VIR) and a later one (1979–1981) made by Y. Mattatia (IPLR-M). Though conducted in the same eco-geographic space, these two collections were subjected to considerably different conservation pathways. IPLR-M, which underwent only one propagation cycle, demonstrated marked genetic and phenotypic variability (within and between accessions) in comparison with IPLR-VIR, which had been regularly regenerated over ~90 years.

CONCLUSION: We postulate that long-term *ex situ* conservation involving human and genotype × environment selection may significantly reduce accession heterogeneity and allelic diversity. Results are further discussed in a broader context of pre-breeding and conservation.

Supporting information may be found in the online version of this article.

Keywords: landraces; bread wheat; durum wheat; phenotypic diversity; genetic diversity; gene bank; conservation

INTRODUCTION

In the past decade, there has been increasing interest in wheat landraces as a genetic resource for improving crop yields and resilience, especially in the context of climate change.[1] There is also a growing interest in landraces as a source of beneficial nutritional traits and flavour repertoire,[2] for breeding applications,[3,4] and even for 'as is' landrace cultivation for niche markets. Wheat landraces cultivated in the Fertile Crescent since the time of domestication thousands of years ago[5] contain considerable genetic diversity in adaptive and important agronomical traits.[6] Despite worldwide awareness of the major role landraces play in preserving crop biodiversity, modernization processes[7] have led to landraces (across all crops) undergoing considerable genetic erosion, and in many cases extinction. In addition, as a consequence of conservation methods, the landraces that are found today in gene banks, and which are utilized in research and breeding, may differ considerably in agronomical and genetic aspects from their historical ancestors sampled decades ago. Traditionally,

* *Correspondence to: R Ben-David, Department of Vegetable and Field Crops, Institute of Plant Sciences, Agricultural Research Organization (ARO) – Volcani Center, 68 HaMaccabim Road, P.O.B. 15159, Rishon LeZion 7505101, Israel. E-mail: roib@volcani.agri.gov.il*

a *Department of Vegetables and Field Crop, Institute of Plant Sciences, Agricultural Research Organization, Volcani Center, Rishon LeZion, Israel*

b *The Robert H. Smith Institute of Plant Sciences and Genetics in Agriculture, The Hebrew University of Jerusalem, Rehovot, Israel*

c *Institute for Cereal Crops Improvement, Tel Aviv University, Tel Aviv, Israel*

d *Land of Wheat "Initiative", Israel*

e *Department of Vegetable and Field Crop Research, Institute of Plant Sciences, Agricultural Research Organization, Gilat Research Center, M.P. Negev, Israel*

f *Department of Plant & Environmental Sciences, Weizmann Institute of Science, Rehovot, Israel*

g *Israel Gene Bank, Agricultural Research Organization, Volcani Center, Rishon LeZion, Israel*

wheat landraces are dynamic, genetically diverse populations[8] containing multiple genotypes.[9] These landrace mixtures or 'meta-populations'[10] harboured buffering capacity that lowered a farmer's risk of failure and minimized yield fluctuation[10,11] and evolved in long-term equilibrium with abiotic and biotic stresses based on their ability to adapt locally. In the area known today as Israel–The Palestinian Authority, M. Jacubziner,[12] L. Pinner[13] and Y. Mattatia[14] all refer to the mixture of genotypes as a routine practice in wheat fields, reaching as high as seven different types in one field.[15] Throughout the 20th century in the Southern Levant, wheat landraces were collected and conserved *ex situ* in gene banks, preserving vital biodiversity, but simultaneously imposing a constant risk for conservation-influenced genetic bottlenecks caused by (i) collecting (intended and unintended human selection during collection), (ii) sample propagation (genetic drift brought about by the impact of genotype × environment, G × E, interactions on sequential propagation cycles),[11] (iii) human selection during propagation (gene bank curators often choose to remove admixture genotypes from the main type found in a sample), and (iv) ongoing sharing of germplasm between national or international institutes[10] that entails small samples, which can be minimized even further when plant protection restrictions limit the number of seeds (genotypes) from which an accession is regenerated. As a result, it is possible that landraces grown, studied, or used for breeding today differ significantly from the stocks that grew in the fields from which they were first collected in terms of sample heterogeneity as well as genetic and phenotypic diversity.

Wheat landraces in the Southern Levant present an extreme example of genetic erosion, and even extinction, of valuable exotic germplasm. In the late 19th century, there were dozens, if not hundreds, of traditional wheat landraces in this region. Since the beginning of the 20th century, the genetic diversity of wheat has been sharply eroded by the continual selection for industry needs as well as high and stable yields, a process that was completed with the transition to semi-dwarf 'Green Revolution' cultivars in the 1960s and 1970s.[16] Although many Israeli and Palestinian wheat landraces were completely lost, some of the representative varieties were sampled and collected throughout the 20th century by plant researchers, including A. Aaronsohn (1913), N.I. Vavilov (1926), L. Pinner (1929), Dickson (1930–1931) and later on by Y. Efrat (1967) and Y. Mattatia (1980). Though some of these collections were transferred to gene banks[7] and private collections worldwide, the majority of this material was not conserved in Israel or Palestine. The Israeli–Palestinian Landrace (IPLR) Wheat project was initiated to apply an integrated approach to restore, conserve, and study local traditional wheat germplasm originating mostly from the region corresponding to current-day Israel, the Palestinian authority, and Gaza Strip. Under the auspices of the project, approximately 1000 wheat accessions were identified, traced, acquired, and restored in the Israel Gene Bank (IGB) from 2015 to 2018. In this study, we: (i) report on the IPLR construction, including some details on the germplasm classification process; (ii) provide historical and agronomic context to the IPLR collection; (iii) characterize and assess its genetic diversity by using a high-throughput single-nucleotide polymorphism (SNP) assay to help determine whether these landraces hold a rich allelic reservoir compared with modern cultivars; (iv) compare data from two distinct sub-collections within the IPLR: the Vavilov collection made in 1926 (IPLR-VIR), one of the oldest and best conserved resources, and the relatively recent collection made by Y. Mattatia in 1979–1981 (IPLR-M). These two collection missions were conducted in the same eco-geographic space yet were, over the years, subjected to considerably different conservation pathways. By analysing phenotypic and genotypic data from these two collections we examine the hypothesis that some long-term *ex situ* conservation practices used in the majority of gene banks might lead to erosion of both genetic and phenotypic diversity.

MATERIAL AND METHODS

Germplasm and IPLR panel construction

The entire IPLR collection ($n = 932$) consists of germplasm from collections made in the early 20th century up until 1981. Those are held in gene banks and private collections worldwide, including (see Table 1) the Australian Grains Genebank (AGC; $n = 98$), the National Database Switzerland (BDN; $n = 8$), the Palestinian Biodiversity & Environmental Research Center (BERC; n = 8), the Centre for Genetic Resources, the Netherlands (CGN; $n = 14$), the International Maize and Wheat Improvement Center (CIMMYT) in Mexico ($n = 64$), the Crop Research Institute in the Czech Republic (CRI; $n = 3$), the Institute for Cereal Crops Improvement, Israel (ICCI; $n = 25$), the Israel Gene Bank (IGB; $n = 194$ – includes the Y. Mattatia collection, $n = 130$), the French National Institute for Agricultural Research (INRA; $n = 2$), the Leibniz Institute of Plant Genetics and Crop Plant Research, Germany (IPK; $n = 44$), the Watkins Collection of the John Innes Centre, Norwich, UK (JIC; $n = 23$), the Jordanian National Agricultural Research Center (NARC; $n = 3$), the Nordic Gene Bank (NGB; $n = 4$), the National Small Grains Collection of the United States Department of Agriculture (USDA; $n = 90$), the N.I. Vavilov All-Russian Institute of Plant Genetic Resources (VIR; $n = 78$), the private collection of Professor Moshe Feldman, held at the Weizmann Institute of Science in Israel (MF; $n = 270$), and collections made in recent years from local Druze and Palestinian farmers (LF; $n = 4$). In a small number of cases ($n = 24$) we included accessions that originated *outside* the region (Algeria, Tunisia, Yemen, Iraq, Pakistan, Italy, UK, and France), when the given accession name was identical to a name of a local 'variety' we knew to be in cultivation in the region based on historical texts. We assumed these to be cases of germplasm transfer in one or two directions, and because of the scarcity of authentic local landraces, we chose to include these lines as well. Accessions were transferred to the collection under material transfer agreements when necessary. In some cases, materials from a given gene bank were transferred via a second gene bank in order to consolidate, expedite, and simplify transfer logistics and plant protection protocols. Germplasm arriving from outside of Israel was first propagated under quarantine. After receipt of plant protection certification for the imported samples, the entire collection went through field propagation and characterization. Combined passport data and morphological, phenotypical and genotypic data (details are given later) were used to assess the 'authenticity' of each accession in the collection.

