



## Yield, yield stability and farmers' preferences of evolutionary populations of bread wheat: A dynamic solution to climate change

Riccardo Bocci<sup>a</sup>, Bettina Bussi<sup>a</sup>, Matteo Petitti<sup>a</sup>, Riccardo Franciolini<sup>a</sup>, Virginia Altavilla<sup>a</sup>, Gea Galluzzi<sup>a</sup>, Paolo Di Luzio<sup>b</sup>, Paola Migliorini<sup>c</sup>, Sandra Spagnolo<sup>d</sup>, Rosario Floriddia<sup>e</sup>, Giuseppe Li Rosi<sup>f</sup>, Modesto Petacciato<sup>g</sup>, Vincenzo Battezzato<sup>h</sup>, Andrea Albino<sup>i</sup>, Giovanni Faggio<sup>j</sup>, Carlo Arcostanzo<sup>j</sup>, Salvatore Ceccarelli<sup>a,\*</sup>

<sup>a</sup> Rete Semi Rurali, Via di Casignano, 25, 50018, Scandicci (FI), Italy

<sup>b</sup> AIAB Molise, Contrada Solagne-S. Giacomo degli Schiavoni (CB), Italy

<sup>c</sup> University of Gastronomic Sciences, Piazza Vittorio Emanuele, 9, 12060, Pollenzo, Bra (CN), Italy

<sup>d</sup> 4AIAB Piemonte, Via Stura, 24, 10098, Rivoli (TO), Italy

<sup>e</sup> Azienda Agricola Floriddia, Via della Bonifica, 171 56037, 56037, Peccioli (PI), Italy

<sup>f</sup> Simenza, Raddusa (CT), 95049, Italy

<sup>g</sup> Azienda Agricola Biologica Petacciato Modesto, Contrada Giammetta, 86040, San Giuliano di Puglia (CB), Italy

<sup>h</sup> Azienda agricola Vincenzo Battezzato, Contrada Colle Calcare, 53, 86100, (CB), Italy

<sup>i</sup> Soc. Agr. Fonte Santa Maria snc, Via dei Mille 121, 86040, Montorio nei Frentani (CB), Italy

<sup>j</sup> Il Papavero Rosso, Via Valle Maira 109, 12100, Confreria (CN), Italy

### ARTICLE INFO

#### Keywords:

Evolutionary plant breeding  
Heterogeneous plant populations  
Climate change  
Landraces  
Local adaptation  
Biodiversity  
Farmers' preferences  
Directional selection

### ABSTRACT

The role of biodiversity in plant breeding needs to be reconsidered to cope with the complexity, location specificity and combined challenges of climate change, human population increase, human health and food security, safety and sovereignty. From a biological viewpoint, heterogeneous plant populations derived from evolutionary plant breeding may address most of these grand challenges. Field trials were conducted over four years and four locations under organic farming conditions to test the hypothesis that evolutionary populations planted in contrasting locations, evolve adapting to the local conditions and becoming distinct from one another. The experiment also included mixtures, landraces and a modern variety of bread wheat. The results show evidence of divergent selection for grain yield under the sole effect of natural selection by which the best performing evolutionary populations yielded as much or more than the commercial variety. Farmers' selection within one of the evolutionary populations was effective in improving yield and yield stability above those of the original population across years and locations. Farmers' preference was not always associated with grain yield and was not gender dependent. We conclude that evolutionary populations are able to gradually evolve, adapting to each environment in which their seed is multiplied, reaching high and stable yield levels thus ensuring income to farmers, both as seed and as grain.

### 1. Introduction

Biodiversity is key to food security, making production systems more resilient (FAO, 2019); it is also an essential resource for crop improvement, allowing the adaptation of agriculture to a changing climate and consumer preferences (Hufford et al., 2019; van Frank et al., 2020). However, biodiversity in general, and crop biodiversity in particular, are threatened and plant breeding is thought of being one of the causes (van

der Wouw et al., 2010). Indeed, already in 1950, Sir Otto Frankel warned, "From the early days of plant breeding, uniformity has been sought after with great determination. For this there are many reasons – technical, commercial, historical, psychological, and aesthetic. The concept of purity has not only been carried to unnecessary length but that it may be inimical to the attainment of highest production" (Frankel, 1950).

In 2009, the concept of 'planetary boundaries' was proposed to

\* Corresponding author.

E-mail address: [ceccarelli.salvatore83@gmail.com](mailto:ceccarelli.salvatore83@gmail.com) (S. Ceccarelli).

<https://doi.org/10.1016/j.eja.2020.126156>

Received 29 January 2020; Received in revised form 24 July 2020; Accepted 17 August 2020

Available online 28 September 2020

1161-0301/© 2020 Elsevier B.V. All rights reserved.

define a ‘safe operating space for humanity’ with respect to nine processes including climate change, rate of biodiversity loss, ozone depletion, ocean acidification, human interference with nitrogen and phosphorus cycles, global freshwater use, changes in land use, chemical pollution and atmospheric aerosol loading (Rockström et al., 2009). This article, as well as a paper published six years later (Steffen et al., 2015), shows that the boundaries of three of the above processes, namely climate change, rate of biodiversity loss and human interference with nitrogen and phosphorus cycles, have been already overstepped. According to Jaramillo and Destouni (2015), the use of fresh water has also gone beyond the associated planetary boundary. Conventional food systems, from production to processing, consumption, loss and waste are also major drivers of climate change (Vermeulen et al., 2012). Hence, many of the major challenges of the 21st century are closely related to each other, and particularly associated with industrial agriculture, which contributes greatly to, but is also greatly affected by, climate change (Springmann et al., 2018). One of the effects of climatic changes on agriculture is a decline in resilience of crop systems, as recently suggested for wheat in Europe (Kahiluoto et al., 2019; but see Piepho, 2019). The same authors argue that cultivating “a set of varieties with diverse responses to critical weather conditions is required to promote the climate resilience of a crop”. This approach, which uses agro-biodiversity in the form of crop heterogeneity to achieve food security, has already proven its effectiveness in reducing powdery mildew in barley (Wolfe et al., 1992) and in reducing rice blast, thereby increasing grain yield in China (Zhu et al., 2000).

In order to cope with the complex and combined effects of challenges such as human population increase, food security, food safety and human health, we need to reconsider the role of diversity in plant breeding. One strategy, which has proved to be efficient in maximizing genetic gains while deploying crop diversity has been decentralized selection, defined as selection in the target environment(s) (Ceccarelli, 2015). With its emphasis on specific adaptation, decentralized selection differs from centralized breeding by adapting crops to each target environment through different, locally adapted, varieties. Decentralized selection recognizes the importance of keeping genotype  $\times$  location (GL) interaction distinct from genotype  $\times$  year (GY) interaction, as already recommended more than 50 years ago by Allard and Hansche (1964). By subdividing the target population of environments in sub groups, each characterized by repeatable GL interaction, we can exploit genotype  $\times$  environment interactions (GEI). In fact, in this way, the variance component due to GEI is minimized, and therefore both heritability and selection gains are maximized. However, a distinction is rarely made in the literature between GL and GY (see for example Arief et al., 2020). This is unfortunate because GY is largely unpredictable while GL can to some extent be repeatable and therefore predictable. Moreover, decentralized selection can make a positive use of GL interactions by selecting for specific adaptation, while the solution to GY are well buffered cultivars against unpredictable fluctuations of the environment, which are increasing in frequency and extent as a consequence of climate change (Baethgen, 2010; Thornton et al., 2014). The buffering against environmental fluctuations can be achieved by growing heterogeneous populations, which combine both individual and population buffering. While individual buffering is a property of specific genotypes, and particularly of heterozygotes, population buffering arises from the interactions among the different genotypes within a population, beyond the individual buffering of specific genotypes; heterogeneous populations have the advantage of exploiting both individual and population buffering (Allard and Hansche, 1964).