Historical background on IPLR-VIR and IPLR-M sub-collections

To address the question of sample heterogeneity and its preservation, we focus herein on two IPLR sub-collections, which took place in Israel/Palestine: IPLR-VIR collected by N.I. Vavilov in 1926 ($n = 78$) and IPLR-M collected by Y. Mattatia from 1979 to 1981 ($n = 130$). Nikolai I. Vavilov conducted his expedition to the Near East in 1926 as he travelled from Syria to the Sinai Desert and Egypt. His expedition in British Mandatory Palestine covered Jerusalem, the Jezre'el Valley, the Sea of Galilee, the Jordan river, the Dead

Table 1. Restoration of the IPLR collection ($n = 932$)

Germplasm donor[a]	No. of accessions	Collection years	Region of origin
AGC	98	1924–1973	Israel, Jordan, Algeria, Syria, Egypt
BDN	8	n.d.	Israel, Syria
BERC	8	n.d.	Palestine
CGN	14	1967–1975	Israel, Syria, France
CIMMYT	64	1967	Israel
CRI	3	n.d.	Israel, Levant
ICCI	25	1987–1990	Israel, Lebanon, Egypt
IGB	64	1926–1997	Israel, Jordan
MATTATIA	130	1979–1981	Israel, Palestine, Lebanon, Egypt
INRA	2	Early 20th century	Palestine
IPK	44	20th century	Israel, Syria, Italy, Yemen
JIC	23	1920–1929, 1971	Palestine, Syria, Tunisia
NARC	3	n.d.	Jordan
NGB	4	Early 20th century	Israel
USDA	90	1915–1985	Israel, Palestine, Lebanon, Syria, Tunisia, Iraq, Turkey, Jordan, Egypt, Pakistan, UK
VIR	78	1926, 1959	Israel, Palestine, Lebanon, Syria, Jordan
MF	270	n.d.	Israel, Lebanon
LF	4	2014–2018	Israel
Total	932		

[a] AGC: Australian Grains Genebank; BDN: National Database Switzerland; BERC: Palestinian Biodiversity & Environmental Research Center; CGN: Center for Genetic Resources, the Netherlands; CIMMYT: International Maize and Wheat Improvement Center, in Mexico; CRI: Crop Research Institute, in the Czech Republic; ICCI: Institute for Cereal Crops Improvement, Israel; IGB: Israel Gene Bank; MATTATIA: Y. Mattatia collection, held in the IGB; INRA: French National Institute for Agricultural Research; IPK: Leibniz Institute of Plant Genetics and Crop Plant Research, Germany; JIC: the Watkins Collection of The John Innes Centre, Norwich, UK; NARC: Jordanian National Agricultural Research Center; NGB: Nordic Gene Bank; USDA: National Small Grains Collection of the United States Department of Agriculture; VIR: N.I. Vavilov All-Russian Institute of Plant Genetic Resources; MF: private collection of Professor Moshe Feldman, held at the Weizmann Institute of Science in Israel; LF: local Druze and Palestinian farmers; n.d.: no data available.

Sea, and the coastal plain including Gaza.[17] Vavilov (and/or his collaborators) recorded extensive data for each accession, including names, taxonomic/botanical descriptions, collection sites, map coordinates, and so on. From 1979 to 1981, Yaakov Mattatia carried out collecting trips in villages in the West Bank, the Golan Heights and the Sinai Desert. His samples were taken directly from the field and accompanied by written data gathered from the farmers, including precise coordinates, the farmer's name, accession name(s), agronomic cultivation methods, and culinary use. IPLR-VIR was regenerated repeatedly at VIR (apparently at least once every 5–10 years) over the course of ~90 years. In contrast, the samples in IPLR-M went through a single regeneration in 1996.[18] Both the original IPLR-M germplasm and germplasm from the 1996 regeneration have been stored in the IGB since collection or regeneration.

Field experiment

A common garden experiment was carried out over two seasons in a protected net house containing Rhodoxeralf soil at Rishon LeZion, Volcani Agricultural Research Organization (ARO), Israel (31°59′32.7″N, 34°49′2.6″E). Supplementary irrigation was used to maintain optimal growing conditions. Herbicides were applied before sowing to mitigate weed pressure. The nursery was treated periodically with fungicides and pesticides to prevent the development of fungal pathogens or insect pests. Beds were weeded manually on a monthly basis. Nitrogen fertilization was applied both at pre-sowing and pre-heading (80 kg ha^{-1} at each stage).

2016–2017

Data were collected from 78 and 62 lines from VIR (IPLR-VIR) and Mattatia (IPLR-M) collections respectively, in addition to five modern Israeli cultivars. Seeds of IPLR-M had gone through a single propagation cycle[18] and had since been stored in low-humidity, low-temperature (15% relative humidity, 4 °C) facilities in the IGB at the Volcani Center, Rishon LeZion, Israel. As germination rate of the original collection (1979–1981) was poor (results not shown), seeds from the 1996 propagation cycle were used. Seed propagation was conducted using a single plot per accession, in 0.5 m^2 plots containing three rows spaced 20 cm apart. Annual rainfall was 404 mm, with 100 mm supplemental irrigation during the vegetative stage. Intra-sample variation of morphological uniformity index was calculated based on plant height and spike morphology (spike density, colour, and awns length). In cases where more than one phenotype was observed within the same plot, heterogeneity was recorded for those lines and was calculated as percentage out of total. Presence/absence of glaucousness on various plant organs (spike, stem, leaf) was recorded. These phenotypic parameters were used to assess intra-sample heterogeneity.

2017–2018

Data were collected from 75 and 56 lines from VIR (IPLR-VIR) and Mattatia (IPLR-M) collections respectively. In addition, 51 modern Israeli cultivars were included as a reference (MOD) for assessing the extent of phenotypic and genetic diversity in the collections. The experimental design was identical to 2016–2017 (as already noted). Annual rainfall was 398 mm, with an additional 100 mm supplemental irrigation during the vegetative stage.

Phenotypic characterization was assessed during the two seasons at: (i) vegetative stage by third leaf width (TLW – only in 2016–2017), canopy temperature at 75 days from sowing (CT; IR 200, IR thermometer, Extech Instruments, Nashua, NH, USA),

chlorophyll content (ChC; SPAD 502, Minolta chlorophyll meter), and flag leaf area (FLA; only in 2016–2017); and (ii) at reproductive stage, by plant height (PLH), days to heading (DtH), for each plot, days were counted from emergence until spike was fully exposed in 80% of plants, grain yield (GY) and 1000 kernel weight (TKW); (only for 2017–2018). In cases where morphological variation was observed, the dominant morph type was marked and collected separately from the other variants. Each variant within each sample was maintained separately for future conservation.