There are two ways to developing heterogeneous populations: either by crossing in all or several combinations a number of varieties, obtaining what are known as Composite Crosses or Evolutionary or Bulk populations, or by mixing the seed of different varieties obtaining mixtures, which can be either static or dynamic (Wolfe and Ceccarelli, 2020).

Research on evolutionary populations (EPs) and mixtures has

spanned several decades, from the original paper of Harlan and Martini (1929) to recent years (Goldringer et al., 2006; Raggi et al., 2017). This large body of work, reviewed in 2011 (Döring et al., 2011), demonstrates that natural selection in EPs and mixtures is effective in changing phenology (Allard and Hansche, 1964; Goldringer et al., 2006), improving yield (Suneson, 1956; Rasmusson et al., 1967; Raggi et al., 2017), yield stability (Allard, 1961; Döring et al., 2015; Reiss and Drinkwater, 2018) and host plant resistance to pathogens (Wolfe et al., 1992; Smithson and Lenné, 1996; Finckh et al., 2000; Mundt, 2002). Research has also shown that the evolutionary potential of populations is higher than that of mixtures (Patel et al., 1987).

Therefore, there is little doubt that, from a biological viewpoint, EPs have the potential to address many of the grand challenges indicated above. However, they represent a challenge for the prevalent seed systems, which require distinctness, uniformity and stability (Bocci, 2009); in fact they are neither uniform nor stable, although they can be distinct. This could be one of the reasons explaining why in the past heterogeneous populations never spread among farmers as an alternative type of cultivars.

Despite the evidence of the beneficial effects of EPs and mixtures, research has not addressed their potential to respond to different selection pressures; these may occur across time within the same location as well as between different locations as expected based on field studies of the effects of natural selection (Kingsolver et al., 2001; Siepielski et al., 2009). If the same EP and/or mixture is able to evolve differently when planted continuously in different locations, exploiting this ability could become a cost-effective strategy to generate a continuous flow of sub-populations through a process of adaptive evolution, thus contributing to increase both temporal and spatial agrobiodiversity. This has been shown to be the case for an important adaptive character such as phenology (Goldringer et al., 2006), but we are not aware that this evolutionary ability applies to grain yield to a level that makes evolutionary populations commercially competitive.

Therefore, the main objective of this paper was to test the hypothesis that when the same evolutionary population with a large genetic basis is planted continuously in different locations, it is subjected to divergent selection for a number of traits including grain yield and that, as a consequence, it evolves into distinct subpopulations each with a yield advantage in the location where they evolved. In addition to testing the evidence for divergent evolution, the experiment involved a wide range of potential users (farmers, consumers, millers, bakers) both men and women, with the objective of detecting possible gender differences in the evaluation of an unusual genetic material.

## 2. Material and methods

### 2.1. Plant material

The experiment was conducted in Italy for four years (2015 through 2018) within the framework of the Horizon 2020 project DIVERSIFOOD ([www.diversifood.eu](http://www.diversifood.eu)). In the first year the experiment was conducted in three locations (Raddusa in Sicily, Peccioli in Tuscany and Campobasso in Molise): in the second year we added a fourth location (Caraglio, in Piedmont). The experimental material included 7 bread wheat evolutionary populations (EPs), three mixtures, three Italian landraces and one modern variety recommended for organic agriculture (Table 1).

Four evolutionary populations were derived from the same bread wheat EP made by the corresponding author and Dr. Stefania Grando in Syria, while working at the International Center for Agricultural Research in the Dry Areas (ICARDA) and therefore named ICA\_EP. The ICA\_EP was obtained in 2009 by mixing an approximately equal amount of seed of F<sub>2</sub>, F<sub>3</sub> and F<sub>4</sub> of 1996 crosses made available by the ICARDA's bread wheat breeder and obtained from 256 parents representing a wide range of germplasm types. In October 2010, in the framework of the H2020 EU project SOLIBAM ([www.solibam.eu](http://www.solibam.eu)), about 20 kg of seed of the population were sent to Italy and distributed to two organic farmers,

**Table 1**  
The experimental material.

ENTRY	NAME	ABBREVIATION	TYPE
1	Evolutionary Population ICARDA Floriddia <sup>a</sup>	ICA_FL	Composite Cross Population
2	Evolutionary Population ICARDA Li Rosi <sup>a</sup>	ICA_LR	Composite Cross Population
3	Evolutionary Population ICARDA Rosati <sup>b</sup>	ICA_RO	Composite Cross Population
4	Selection Floriddia	ICA_FLS	Farmer selection from ICA_FL
5	CCP_YQ	C_YQ	Composite Cross Population
6	CCP_Y	C_Y	Composite Cross Population
7	CCP_Q	C_Q	Composite Cross Population
8	Piedmont Mix <sup>c</sup>	P_MIX	Dynamic Mixture of landraces, old and modern varieties
9	Mix of old varieties <sup>d</sup>	OV_Mix	Static Mixture of old varieties
10	Mix of modern varieties <sup>e</sup>	MV_Mix	Static Mixture of modern varieties
11	Maiorca	Maiorca	Sicilian landrace
12	Andriolo	Andriolo	Tuscany landrace
13	Frassineto <sup>f</sup>	Frassineto	Old variety selected from the Tuscany landrace Gentilrosso
14	Gentilrosso awned <sup>g</sup>	Gentilrosso	Selected from the Tuscany landrace Gentilrosso
15	Emese	Emese	Modern variety for organic agriculture

<sup>a</sup> These two populations were authorized to participate in the temporary experiment organized as indicated by the Commission Implementing Decision of 18 March 2014 and provisionally named SOLIBAM TENERO FLORIDDIA and SOLIBAM TENERO LI ROSI.

<sup>b</sup> Introduced in 2016.

<sup>c</sup> Introduced in 2016 and composed by Andriolo, Gentilrosso, Frassineto, Gamba di Ferro, Abbondanza, Verna e Sieve.

<sup>d</sup> Introduced in 2016 and composed of Andriolo, Inallettibile and Frassineto (replaced by Gentilrosso awned in 2017 and 2018).

<sup>e</sup> Introduced in 2016 and composed of Emese, Bologna (replaced by PR22R58 in 2017 and 2018) and A416.

<sup>f</sup> Only in the first two years and then replaced by Gentilrosso awned in 2017 and 2018 because susceptible to bunt.

<sup>g</sup> Introduced in 2017 and 2018 as the genetically closest relative of Frassineto.