Molecular analysis

All lines were genotyped during the 2016–2017 season using the 96-marker 'Durum Wheat Reference Collection KASP Markers' (http://www.cerealsdb.uk.net/cerealgenomics/CerealsDB/wheat_durum_ref.php). Five plants from each IPLR accession were grown in pots at the Weizmann Institute of Science, Rehovot, Israel, out of which one plant was randomly selected for genotyping. Leaf samples were collected following the protocol of the LGC KASP SNP assay for genotyping kit (three leaf spots per sample from a single plant). The raw genotyping data were filtered to remove low-quality and missing data and resulted in 83 informative KASP markers, including the *Rht1* locus. The presence of *Rht1* (the 'Reduced Height-1' dwarfing alleles introduced during the 'Green Revolution' from 'Norin-10', a Japanese dwarf cultivar) was combined with measurements of PLH as a marker to identify modern semi-dwarf lines. Genetic distances and relatedness were estimated using the principal coordinate analysis (PCoA) as described by Abu-Zaitoun *et al*.[1] Clustering method (constellation plot, JMP® Pro v.14.0.0) was used to estimate genetic relations between accessions.

Statistical analysis

All the statistical analyses were performed using the JMP Pro v.14.0.0 statistical package (SAS Institute, Cary, NC, USA). Descriptive statistics were performed on the full dataset to illustrate variable distribution and to calculate coefficient of variance (CV), enabling comparison of phenotypic variability across traits.

RESULTS

Construction of the IPLR collection

The restoration and construction of the IPLR collection drew on germplasm donated by 18 gene banks and private collections worldwide. Information on germplasm donors and institutes, including collection date and region, appears in Table 1 and Fig. 1. The IPLR collection ($n = 932$) consists of a majority of *Triticum durum* lines ($n = 691$), as well as lines of *Triticum aestivum* ($n = 192$), *Triticum dicoccum* ($n = 26$) and a small number of species that are rare in this region (*Triticum spelta*, *Triticum polonicum*, *Triticum compactum* and *Triticum monococcum*; $n = 23$).

Passport data regarding the species of each of the accessions was verified (and corrected when needed to) using two tools: (i) SNP-based PCoA analysis (which separated the collection into tetraploid and hexaploid groups (Fig. 2(a)), and (ii) morphologic characterization of spike form (Fig. 2(b)–(e)). Using these tools, we were able to reclassify lines originally identified by gene bank passport data as *T. durum* (Fig. 2(b, c)) that were in fact *T. aestivum* ($n = 25$) and vice versa ($n = 46$) (Fig. 2(d, e)).

To eliminate duplicate accessions (e.g. in cases where two gene banks had donated germplasm from identical original accessions), molecular analysis (based on the 83 useful SNPs from the KASP assay) was cross-referenced with passport data (e.g. accession numbers, names) and phenotypic parameters. Out of the total collection, 32 accessions were found to be duplicates and removed from the collection. Landrace 'authenticity' was verified by the presence of the *Rht* null (tall) allele, cross-referenced with measurement of PLH. Altogether, 29 accessions were identified as 'non-authentic' using the aforementioned criteria. Most of the 'non-authentic' samples originated in fairly later collections, such as IPLR-M ($n = 4$), BERC ($n = 7$), IGB ($n = 9$), LF ($n = 1$), and MF ($n = 1$). Few, however, came from much older collections (USDA, $n = 4$; VIR, $n = 2$; and ACG, $n = 1$), which we imagine can only be due to human error (e.g. in data entry of collection year).

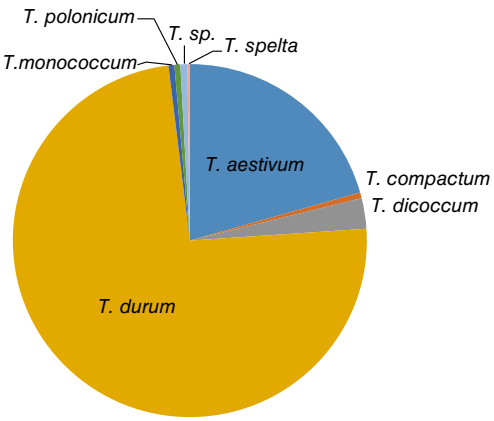

Figure 1. The portion of different *Triticum* species represented in the IPLR collection. The IPLR collection is constructed from *Triticum durum* ($n = 691$), *Triticum aestivum* ($n = 192$), *Triticum dicoccum* ($n = 26$) and scarce species (*Triticum spelta*, *Triticum polonicum*, *Triticum compactum* and *Triticum monococcum*; $n = 23$).

A number of common Arabic names of local landraces – e.g. 'Horani', 'Juljuli', 'Nursi', 'Gaza', 'Doubbi', 'Hatti' – are represented multiple times in the accessions' passport data (spikes of randomly selected IPLR accessions that have associated names are presented in Fig. 3(a)). However, the PCoA shows high genetic variability within accessions holding the same common name, and in many cases no clear relatedness between them was found (Fig. 3(b)). The accessions named 'Nursi' (said to originate from the village of Nuris on the western slopes of Gilboa Mountain) are an exception; they can be seen to cluster in the right upper part of the PCoA chart (blue dots). Also notable are accessions titled 'Hatti' (green dots) and 'Horani' (pink dots), which both tend to aggregate in a particular location in the genetic space, although in these cases certain individual accessions within each name-group are scattered across the chart.

The IPLR-VIR and IPLR-M case study

A detailed investigation of the genotypic and phenotypic variability of the IPLR-M and IPLR-VIR sub-collections has brought to light some notable differences between these two gene pools. As in the genetic structure of the wide collection (Fig. 2(a)), the genotypic data show a clear separation into *T. aestivum* and *T. durum* clusters (Fig. 4). As can be seen on the phylogenetic tree, the two sub-collections (IPLR-VIR and IPLR-M) and the modern reference group MOD, are generally located on separate tree branches. Both IPLR-VIR and IPLR-M occupy a wider genetic spectrum than MOD, expressed both in number of branches and in branch length (Fig. 4). The *T. aestivum* section bifurcates into two branches; most of the MODs are located on one of these branches ($n = 40$),