Mr. Rosario Floriddia (FL) in Tuscany and Mr. Giuseppe Li Rosi (LR) in Sicily who received 10 kg of seed each. Both farmers have planted and harvested the EP in their respective farms ever since using their own seed (Entries 1 = ICA\_FL and 2 = ICA\_LR); they continue to do so at the moment of writing. After 2010, the seed of the EPs spread from Sicily and Tuscany to several farmers in Italy through informal seed exchanges facilitated by Rete Semi Rurali (RSR). In 2014, 100 kg of the seed of the EP, which was evolving in Sicily, started being planted by Mr. Gianfranco Rosati (RO) in Abruzzo, generating a new EP named ICA\_RO (Entry 3). Furthermore, during the second year of cultivation (2012), Mr. Floriddia bulked a number of single heads personally selected from his EP, generating a new selected EP (Entry 4 = ICA\_FLS) while also maintaining the original population.

Entries 5 to 7 are three composite cross populations (CCP) created in 2001 in the UK by Prof. Martin Wolfe at the John Innes Centre in cooperation with the Organic Research Centre. In 2005, the three populations were sent to Germany (Bertholdsson et al., 2016) and became part of a joint research program (COBRA <https://orgprints.org/29766/>) with the participation of the Italian Association of Organic Agriculture (AIAB). The CQ was obtained by 66 half-diallel crosses between twelve varieties with high baking quality, the CY was obtained by 36 half-diallel crosses of nine varieties with high yield, while the CYQ was obtained by 190 half-diallel crosses between Y and Q parents (Döring et al., 2015). We will refer to entries 5 to 7 as the North European CCPs. Entries 8 to 10 are mixtures: in the case of entry 8, the mixture, made of varieties and landraces adapted to Piedmont (Andriolo, Gentilrosso, Frassineto, Gamba di Ferro, Abbondanza, Verna e Sieve) was propagated using the seed harvested from the experiment in each location, hence it was a dynamic mixture according to the definition of Wolfe and Caccarelli (2020). The other two mixtures (entries 9 and 10) were static mixtures as they were reconstituted at the beginning of each cropping season. Entry 9 is a mixture of three landraces (Andriolo, Inallettibile and Frassineto, with the latter replaced by Gentilrosso in 2017 and 2018). Entry 10 is a mixture of three modern varieties (Emese, A416 and Bologna, with the latter replaced by PR22R58 in 2017 and 2018); entry 11 (Maiorca) is a Sicilian landrace still widely cultivated in Sicily; entry 12 (Andriolo) is a landrace of unknown origin and common in Tuscany; entry 13 (Frassineto) is also a landrace from Tuscany obtained in 1922 by pedigree selection from Gentilrosso; entry 14 (Gentilrosso) is a landrace from Tuscany, which by the middle of the 19th century spread to other regions in Central Italy. Some of the landraces mentioned above

are now part of the list of conservation varieties (Commission Directive 1998/95/EC). However, in the paper they will be referred to as landraces. Entry 15 (Emese) is a modern variety bred in 2000 by the Agricultural Research Institute of the Hungarian Academy of Sciences (one of partners of the SOLIBAM project), with drought tolerance, early maturity, very good baking quality, excellent frost and lodging resistance, good resistance to powdery mildew, and moderate resistance to leaf rust and stem rust (Bedö et al., 2001) (Table 1).

## 2.2. Agronomic management

All four experiments were conducted on certified organic farms. Soil preparation and agronomic practices were carried out according to the routine farm practices and the rotations followed the farms' overall rotation scheme. Sowing was done by hand at a rate of 20 g/m<sup>2</sup>.

The experiments in Tuscany were hosted on Mr. Rosario Floriddia's farms, on the hills of Cedri, near Peccioli, province of Pisa (43° 29' 37.21"N, 10° 47' 51, 98"E, 124 m a.s.l.). The farm is managed following a three year rotation, wheat (bread or durum wheat) - chickpea - clover or Italian sainfoin, without externally sourced fertilizers; therefore the crop preceding wheat was either clover or sainfoin; however, in the first year, the experiment was planted after wheat as no other areas in the farm were available. The soil is 31.0 % sand, 33.4 % silt and 35.6 % clay (Table 2). Soil tillage includes ploughing or ripping with a pneumatic cultivator below 20 cm, disc harrowing or a vibrating cultivator. Weeding was done manually with a hoe. Sowing dates were January 5, 2015, November 6, 2015, November 16, 2016 and October 30, 2017.

In Sicily, the experiments were hosted on Mr. Giuseppe Li Rosi's

**Table 2**  
The main soil characteristics of the experimental sites.

Location	Sand	Silt	Clay	pH	Organic matter	Total (N g Kg <sup>-1</sup> )	P (mg Kg <sup>-1</sup> )
Tuscany	31.0	33.4	35.6	8.13	1.75	0.93	15.0
Sicily	50.2	7.7	42.2	–	1.97	1.26	14.0
Molise <sup>a</sup>	35.6	16.7	47.7	7.00	2.38	0.90	27.0
Molise <sup>b</sup>	38.2	10.6	51.3	8.50	1.73	1.00	19.0
Piedmont	55.0	25.5	19.5	6.48	3.12	2.00	28.0

<sup>a</sup> Location used in 2015 and 2016.

<sup>b</sup> Location used in 2017 and 2018.

farm, on the hills of Raddusa, province of Catania (37°29'37", 01°N, 14°29'12", 50°E 315 m a.s.l.). On his organic farm, Mr. Li Rosi follows a rotation with bread, durum and Khorasan wheat, alternated with legumes (chickpea, pea or lentil). The preceding crop was always a grain legume, either chickpea, or lentil or pea. The soil is a sandy/clay with 50.2 % sand, 7.7 % silt and 42.2 % clay (Table 2). The soil is first ploughed with a chisel tiller to 30 cm depth or with a subsoiler to 45 cm depth and then once or twice with a vibro cultivator. Weeding was done manually only when needed. Sowing dates were December 15, 2014, December 23, 2015, and between the second and the third week of December in 2016 and 2017 (no records were available with the exact dates).

In Molise, the experiments were hosted for the first two years by Mr. Vincenzo Battezzato, Contrada Colle Calcare, in Campobasso (41°32'49.4"N, 14°41'50.4"E, 530 m a.s.l.), while in 2017 and 2018 they were hosted by Mr Andrea Albino near Montorio nei Frentani (41°46'08", 36 N, 14°56'24", 40 E) in 2017, and near Larino (41°46'08", 28 N, 14°54'20", 63 E) in 2018. The soil on both farms is deep and with average fertility, but with different pH, organic matter and phosphorus (Table 2). Both farms use bio fertilizer made of manure and phosphorus at the rate of 250 and 300 kg/ha, respectively. The former is organically managed since more than 10 years, with a rotation of mixed vegetable crops (onion, tomatoes, zucchini, salads, etc.), legumes (often beans) and cereals (both maize and winter cereals). The preceding crop was onion; soil was ploughed with a spading machine at 20 cm depth while weeding was done with a rotary tiller at 10 cm depth. The latter is also organically managed since 15 years, with a typically extensive rotation: durum wheat (or spelt wheat), legumes (chickpeas, field beans, lentils) and sunflowers. The preceding crops were beans in 2017 and chickpea in 2018. Soils were ploughed with a subsoiler at 40 cm and weeding was done mechanically with a grubber and then with a vibro cultivator. Tillage is shallow with a multi-plough while weeding was done manually. In 2017 the experiment was planted at about 460 m a.s.l. while in 2018 the experiment was planted at 520 m a.s.l. Sowing dates were December 13, 2014, December 16, 2015, December 26, 2016 and January 5, 2018.

The experiments in the north-western region of Piedmont started in 2016 and were hosted on an organically managed farm on the outskirts of Caraglio, province of Cuneo (44°25'7", 37°N, 7°26'11.79"E, 555 m a.s.l.). The land belongs to the craft bakery enterprise "Il Papavero Rosso". The farm is located at the bottom of a steep valley and the crop rotation includes bread wheat, emmer, rye, durum wheat alternated with leguminous fodder crops (clover, alfalfa) without externally sourced fertilizers. The soil is very fertile with about 3 % organic matter, 55.0 % sand, 25.5 % lime and 19.5 % clay, rich in phosphorus (Table 2) and with average depth. The wheat crop is always preceded by two years of red clover. Soils were ploughed with a subsoiler at 40 cm and weeding was done mechanically with a grubber and then with vibro cultivator. Sowing dates were November 13, 2015, November 18, 2016 and November 24, 2017.

In 2015, we included 10 entries in the experiment (Entries 1, 2, 4, 5, 6, 7, 11, 12, 13 and 15 in Table 1), which were replicated twice with 20 plots of 10 m<sup>2</sup> in a row and column design with 4 rows and 5 columns. In 2016, we added four entries (Entries 3, 8, 9, 10 in Table 1) and in all subsequent locations and years we used 14 entries in a randomized block design with two replications and 28 plots of 10 m<sup>2</sup> arranged in 4 rows and 7 columns. However, as indicated earlier, one of the entries used in the first two years (Frassineto) was replaced by Gentilrosso in 2017 and 2018. Therefore, although we tested a total of 15 entries, in the last three years we always had only 14 entries. Randomization was different in each year and each location and was generated by DiGger (R Core, 2019), a program that generates efficient experimental designs for non-factorial experiments with plots arranged in rows and columns (Coombes, 2009; Cullis et al., 2006).

In the case of evolutionary populations and of dynamic mixtures, the seed for the experiment in any given year and location was a random

sample of the seed produced by the same population and/or mixture in that location the year before.

### 2.3. Measurements

In all plots and in all location-year combinations, we measured grain yield (expressed as kg ha<sup>-1</sup>) and participants' preferences. In some locations and years we also measured plant height, early ground cover and weed density. Plant height was measured at maturity during the first 3 years of the experiment only (the first two in the case of Piedmont) from ground level to the bottom of the spike on 10 random plants in the inner part of the plots to avoid border effects. Early ground cover was estimated in 2015 in Tuscany and Molise and in 2016 in Tuscany, Molise and Piedmont, with a visual score from 1 = minimum soil cover to 5 = maximum soil cover at the beginning of tillering; weed density was estimated at flowering in 2015 and 2016 in Tuscany and Molise by counting the number of weeds in 1 m<sup>2</sup> randomly chosen inside the plot.

### 2.4. Public evaluation of performance in the field

All plots were evaluated at or near full maturity (peduncle completely yellow) by a total of nearly 500 participants during field days organized by RSR with the collaboration of the host farmers (Table 3). Participants included farmers, agronomists, scientists, students, millers, bakers, consumers and Ministry of Agriculture officials. No limits were imposed to either the number or the gender or the profession of the participants; however, the first four categories were invariably the most represented. Before the evaluation, the participants assembled near the experiment and one RSR staff introduced the experiment's objectives, the type of material being evaluated (but not the names or the position of the entries), the layout and the most convenient path to follow inside the experimental area. Before starting the evaluation, the participants were informed that their evaluation and the data collected would be analyzed and the results submitted for publication. In most locations there was another experiment on one side of the bread wheat experiment, which was also evaluated: therefore, the participants were divided in two groups, with one group starting with the side experiment and the other group with the bread wheat experiment. At the end of the evaluation in one experiment, the two groups switched over and in this way the number of people walking inside each experiment at any given moment was reduced. Participants used a form, with only plot numbers and no information on the plant material in each plot. The form was divided in four columns with the following headers: bad, medium, good, and excellent. Each participant classified each plot by marking one of the four columns; the marks were transformed, for computational purposes, into numerical scores from 1 = bad to 4 = excellent. Adding name and profession to the top of the form was left optional. The form also had a blank space available for the participants to enter the reason(s) for their evaluation but this space was so sparsely filled that this information was only used for discussion. The participants were requested to do the evaluation individually without consulting with each other. RSR staff was available at different points of the field to facilitate the movement of the participants through the plots and to answer questions. Only when all the participants completed the evaluation, did RSR staff reveal, when requested, which material was planted in individual plots. In Tuscany and Sicily, because of the possibility of comfortably assembling the participants, immediately after the evaluation RSR staff shared a quick summary of the evaluation indicating which entries had received more often the minimum and the maximum score. We analyzed the average score of all participants, as well as the average score of women and of men separately. We did not disaggregate the score by the other categories of the participants as this will be the subject of another study.

### 2.5. Statistical analysis

After testing for normality, the data were first analyzed on a single



**Table 3**Number and gender (m = men; w = women)<sup>a</sup> of people who did the evaluation in the four years and the four locations.

Location	2015	2016	2017	2018	Total
Tuscany	39 (23 m - 16 w)	76 (52 m - 17 w)	47 (31 m - 16 w)	49 (29 m - 20 w)	211 (135 m - 69 w)
Sicily	30 (22 m - 8 w)	43 (29 m - 14 w)	40 (30 m - 8 w)	43 (36 m - 7 w)	156 (117 m - 37 w)
Molise	13 (9 m - 4 w)	15 (12 m - 3 w)	12 (7 m - 1 w)	23 (15 m - 8 w)	63 (44 m - 19 w)
Piedmont	–	29 (17 m - 7 w)	13 (10 m - 3 w)	26 (17 m - 9 w)	68 (44 m - 19 w)
Total	82 (56 m - 26 w)	163 (110 m - 41 w)	112 (78 m 28 w)	141 (97 m 44 w)	498 (341 m - 139 w)

<sup>a</sup> The sum of men and women participants within parenthesis does not always match the total number of participants in a given year/location combination in front of the parenthesis (as, for example, in the case of Tuscany 2016) because the gender was not recorded for all the participants.

environment (year-location combination) basis using the spatial analysis described by Singh et al. (2003) and by Rollins et al. (2013). The spatial analysis generated the Best Linear Estimates (BLUEs) of the Entries' effects, obtained by using the most suitable spatial model given the experimental layout. The entries' BLUEs were analyzed by an analysis of variance (ANOVA) using a linear mixed model (LMM), where the entry yield BLUE  $Y_{ijz}$  is a function of the grand mean  $\mu$ , of the Entry (E) effect of the  $i$  Entry, of the Location (L) effect of the  $j$  location, of the Year (Y) effect of the  $z$  Year, of the two order interactions effects (EL), (EY), (LY) and of the effect of the (EYL) used as an estimate of the experimental error:

$$Y_{ijz} = \mu + E_i + L_j + Y_z + (EL)_{ij} + (EY)_{iz} + (LY)_{jz} + (EYL)_{ijz}$$

In the analysis Entry was considered as fixed, while Location and Year were considered as random.