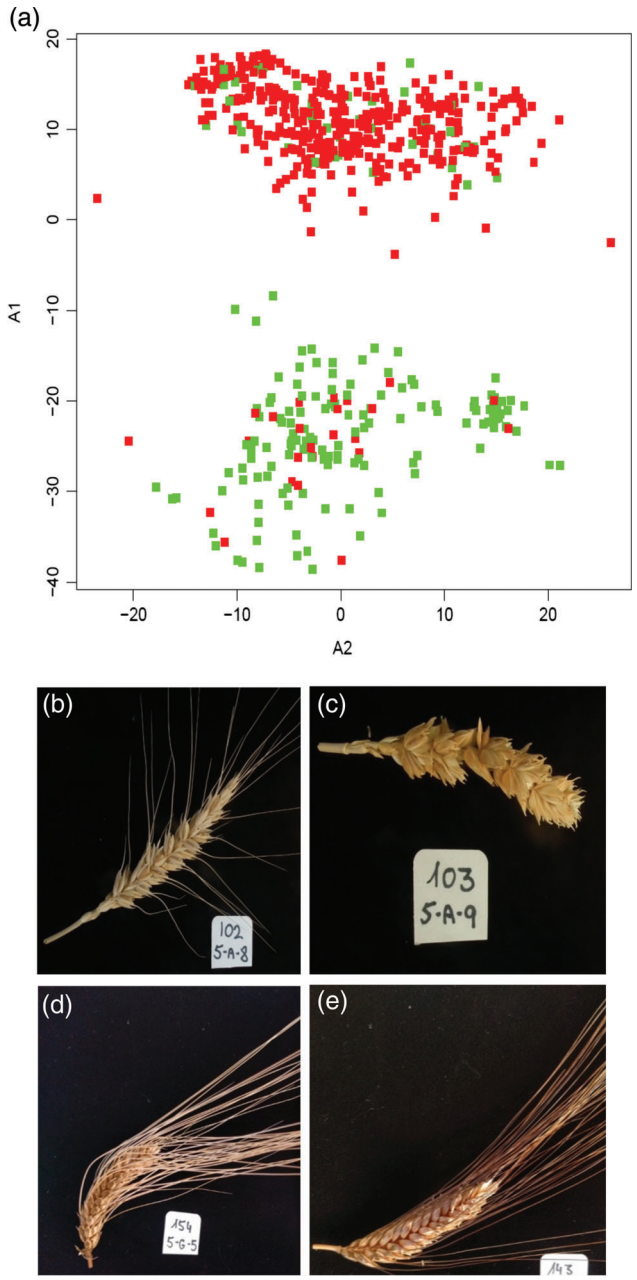

Figure 2. IPLR classification into tetraploid and hexaploid accessions. (a) Principal coordinate analysis (based on 83 KASP markers data) separating accessions into tetraploid (red) and hexaploid groups (green). (b–e) Morphologic characterization of spike form appearance, representing cases in which lines that were defined as *Triticum durum* (b, c) are reclassified to be *Triticum aestivum* and vice versa (d, e).

along with a few IPLR-VIR ($n = 7$) and IPLR-M ($n = 5$) accessions. The second branch is populated mostly by IPLR-VIR data points ($n = 39$), with only three accessions from IPLR-M and one from MOD. The *T. durum* section also splits into two main branches. The *T. durum* MOD cultivars occupy a much denser space than IPLR-M and IPLR-VIR do, both of which are mostly scattered on separate branches distinct to each collection. Interestingly, in the tetraploid *T. durum*, 9 of the 11 MOD cultivars do cluster to a single branch, whereas the hexaploid *T. aestivum* MOD cultivars ($n = 40$) are more variable and are present on several branches (Fig. 4).

IPLR-VIR and IPLR-M differ not only by chronology (a time gap of over 50 years between collections) but also by expedition routes within the region of Israel and Palestine. More than half of the IPLR-VIR accessions were collected along the Israeli coastal plain. All IPLR-M accessions were collected in the central mountain belt (e.g. Judean Mountains, Samaria, and eastern Upper Galilee). In some cases, however, the two sub-collections do overlap geographically: Around Jerusalem, Mattatia collected numerous samples of both *T. durum* and *T. aestivum* (Fig. 5(f, g)), whereas IPLR-VIR holds only one *T. aestivum* wheat line from a nearby site (Fig. 5(a)). Generally, both sub-collections show morphologic variability and distinct heterogeneity *between* accessions; however, IPLR-M shows a higher *intra-accession* variability relative to the homogeneous morphology of IPLR-VIR accessions (Fig. 5). In the southern Samaria mountains, IPLR-M samples were collected at five different field locations, all reflecting in-field heterogeneous mixtures. These mixtures contain between two and five different spike morphs. This diversity of spike morphology is expressed in spike length, spike density, awn length, and spike colour (white *versus* black; Fig. 5(h–l)). Interestingly, the majority of these mixtures contain both *T. durum* and *T. aestivum* accessions reflecting not only intra-crop variability but also actual crop species mixture (Fig. 5(h–j)). These crop mixtures are evident across all the IPLR-M samples in the collection. Moreover, a few IPLR-M samples also contained barley plants. Hence, we can assume that (at least at those sites that were sampled), IPLR-M represents seeds/spikes collected from farmers who used their own varietal mixture in the field as opposed to modern cultivars (which at that time were already available for purchase). In contrast, the IPLR-VIR samples collected in proximity to these IPLR-M samples were all homogeneous bread wheat accessions (Fig. 5(b–e)).

IPLR-M demonstrated marked morphological intra-accession variability in spike morphology, PLH, and canopy colour (Fig. 6(a)) leading to a uniformity index of only 16%, whereas IPLR-VIR showed a uniformity score of 86%. The presence of glaucousness (waxiness) on stem, leaves, and spikes was found to be highly variable within samples of IPLR-M in contrast to IPLR-VIR (Fig. 6(b)). In IPLR-M, 43.5%, 9.6%, and 27.4% of the accessions showed glaucousness in the stem, leaf, and spike respectively, whereas 62.9% of the accessions had spikes with partial (mixed) glaucousness. In IPLR-VIR, about 93.5%, 92.3%, and 75.6% of the accessions demonstrated glaucousness in stem, leaf, and spike respectively, whereas 16.6% of the accessions had mixed glaucousness in spike and 2.56% had mixed glaucousness in stem and leaf.

Independently of morphological appearance, the phenotypic data from two seasons (2016–2017 and 2017–2018) shows higher *intra-accession* phenotypic variability in IPLR-M than in IPLR-VIR and MOD. This is expressed most significantly in higher CV values in PLH and GY in 2016–2017 (Fig. 7(a)) and in ChC, PLH, DtH, and CT in 2017–2018 (Fig. 7(b)). Interestingly, values of the IPLR-VIR were only slightly higher in CT and ChC than MOD and IPLR-M (2016–2017) and in TKW than IPLR-M (2017–2018). Only the CV value of TKW was higher in MOD than in IPLR-VIR and IPLR-M (2017–2018). The TKW mean differences between durum and bread wheat are the main driver of high CV. This could explain the detection of low CV of TKW in IPLR-M compared with IPLR-VIR and MOD as it consists mainly of durum lines ($n = 51$) and has a very small portion of bread wheat ($n = 5$). The CV value of DtH in 2016–2017 was similar in both IPLR sub-collections and higher than in MOD.