Both analyses were done using ANOVA for unbalanced design in GenStat 20<sup>th</sup> Edition (VSN, 2019).

The interactions between the Entries and the 15 combinations of Years and Locations were further investigated in the case of grain yield and participants' score by using the GGEBIplot R package (Yan et al., 2000; Frutos et al., 2014; R Core, 2019). We used two features of the GGE biplot, namely the relations between traits and the "mean and stability" feature. In the first feature, and when the data are sufficiently approximated by the biplot, the cosine of the angle between the vectors of two traits approximates the correlation coefficient between them. In particular, two traits are positively correlated if the angle between their vectors is  $<90^\circ$ ; two traits are negatively correlated if the angle between their vectors is  $>90^\circ$ ; two traits are independent if the angle between them is near  $90^\circ$  (Yan and Rajcan, 2002). In the second feature, whose properties are described by Yan and Hunt (2020), a line (called the mean environment axis), which passes through the biplot origin is drawn in the biplot. A second line is drawn, perpendicular to the mean environmental axis, which passes through the origin. The projections of the entries tested in the experiment on the mean environment axis approximate their mean yields (or the mean of the trait being plotted), while the projection to the perpendicular axis approximates the GEI associated with the entry. The longer the projection, the greater is the GEI, which is a measure of the instability. In this graphical representation the ideal entry, in terms of grain yield and stability, is one that has the longest positive projection on the mean environment axis (high mean) and a zero projection on the perpendicular axis (high stability). Thus, in the resulting graph, the genotypes are evaluated for a combination of high mean and high stability in the sense of low GEI.

The information on the stability of grain yield generated by the GGE biplot, was validated using the environmental variance (EV) (Döring et al., 2015) as an estimate of static stability, the cultivar-superiority measure proposed by Lin and Binns (1988), and the mean rank (Nassar and Hühn, 1987). The three measures of stability were calculated using the "GEstability" procedure of the software GenStat version 20.1 (VSN, 2019). Ekbohm's (Ekbohm, 1981) test was used as test of significance for the environmental variance, while for the cultivar superiority measure we used the test proposed by Lin and Binns (1988). The entry mean ranks are presented with their standard errors. An entry with a low mean rank for grain yield is one of the top yielders (low ranking) in most

year-locations combinations. We excluded from the stability analysis the two entries (Gentilrosso and Frassineto) for which we only had two years data.

In the case of grain yield, we also used the yield reliability measure proposed by Kataoka (1963). The measure combines mean yield across environments with EV across the same environments as shown below, where the  $I_i$  is the reliability index for the entry  $i$ ,  $m_i$  is the entry mean for the entry  $i$ ,  $Z_{(P)}$  is the percentile from the standard normal distribution for which the cumulative distribution function reaches the value  $P$ , the probability of the lowest yield, (Annicchiarico, 2002), and  $EV_i$  is the environmental variance of entry  $i$ :

$$I_i = m_i - Z_{(P)} \sqrt{EV_i}$$

The yield reliability measure estimates the lowest yield, which is expected with a given probability chosen according to the level of risk aversion by farmers: the values of  $Z_{(P)}$  range from 0.675 for  $P = 0.75$ , namely a moderate risk aversion, to 1.645 for  $P = 0.95$ , namely a strong risk aversion (Annicchiarico, 2002). We used  $P = 0.80$  instead of the suggested 0.75 for modern agriculture in favorable conditions to take into account the increased level of risk due to an increased climate variability (Thornton et al., 2014).

Because of the evolving nature of some of the experimental material, it was expected that the original seed sample sent from Syria would evolve differently in different locations in response to specific selection pressures. Therefore, the most interesting comparisons for these materials were considered to be those across different locations: among those we extracted the following specific contrasts to be examined in detail.

Specific contrasts of interest were:

- 1 The three ICA populations versus each other;
- 2 The three ICA populations vs. the modern variety Emese;
- 3 The three ICA populations vs. the three mixtures;
- 4 The three ICA populations vs. the landraces.

The level of significance of the contrasts were tested using the LSD based on the error term from the ANOVA.

These comparisons were limited to grain yield, plant height and participants' preference because of the limited data available for ground cover and weed density. We attempted to answer the question frequently asked by farmers interested in growing evolutionary populations: "how rapidly do they evolve to become adapted to the conditions of my area?" by analyzing the 2015 data, namely the first year of the experiment. Three of the evolutionary populations, namely ICA\_LR, ICA\_FL and ICA\_FLS, by 2015 had evolved independently for at least three cropping seasons (ICA\_FLS) up to five (ICA\_LR and ICA\_FL). The analysis of the 2015 data was conducted on both grain yield and plant height.

### 3. Results

#### 3.1. Climatic data

Meteorological data was collected for the two locations where the bread wheat EPs had been evolving since its introduction to Italy (i.e. Sicily and Tuscany) since October 1st 2010. Meteorological data for the

two other locations was collected for the time period in which the experiment was conducted there (Molise 2015–2018; Piedmont 2016–2018).

Daily maximum and minimum temperatures and rainfall were obtained from the closest meteorological stations. The thermal sum for the growing season was obtained adopting a TBASE = 0 °C, as in [McMaster and Wilhelm \(1997\)](#). No upper temperature threshold TUT was adopted, in order to highlight potential climatic trends, in line with [Costanzo and Bärberi \(2016\)](#). Monthly and seasonal rainfall were calculated for each location-year.

In the two locations (Tuscany and Sicily) where the ICA\_FL and the ICA\_LR evolved for a longer and similar period of time, the yearly mean seasonal (October-June) rainfall (in mm) and temperature (expressed as thermal time) show different patterns. In Tuscany ([Fig. 1A](#)), rainfall was very variable with no consistent pattern as indicated by the low regression coefficient ( $b = -0.002$ ,  $R^2 = 0.0256$ ,  $P = 0.705$ , d.f. = 6). Temperature showed a slight tendency to increase ( $b = 0.006$ ,  $R^2 = 0.2456$ ,  $P = 0.211$ , d.f. = 6) probably due to the 2014–2016 period being warmer than average. In Sicily ([Fig. 1B](#)), by contrast, rainfall decreased over time, but not significantly ( $b = -0.0106$ ,  $R^2 = 0.3674$ ,  $P = 0.111$ , d.f. = 6), despite the wetter than usual 2014. A similar consistent pattern was observed for temperature that, on average, was always higher than in Tuscany, and increasing over time though not significantly ( $b = 0.0152$ ,  $R^2 = 0.3630$ ,  $P = 0.114$ , d.f. = 6).

In the case of Molise and Piedmont ([Fig. 1C and D](#), respectively) we only had four and three years meteorological data, respectively, which are only shown graphically as they could explain some of the agronomic results discussed later.