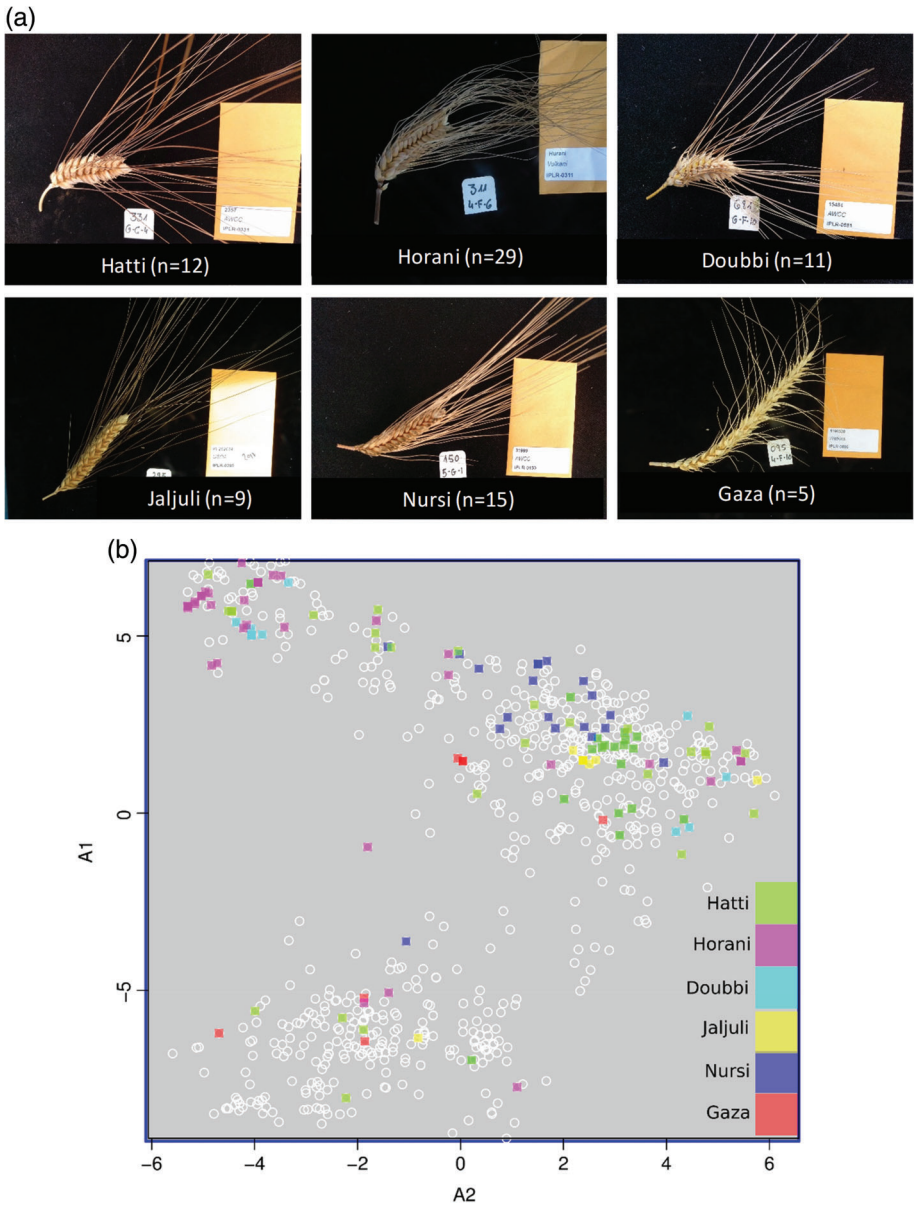

Figure 3. Frequent traditional Arabic names of IPLR landraces. (a) Spike morphology. Each frequent name group (n = No. of accession carrying the same name) represented by spike of random accession. (b) PCoA based on genetic distances obtained from 83 SNP. Colors represent frequent names. The intensity of the color reflects the number of accessions with the same genotype.

DISCUSSION

It is widely agreed that diversity of wheat landraces, such as those collected from fields of Palestinian *falaheen* (peasants), might provide a wide source of adaptive alleles that could serve wheat research and pre-breeding programmes.[6,10,19] This wide allelic diversity is prominent when compared with the diversity of modern germplasm, underlining the significant potential that landraces may possess for present and future crop improvement. However, how much of the allelic repertoire of the wheat grown by historical *falaheen* is actually still available today? We know that many traditional 'varieties' were totally lost in the transition to modern cultivars, but to what extent have collection and conservation strategies affected the diversity of landrace accessions that *were* preserved? And what sampling and conservation methodologies might maximize the preservation of such diversity in the future?

The IPLR collection from a historical view

The IPLR collection ($n = 932$) consists mostly of *T. durum* ($n = 691$), a portion that corresponds to the importance of durum wheat in the region. This holds not only to the Levant but also across the Mediterranean basin, where durum wheat is used for traditional Mediterranean foods, such as pasta, couscous, bulgur, and some flat breads, and is considered a dominant crop.[20,21] The restoration of the IPLR collection was based upon collections spanning six decades; the earliest recorded collections of wheat landraces germplasm are from the beginning of the 20th century, and the collecting practice continued up until the early 1980s. Two early explorers and collectors of wheat landraces germplasm in Palestine articulated a clear vision with far-reaching consequences as to the potential significance of this biodiversity; visions still highly relevant today. Aaron Aaronsohn predicted the technological progress of the local population and stressed the danger of

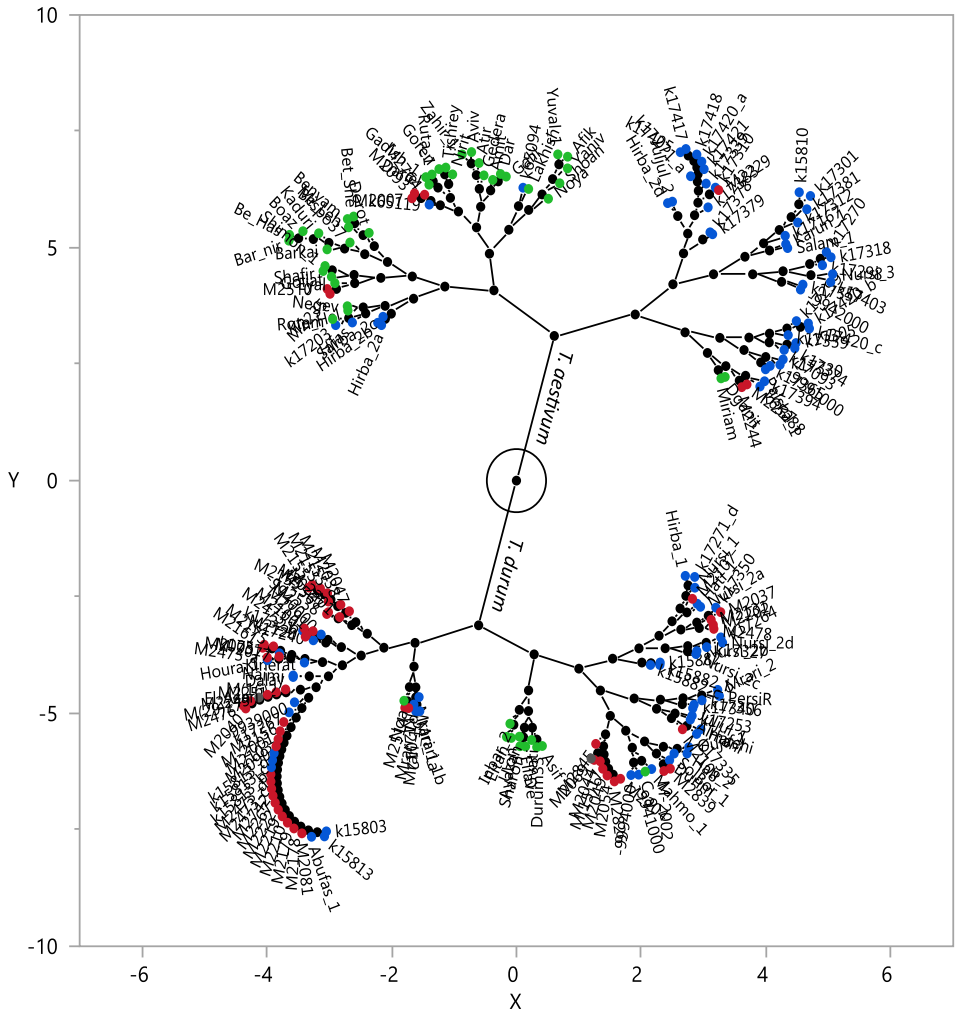

Figure 4. Genetic distances based on 83 KASP single-nucleotide polymorphism assay between IPLR-VIR, IPLR-M and MOD. Constellation plot of genetic distances between IPLR-M (red), IPLR-VIR (blue) and modern reference set (MOD) (green) marks the genetic space each of the collections.

extinction of local wheat landraces as a direct consequence. For this reason, in 1910 Aaronsohn urged for 'botanical–agronomic exploration of The Orient' in order to collect and preserve samples of local wheats.[22] A decade later, Nikolai I. Vavilov stressed the agronomical value of sample collection and *ex situ* conservation of crop plants and their wild relatives in gene banks. His vision and predictions provided inspiration for the establishment of important institutes for germplasm conservation and national and international gene banks,[7] some of whom donated germplasm to the IPLR collection (Table 1).

It is worth mentioning that some major collections of local wheat landraces made in expeditions during the 20th century did *not* survive as vital germplasm. This includes Aaronsohn's original collection (1913), now stored at the Herbarium at the Hebrew University of Jerusalem, and the collection of Dr Ludwig Pinner (1929), which is stored at the Department of Land of Israel Studies and Archaeology, at Bar-Ilan University. The Efrat collection (1967) ($n = 1553$) was also completely lost except for a somewhat representative core-collection ($n = 64$)[16] that was sent to CIMMYT, Mexico, for conservation by A. Blum two decades later.

At the beginning of 20th century, the dominant 'varieties' in the fields of Palestinian *falaheen* were mostly *T. durum* landraces, such as 'Horani', 'Jaljuli', 'Nursi', 'Gaza', 'Doubbi', and 'Hatti' (Fig. 3).[23] A possible explanation for the phenomenon of genetic variability within a given name group (Fig. 3(b)) might be the fact that landrace names in this region are often used to denote the region or village from which a given 'variety' comes from. Thus, wheat of any variety from the village of Jaljuliah might be called 'Jaljuli', wheat originating from (or close to) the Gaza strip was called 'Gaza', and the various forms of 'Horani' are wheats that came from the Horan Heights (also known as Houran, Hauran, etc., a region that today spans southern Syria and northern Jordan). However, despite references to rigid 'hyperlocality' of wheat landraces in this region,[22] when taking into account millennia of cultivation accompanied by movement of people, germplasm, and knowledge, it is hard to believe that these landrace families were truly restricted to a specific narrow geographic region. One of the few authentic landraces that is still cultivated today in the Akko Valley by Druze farmers is called 'Nursi', although lines with this name supposedly originated from Nuris, a village 70 km east (a significant distance in local terms) on the western slopes of Mount Gilboa.[22]

Introduction of foreign varieties into Palestine is documented as early as the 1920s,[24] and became quite frequent in the 1940s.[25] This, along with a growing demand for varieties suitable for bread baking (i.e. *T. aestivum*), fuelled a gradual transition away from local

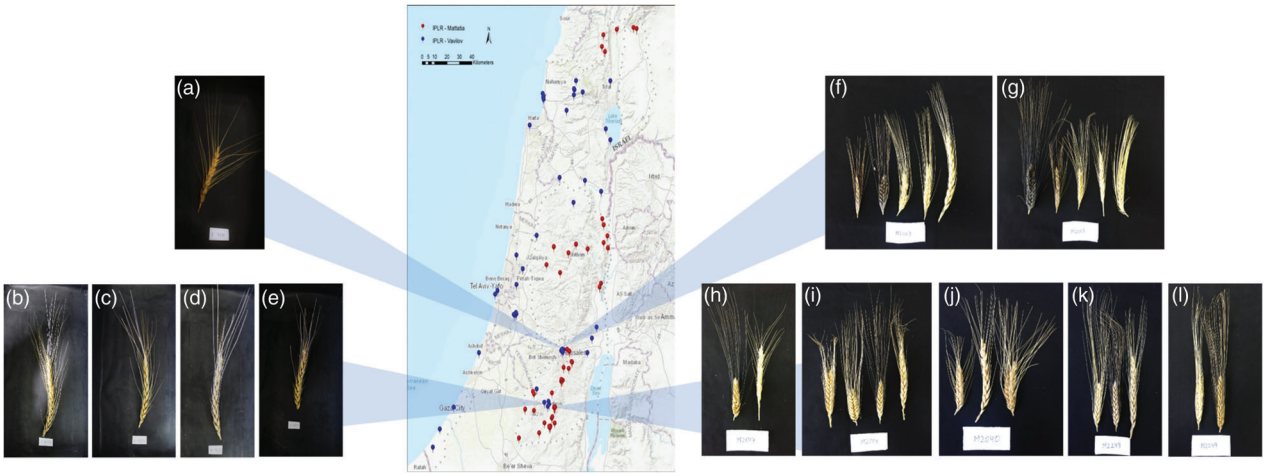

Figure 5. Collection sites of IPLR-VIR (blue) and IPLR-M (red). IPLR-VIR accessions (a–e) and IPLR-M accessions (f–l) that were collected from same geographic area manifesting morphologic variability.

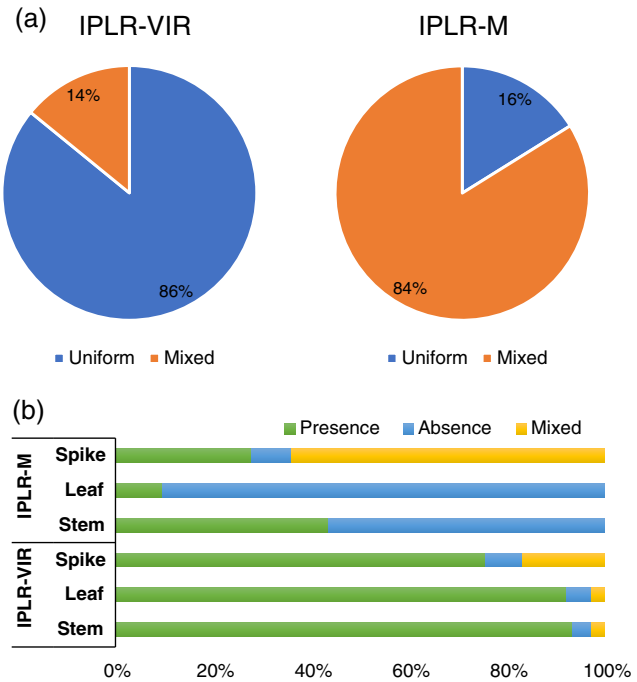

Figure 6. Comparative view of phenotypic traits coefficient of variance between collections. (a) 2016–2017: third leaf width (TLW), canopy temperature (CT, °C), chlorophyll content (ChC), flag leaf area (FLA), days to heading (DtH), plant height (PLH, cm) and grain yield (GY, g). (b) 2017–2018: ChC, 1000 kernel weight (TKW, g), PLH (cm), DtH, and CT (°C). Intra-sample heterogeneity in the IPLR-VIR and IPLR-M accessions. (a) Variation within accessions among IPLR-M and IPLR-VIR (Index combining plant height and spike morphology). In cases where more than one phenotype was observed within the same plot,heterogeneity was recorded for those lines and was calculated as percentage out of total. (b) Glaucousness variation (%) within stem, leaf and spike among IPLR-M and IPLR-VIR.

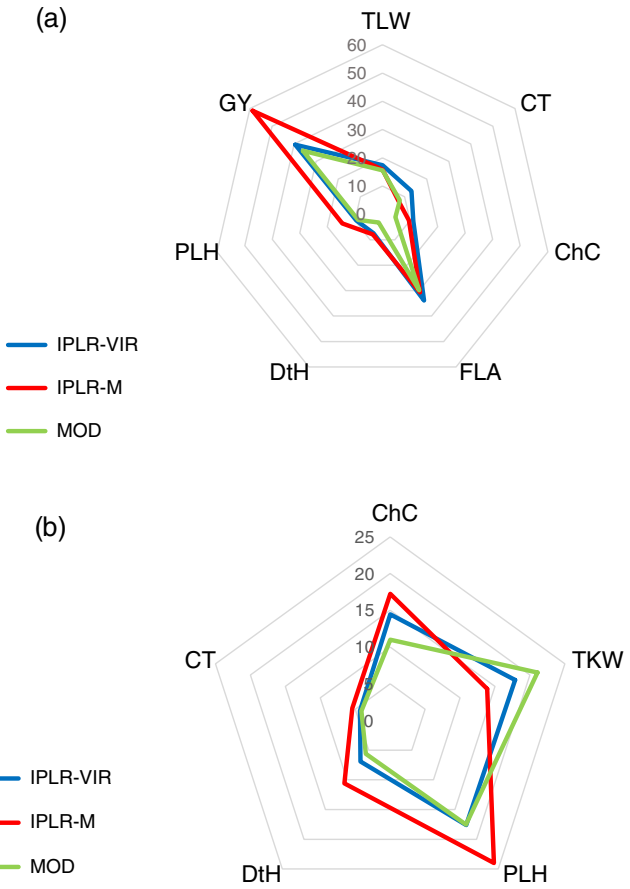

Figure 7. Comparative view of phenotypic traits coefficient of variance between collections. (a) 2016–2017: third leaf width (TLW), canopy temperature (CT, °C), chlorophyll content (ChC), flag leaf area (FLA), days to heading (DtH), plant height (PLH, cm) and grain yield (GY, g). (b) 2017–2018: ChC, 1000 kernel weight (TKW, g), PLH (cm), DtH, and CT (°C).