### 3.2. Agronomic characteristics

The location mean yield varied from a minimum of 359 kg ha<sup>-1</sup> in Tuscany in 2015 and of 473 kg ha<sup>-1</sup> in Piedmont in 2018 to a maximum of 3386 kg ha<sup>-1</sup> in Piedmont in 2017, an almost ten-fold difference

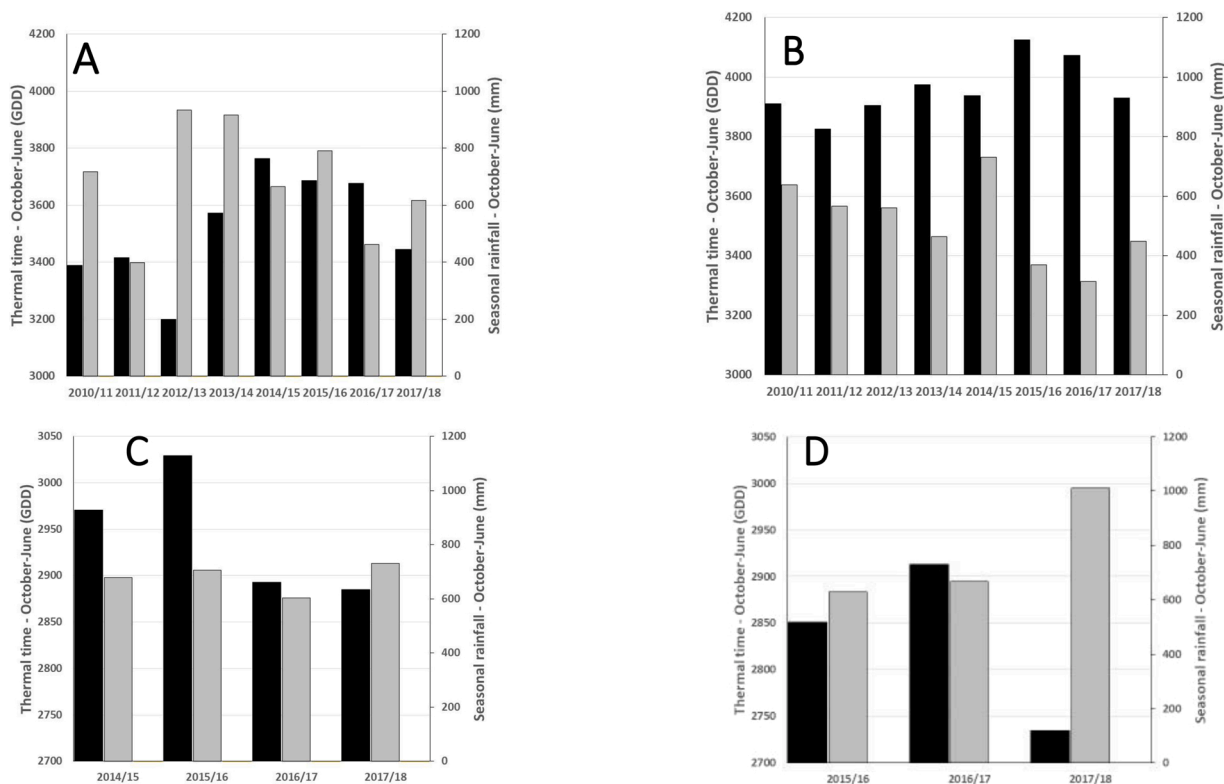
([Table 4](#)). A four-fold yield difference due to yearly weather difference was also observed in Tuscany, while in Sicily and Molise year to year differences were less pronounced. However, the lowest yielding year was not the same in all locations – being 2015 in Tuscany, 2016 in Sicily, and 2018 in both Molise and Piedmont. The shortest crop was observed in the lowest yielding environment (Tuscany 2015), while the tallest was observed in the northernmost location (Piedmont). In the two years where the comparison was possible, plant height in Piedmont was markedly different from the other three locations. Although differences between locations were observed for both early ground cover and weed density, they seem to be season-dependent.

There were significant differences between entries and locations for all traits except for weed density; years differed for all traits except for weed density and participants' agronomic score. Interactions between locations and years were significant for all traits, while interactions between entries and both locations and years were significant in the case of grain yield ([Table 5](#)). Only entry by location interaction was significant in the case of plant height, while entry by year interaction was significant ( $P < 0.04$ ) for the participants' agronomic score.

The highest yielding entry across the 15 combinations of years and locations was the ICA\_FLS, namely the population selected by the farmer from the EP which evolved in Tuscany ([Table 6](#)). The selection was significantly higher yielding than the population which evolved in the same location only as result of natural selection (ICA\_FL), but not significantly higher yielding than the EP which evolved in Sicily (ICA\_LR) without any farmer selection.

The lowest yielding entries were the landraces Andriolo and Gentilrosso, although not significantly lower yielding than Maiorca, the modern variety Emese and the three North European CCPs. The mixtures (entries 8–10) were intermediate in yield, significantly higher yielding than the landraces and significantly lower yielding than ICA\_FLS only.

The mixture of modern varieties was significantly shorter than all the other entries, considerably shorter (between 12 and 17 cm) than the North European CCPs and the Piedmont mixture ([Table 6](#)) and shorter



**Fig. 1.** Seasonal (October – June) rainfall (mm) (light grey bars) and temperature (as thermal time) (black bars) at the locations where the experiment was conducted: A = Tuscany, B = Sicily, C = Molise, D = Piedmont.

**Table 4**

Average grain yield, plant height, ground cover and weed density of the 15 entries listed in Table 1: Grain yield was measured in four locations (three in 2015) and four years, plant height in the first three years.

Location	Grain Yield (kg ha <sup>-1</sup> ± se)				
	2015	2016	2017	2018	Mean
Tuscany	359 ± 30	1503 ± 119	1365 ± 87	720 ± 77	987 ± 73
Sicily	2560 ± 186	1842 ± 158	1927 ± 262	2001 ± 212	2058 ± 37
Molise	1462 ± 45	942 ± 45	1055 ± 67	778 ± 93	1059 ± 39
Piedmont	–	2192 ± 75	3393 ± 217	473 ± 42	2019 ± 182
			Plant Height (cm ± se)		
Tuscany	52.6 ± 2.8	92.0 ± 3.9	90.6 ± 4.8	–	78.4 ± 3.7
Sicily	83.4 ± 5.1	86.3 ± 5.1	86.7 ± 5.5	–	85.5 ± 3.0
Molise	74.2 ± 3.1	85.5 ± 3.1	74.0 ± 4.2	–	77.9 ± 2.2
Piedmont	–	110.6 ± 4.0	102.5 ± 5.6	–	106.6 ± 4.5
			Ground Cover (score 1=poor to 5=good ± se)		
Tuscany	2.00 ± 0.17	3.31 ± 0.10	–	–	2.65
Sicily	–	–	–	–	–
Molise	2.99 ± 0.21	2.96 ± 0.11	–	–	–
Piedmont	–	3.20 ± 0.13	–	–	–
			Weed Density (number per m <sup>2</sup> ± se)		
Tuscany	12.2 ± 2.16	4.4 ± 0.30	–	–	–
Sicily	–	–	–	–	–
Molise	6.15 ± 0.54	11.97 ± 1.38	–	–	–
Piedmont	–	–	–	–	–

**Table 5**

Unbalanced Analysis of variance mean squares and F probability levels for grain yield (GY in kg ha<sup>-1</sup>), plant height (PH in cm), ground cover (GC from 1=min to 5=max), weed density (WE as number of weeds in 1 m<sup>2</sup>) and participants agronomic score (AS from 1= min to 4=max) in the four locations and four years (degree of freedom are adjusted for missing values).