(mostly durum) landraces. This transition was paralleled by the dramatic intensification and mechanization processes that took place across Palestine at that time. As occurred the world over, the introduction of external germplasm peaked of course with the impact of 'Green Revolution' wheat varieties.[26] For Israeli breeders, the transition to semi-dwarf modern cultivars was virtually exhaustive; the process was completed in Israeli wheat fields during the 1960s, with a similar transition occurring in Palestinian farms in the West Bank and Gaza from 1967 till the early 1980s. Due to ethno-cultural barriers, the process of replacing traditional landraces with modern semi-dwarf cultivars seems to have been more gradual.

The last two major collection expeditions in this region were conducted by Y. Efrat in 1967[27] and Y. Mattatia from the late

1970s to early 1980s.[14] With regard to these collections made after the 'Green Revolution', in a time where diffusion of modern semi-dwarf cultivars into local farmers fields presumably occurred, it was important to assure the authenticity of sampled wheat landraces (i.e. to ascertain that these varieties believed by the farmers themselves to be 'traditional', or even 'ancestral' – passed on from parents to children over multiple generations – were not in fact modern cultivars, or crosses with modern cultivars, introduced fairly recently into the field). To this end, we combined PLH field data with the allelic profile of the *Rht* locus. This composite 'filtering tool' identified 29 accessions that were both of short stature and carried the *Rht1-B1* dwarf allele (recruited by Borlaug from the Japanese cultivar 'Norin 10'). Most of these outliers (~93%) indeed originated from late collections made from the 1970s and onwards.

Additionally, 32 lines were identified as duplicates based on their identical passport data (accession numbers, names, etc.), identical spike morphology, and zero or near-zero genetic distance (based albeit on only 83 markers). Out of 32 lines, 22 duplicates were from the USDA. This might be the consequence of germplasm transfer from the USDA to other gene banks where the receiving gene bank would have assigned new accession IDs compatible with the receiving gene bank's cataloguing system. PCoA analysis of the KASP SNP genotyping assay resulted in two clusters clearly separating *T. durum* and *T. aestivum* (Fig. 2). This information, combined with spike morphology data, facilitated the identification of 71 accessions of misclassified species in the passport data. We can only speculate that such misclassification may have occurred due to human error in classification and/or in data entry, or to situations in which originally heterogeneous samples included both species.

The Vavilov *versus* Mattatia collection case study

In the current study, IPLR-M, IPLR-VIR, and MOD were evaluated for their genotypic and phenotypic diversity. Both the IPLR-M and IPLR-VIR sub-collections show high genetic diversity compared with MOD. This is expressed in wider and more extended branching of the phylogenetic tree in both the tetraploid and hexaploid genetic backgrounds (Fig. 4). Moreover, each collection mainly distributes on distinct branches. These findings could shed light on the potential variability of this genetic resource and support the hypothesis presented in the discussion preface regarding the promising potential of such exotic germplasm.[19]

The phenotypic variability measured during the vegetative and reproductive stages is expressed in this study by CV values. Higher CV values in IPLR-M (Fig. 7) might reflect the potential adaptation of different genotypes composing this sub-collection to different field conditions. High heterogeneity was observed within the IPLR-M accessions compared with IPLR-VIR (Fig. 6). This was notable in morphological uniformity index and degree of glaucousness on various plant organs. Beharav *et al.* also observed sample heterogeneity while propagating and evaluating these lines in the fields in 1996.[18] In addition to heterogeneity of forms within the same species, in some cases ($n = 24$) there were clear crop mixtures of *T. aestivum*, *T. durum*, and even *Hordeum vulgare*. These records are in accordance with previous descriptions of wheat landraces as consisting of mixed plant types.[14,16,28] More than one factor could be speculated as a driver for the low phenotypic heterogeneity of IPLR-VIR compared with IPLR-M. First among these is the type of samples collected. The Vavilov expedition took place in early autumn[17] during the sowing season, too early for direct collection of germplasm from the fields. This fact might have dictated collection of more homogeneous landrace samples (in traditional local markets or taken from wheat sheaves). The second factor is the various propagation and selection methodologies practiced in VIR's gene bank. And third is genetic drift affected by significant G × E interactions, such as the transition of Mediterranean-adapted spring-type genotypes grown for millennia over the winter (November to May) in a relatively mild climate at a lower latitude into propagation cycles that took place over the spring and summer (April to August) at the high latitudes and cold environments of the former USSR.

Based on the foregoing, it appears that certain collection strategies, along with long-term *ex situ* conservation efforts involving human subjective selection and G × E interactions inherent to multiple cycles of propagation, may significantly reduce accession heterogeneity and could lead to loss of genetic diversity. Integration of *in situ* elements into the conservation strategy might be a possible way to address this issue. Such a strategy may also benefit from the true on-farm conditions of *in situ* conservation (which in some cases can even take place under more traditional agronomic management),[29] all which may contribute to more accurate conservation of heterogeneous landrace entities. However, compared with *ex situ* conservation, the *in situ* strategy requires resources, expertise, and long-term infrastructure, which are not always available in conventional institutional gene banks. In that context, fully implemented *in situ* conservation of a wide collection such as IPLR is probably not pragmatic. Nevertheless, our finding regarding the genetic bottleneck associated with century-long *ex situ* practices might underline the value of integrating *in situ* conservation as a complementary tool to standard *ex situ* conservation. Such an approach is being tested currently in the IGB, where in parallel with the long-term (−20 °C) *ex situ* conservation of all IPLR accessions, a subset of heterogeneous accessions has been assigned to on-farm sites (e.g. botanical gardens, national natural parks, community gardens) who are able to support careful *in situ* cultivation of these accessions (https://igb.agri.gov.il/web/?page=47&lang=he). In addition, we are in the process of deeper genotyping of the entire collection, which, along with phenotypical data, will enable us to select a smaller 'core-collection' closely representing the entire IPLR collection in order to facilitate the prioritization of germplasm for both *ex situ* and *in situ* conservation.

CONCLUDING REMARKS

The IPLR collection represents a potentially unique and diverse germplasm source for allelic variation of wheat landraces that evolved over millennia in the Levant.[14] This study enables a first glimpse into the genetic diversity deposited in this collection that appears to reflect long-term local adaptability. Owing to historical agrarian transitions, the majority of this genetic treasure was virtually on the verge of extinction. Gathering germplasm from gene banks and institutes worldwide and construction of the IPLR collection was an urgent and essential phase that has been (almost) successfully completed (a few hundred accessions from the region that are stored in a number of international gene banks are still absent from the collection).