Source of Variation	d.f.	GY	P	d.f.	PH	P	d.f.	GC	P	d.f.	WE	P	d.f.	AS	P
Entries (E)	14	406897	0.031	14	1882	***	13	0.470	0.030	13	24.05	0.253	14	0.995	***
Location (L)	3	18259761	***	3	4983	***	2	0.975	**	1	46.39	0.116	3	0.639	0.003
Years (Y)	3	8921118	***	2	3293	***	1	4.145	***	1	13.77	0.363	3	0.234	0.155
E*L	42	298655	***	42	133	***	26	0.206	0.242	13	24.45	0.242	42	0.167	0.167
E*Y	34	209515	***	21	64	0.101	9	0.126	0.518	9	14.86	0.518	34	0.209	0.040
L*Y	8	5957967	0.004	5	1094	***	1	5.091	***	1	456.5	***	8	0.291	0.033
Residual (E*L*Y)	93	104121		54	41		9	0.130		9	15.37		93	0.131	
Total	197	833773		141	443		61	0.409		47	30.20		197	0.229	

\* P < 0.05; \*\* P < 0.01; \*\*\*; P < 0.001.

**Table 6**

Entry means across locations and years for grain yield (GY in kg ha<sup>-1</sup>), plant height (PH in cm), ground cover (GC from 1=min to 5=max), weed density (WE as number) and participants' agronomic score (AS from 1 to 4). Values with letters in common are not significantly different at P < 0.05 based on LSD test.

Entry	Abbreviation	GY	PH	GC	WE	AS
1	ICA_FL	1549 bcd	98.9 cd	3.20 ab	7.18 a	2.63 de
2	ICA_LR	1754 ab	80.8 fg	3.26 ab	8.18 ab	2.55 ef
3	ICA_RO	1424 cde	93.8 de	3.18 abc	8.15 a	2.45 g
4	ICA_FLS	1953 a	88.9 ef	3.40 a	5.14 a	2.82 bc
5	CYQ	1450 cde	74.5 g	2.61 c	9.45 ab	2.43 gh
6	CY	1453 cde	75.2 g	2.68 bc	12.54 ab	2.45 fg
7	CQ	1375 de	75.0 g	2.63 c	10.54 ab	2.27 i
8	P_Mix <sup>b</sup>	1553 bcd	80.0 g	2.65 bc	15.00 ab	2.22 l
9	OV_Mix	1520 bcd	108.9 b	3.33 ab	5.75 a	2.84 b
10	MV_Mix <sup>b</sup>	1662 bc	62.7 h	3.18 ab	6.87 a	1.91 m
11	Maiorca	1432 cde	98.0 cd	3.25 ab	6.65 a	2.71 cd
12	Andriolo	1249 e	102.5 bc	2.89 abc	7.33 a	2.51 fg
13	Frassineto <sup>a</sup>	1625 bcd	104.0 bc	3.13 abc	7.20 a	2.98 a
14	Gentilrosso <sup>a</sup>	1235 e	119.3 a	*	*	3.00 a
15	Emese	1457 cde	80.8 fg	2.81 bc	10.84 ab	2.33 hi

<sup>a</sup> Only two years' data.

<sup>b</sup> Only three years' data.

than Emese, one of the components of the MV\_Mix. The tallest entries were the landraces Andriolo, Frassineto and Gentilrosso with the latter being significantly the tallest entry (nearly 120 cm) of the whole experiment. There were significant differences between the four

ICA\_EPs: the one which evolved in Sicily was, on average, nearly 20 cm shorter (P < 0.01) than the ICA\_EP which evolved in Tuscany. The population which evolved in Molise was intermediate, while the population derived from farmer's selection (ICA\_FLS) was significantly (10 cm, P < 0.05) shorter than the original population and not significantly different from ICA\_LR.

The highest level of ground cover was found in ICA\_FLS, significantly higher than the North European CCPs, P\_Mix and Emese. By contrast, the lowest level of ground cover was found among some of the shortest entries such as CYQ and CQ, although not significantly different from the ground cover of some of the tallest entries. This is evident in the case of the modern and old varieties mixtures: although significantly different for plant height, they had very similar ground cover estimates. However, there was an overall tendency for a positive but not significant association between plant height and ground cover as indicated by the correlation coefficient (r = 0.491, P = 0.075, 12 d.f.).

No significant differences were observed in the case of weed density. The highest average number of weeds/plot was observed in the plots of four of the shortest varieties, namely Emese, CY, CQ and P\_Mix, but not in the mixture of modern varieties although this is the shortest entry. We observed a tendency to a negative but not significant association between weed density and plant height, as indicated by the correlation coefficient (r = -0.503, P = 0.067, 12 d.f.).

The highest yielding entries were not always those that received the highest score as shown by the low score received by ICA\_LR, and, conversely by the high score received by one the lowest yielding entries such as Gentilrosso, one of the most popular Italian bread wheat landraces. Overall, there was no association between grain yield and

participants' score ( $r = -0.17$ ,  $P = 0.536$ , 13 d.f.).

### 3.3. Adaptation to location: the three ICA populations versus each other

In Tuscany, the population obtained by farmer's selection (ICA\_FLS) and evolving in the region in which was selected, was the highest yielding and significantly different from the other EPs except ICA\_RO. The original population evolving in the same region (ICA\_FL) was marginally (8.9 %) but neither significantly higher yielding than the population that evolved in Sicily (ICA\_LR) nor than the one which moved from Sicily to Abruzzo. In Molise and Piedmont no significant differences were observed for grain yield between the four ICA EPs.

In Sicily, the population which evolved in the region (ICA\_LR) out-yielded significantly the population which evolved in Tuscany (ICA\_FL), by 33.8 %, (Fig. 2). The population which evolved in Sicily from 2010 to 2014 and was then moved to Abruzzo (ICA\_RO), was significantly lower yielding (-46.3 %) when tested back in Sicily than the population which evolved continuously in Sicily. On the other hand, the population selected by the farmer in Tuscany (ICA\_FLS), when tested in Sicily performed as well as the population which evolved *in loco* (Fig. 2).

The population which evolved in Tuscany (ICA\_FL) was always significantly taller than the population which evolved in Sicily (ICA\_LR) except in Sicily where there were no significant height differences between the four ICA EPs (Fig. 2). The population selected by the farmer in Tuscany (ICA\_FLS) was always shorter than the populations from which it was selected (ICA\_FL) although the difference was significant only in Molise and Piedmont.

### 3.4. The three ICA populations vs. the modern variety Emese

Across locations and years, the yield advantage of the ICA populations over Emese ranged from a significant 34.0 % (ICA\_FLS) to a non-significant 6.3 % (ICA\_FL) while ICA\_RO had a similar yield to the modern variety (Table 6). One of the populations which evolved under natural selection only (ICA\_LR), had a significant yield advantage of 20.4 %.