Further study is underway to investigate and characterize this collection as well as to determine the optimal paths for conservation. Evaluation of wheat landraces' genetic diversity can facilitate identification of rare and beneficial alleles[30] and help sustain the crop genetic improvement[31] necessary to provide plant breeders and farmers high-yield stress-tolerant germplasm and crops.[32] The evidence generated by the current study broadens the theoretical context in which cereal landraces are currently defined and can

stimulate discussions on how best to handle intra-sample diversity in the context of modern gene banks designed for optimized conservation.

ACKNOWLEDGEMENTS

Thanks to the 'Eretz Hita Wheat Landrace Initiative', without which this IPLR collection and this study would not have been possible. Thanks to Michal Gurevich for adding useful comments, to Dr Smadar Ezrati and Cathy Melamed-Bessudo for helping to process the genotyping, and to Nadav Yudekovitz for his help in fieldwork and phenotyping. Special thanks to Dr Thomas Payne, the head of CIMMYT gene bank, for his awareness in rescuing the Efrat core-collection. This study was supported by a 2016 IGB grant and Ministry of Agriculture grant 20-01-0097. Translation and access to historical texts, which provided invaluable context for this study, was funded in part by Yad Hanadiv. SF's doctoral studies are funded by Gvati, Danziger, and Sam-Hamburg fellowships.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

SUPPORTING INFORMATION

Supporting information may be found in the online version of this article.

REFERENCES

1 Abu-Zaitoun S, Chandrasekhar K, Assili S, Shtaya M, Jamous R, Mallah O *et al.*, Unlocking the genetic diversity within a Middle-East panel of durum wheat landraces for adaptation to semi-arid climate. *Agronomy* **8**:233 (2018).
2 Roselló M, Royo C, Álvaro F, Villegas D, Nazco R and Soriano JM, Pasta-making quality QTLome from Mediterranean durum wheat landraces. *Front Plant Sci* **9**:1512 (2018).
3 Giunta F, De Vita P, Mastrangelo AM, Sanna G and Motzo R, Environmental and genetic variation for yield-related traits of durum wheat as affected by development. *Front Plant Sci* **9**:8 (2018).
4 Sherman JD, Nash D, Lanning SP, Martin JM, Blake NK, Morris CF *et al.*, Genetics of end-use quality differences between a modern and historical spring wheat. *Crop Sci* **54**:1972–1980 (2014).
5 Abbo S, Lev-Yadun S and Gopher A, Plant domestication in the Near East, in *Encyclopaedia of the History of Science, Technology, and Medicine in Non-Western Cultures*, ed. by Selin H. Springer, Dordrecht, Netherlands, pp. 1–9 (2008).
6 Dwivedi SL, Ceccarelli S, Blair MW, Upadhyaya HD, Are AK and Ortiz R, Landrace germplasm for improving yield and abiotic stress adaptation. *Trends Plant Sci* **21**:31–42 (2016).
7 Tanksley SD and McCouch SR, Seed banks and molecular maps: unlocking genetic potential from the wild. *Science* **277**:1063–1066 (1997).
8 Smithson JB and Lenné JM, Varietal mixtures: a viable strategy for sustainable productivity in subsistence agriculture. *Ann Appl Biol* **128**:127–158 (1996).
9 Abbo S, Lev-Yadun S and Gopher A, Yield stability: an agronomic perspective on the origin of Near Eastern agriculture. *Vegetation History and Archaeobotany* **19**:143–150 (2010).
10 Jaradat A, Wheat landraces: a mini review. *Emirates Journal of Food and Agriculture* **25**:20–29 (2013).
11 González MY, Philipp N, Schulthess AW, Weise S, Zhao Y, Börner A *et al.*, Unlocking historical phenotypic data from an *ex situ* collection to enhance the informed utilization of genetic resources of barley (*Hordeum* sp.). *Theor Appl Genet* **131**:2009–2019 (2018).
12 Jacubziner MM, *The Wheats of Syria, Palestine and Transjordania: Cultivated and Wild*. All-Union Institute of Plant Industry, Leningrad, Russia (1932);(in Russian).
13 Pinner L, *Wheat Culture in Palestine. Bulletin of the Palestine Economic Society*, Vol. **5**. Palestine Economic Society, Tel Aviv, Israel, pp. 33–46 (1930).
14 Mattatia J, *Exploration and Collection of Primitive Land-races of Field and Garden Crops for the Gene Bank (1980/81)*. The Volcani Center, Beit Dagan, Israel: Department of Scientific Publications (1981).
15 Mattatia J, *Exploration and Collection of Primitive Land-races of Field and Garden Crops for the Gene Bank (1981/82)*. The Volcani Center, Beit Dagan, Israel: Department of Scientific Publications (1982).
16 Poiarkova H and Blum A, Land-races of wheat from the northern Negev in Israel. *Euphytica* **32**:257–271 (1983).
17 Vavilov NI, in *Five Continents*, ed. by Reznik S and Stapleton P. International Plant Genetic Resources Institute (IPGRI), Rome, Italy (1997).
18 Beharav A, Golan G and Levy A, Evaluation and variation in response to infection with *Puccinia striiformis* and *Puccinia recondita* of local wheat landraces. *Euphytica* **94**:287–293 (1997).
19 Zamir D, Improving plant breeding with exotic genetic libraries. *Nat Rev Genet* **2**:983–989 (2001).
20 Soriano JM, Villegas D, Sorrells ME and Royo C, Durum wheat landraces from east and west regions of the Mediterranean Basin are genetically distinct for yield components and phenology. *Front Plant Sci* **9**:80 (2018).
21 Nazco R, Villegas D, Ammar K, Peña RJ, Moragues M and Royo C, Can Mediterranean durum wheat landraces contribute to improved grain quality attributes in modern cultivars? *Euphytica* **185**:1–17 (2012).
22 Aaronsohn A, *Agricultural and Botanical Exploration in Palestine. US Department of Agriculture, Bureau of Plant Industry Bulletin*. Government Printing Office, Washington, DC, pp. 1–63 (1910).
23 Pinthus MJ, *History of Agriculture Research in Israel in the Pre-Statehood Era (1920–1949)*. Itai Bahur, Zikhron-Ya'aKov (2011). In Hebrew.
24 Zagorodsky M, Memories from the experimental station. *Palestinias Agriculturist* **5**:10–12 (1919).
25 Plaut M and Zelcbuch B, Quality of wheat in Israel. *Bull Res Counc Isr* **3**:73–79 (1953).
26 Evenson RE and Gollin D, Assessing the impact of the green revolution, 1960 to 2000. *Science* **300**:758–762 (2003).
27 Blum A, Sinmena B, Golan G and Mayer J, The grain quality of landraces of wheat as compared with modern cultivars. *Plant Breed* **99**:226–233 (1987).
28 Kostrinski J. Acclimatization of Grain Varieties. Ha'Sade. 1943;**23**:89–90. In Hebrew.
29 Traditional management of seed and genetic diversity | Genes in the Field | Taylor & Francis Group. [Online]. Available: https://www.taylorfrancis.com/books/e/9781420049824/chapters/10.1201/9781420049824-5 [24 January 2019].
30 Lopes MS, El-Basyoni I, Baenziger PS, Singh S, Royo C, Ozbek K *et al.*, Exploiting genetic diversity from landraces in wheat breeding for adaptation to climate change. *J Exp Bot* **66**:3477–3486 (2015).
31 Sehgal D, Vikram P, Sansaloni CP, Ortiz C, Pierre CS, Payne T *et al.*, Exploring and mobilizing the gene bank biodiversity for wheat improvement. *PLoS One* **10**:e0132112 (2015).
32 Plucknett DL, Smith NJ, Williams JT and Anishetty NM, Crop germplasm conservation and developing countries. *Science* **220**:163–169 (1983).