Both in Tuscany and in Sicily, with the only exception of ICA\_RO in Sicily, the evolutionary populations yielded at least as well as, and in some cases significantly more (33.7 % in the case of ICA\_LR in Sicily, and 48.8 % in the case of ICA\_FLS in Tuscany) than the improved variety Emese (Fig. 2). ICA\_LR out yielded Emese significantly (by 28.9 %) also in Molise. All the other differences in the two regions, Molise and Piedmont, where none of the three EPs evolved, were not significant.

ICA\_FL was significantly taller than Emese in all four locations with

the largest difference, nearly 20 cm, in Piedmont (Fig. 2). The population, which evolved in Molise was also significantly taller than Emese in Tuscany and in Molise. With the exception of ICA\_LR in Piedmont, the ICA EPs were never shorter than Emese.

### 3.5. The three ICA populations vs. the three mixtures

The results of this contrast (Fig. 3) varied considerably with the location: the mixture of old varieties (OV\_Mix) and the mixture of modern varieties (MV\_Mix), but not the P\_Mix, performed as well as the best evolutionary population (ICA\_FLS) in Tuscany. In Sicily, the P\_Mix and the OV\_Mix yielded significantly less than the best two EPs, which yielded as much as the MV\_Mix. In the other two locations, there were small differences in yield between EPs and mixtures, except Piedmont where the MV\_Mix yielded about 25 % less than the best EP. On average, ICA\_LR and ICA\_FLS had a yield advantage over the MV\_Mix of 5.5 % (not significant) and 17.5 % (significant at  $P < 0.05$ ), respectively.

The mixture of old varieties was the tallest entry in all locations while the mixture of modern varieties was always the shortest (Fig. 3) although the differences with the ICA EPs were not always significant. The ICA EPs were almost always taller, sometimes considerably so, than the MV\_Mix with the only exception of ICA\_LR in Tuscany. However, they were never taller, and in some cases were significantly shorter than the OV\_Mix.

### 3.6. The three ICA populations vs the landraces

With the exception of Frassineto in Piedmont, where it was the highest yielding although not significantly different from the four ICA EPs, and of Maiorca in Molise (Fig. 4), where it was the highest yielding but not significantly higher yielding than the ICA EPs, in Tuscany and Sicily landraces were significantly lower yielding than the best EPs (ICA\_FLS in Tuscany and ICA\_FLS and ICA\_LR in Sicily). In Sicily, it was of interest to find that ICA\_FLS and ICA\_LR significantly out-yielded Maiorca, a well-known old Sicilian bread wheat landrace well adapted to the Sicilian environment, by nearly 46 %.

The tallest landrace was always Gentilrosso followed by Frassineto, which was significantly shorter only in Tuscany (Fig. 4). Maiorca and Andriolo were always similar and not significantly shorter than Frassineto. The ICA EPs were shorter, particularly the Sicilian EP (ICA\_LR) in Tuscany, Molise and Piedmont, than the landraces, and never significantly taller. The tallest EP was ICA\_FL, which in Molise and Piedmont was not significantly different from Gentilrosso.

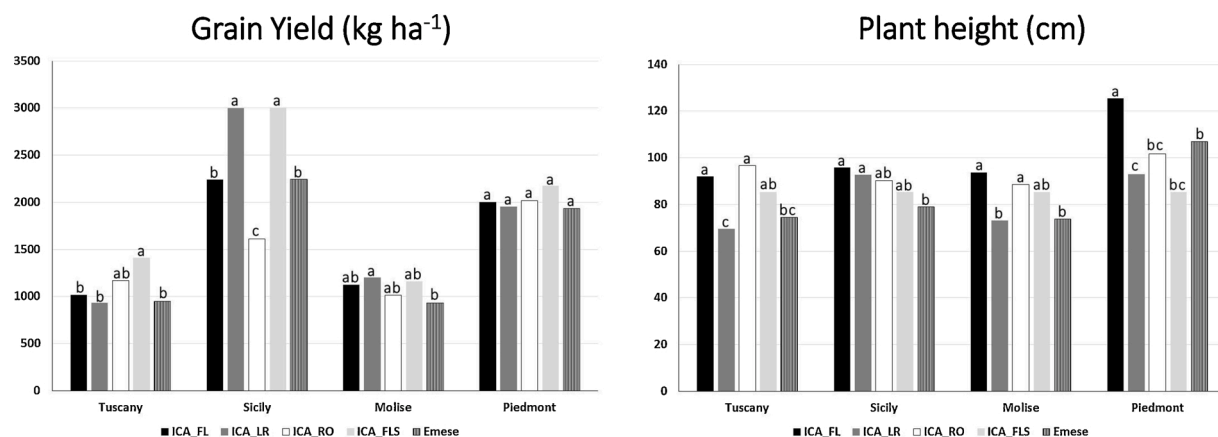


Fig. 2. Grain yields ( $\text{kg ha}^{-1}$ ) averaged over four years (only three years in Piedmont) and plant height (cm) averaged over three years (only two years in Piedmont) of four evolutionary populations and one modern variety in four locations in Italy. Bars with letters in common are not significantly different at  $P < 0.05$  based on LSD test.



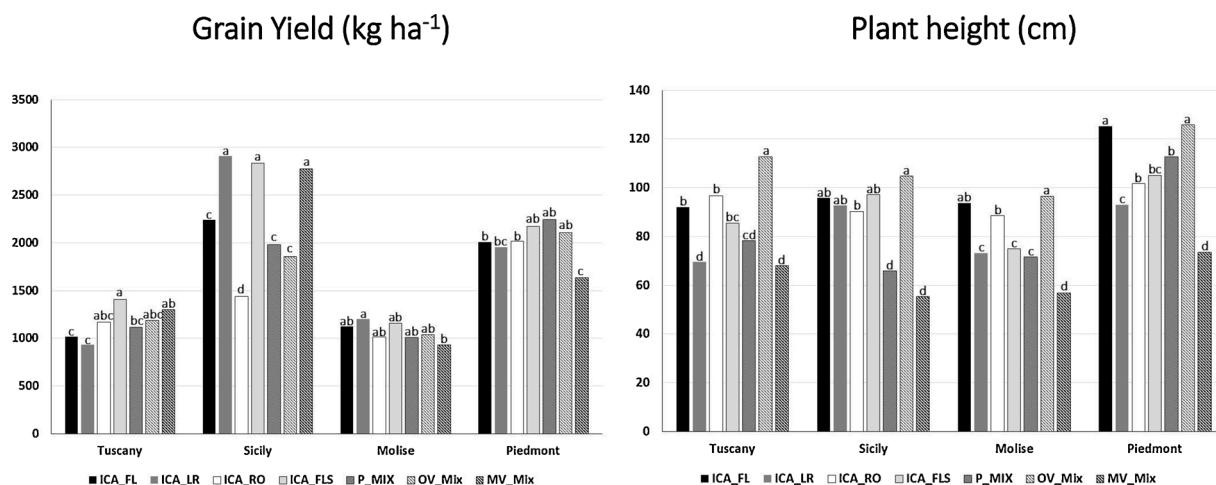


Fig. 3. Grain yields (kg ha<sup>-1</sup>) averaged over four years (only three years in Piedmont) and plant height (cm) averaged over three years (only two years in Piedmont) of four evolutionary populations and three mixtures in four locations in Italy. Bars with letters in common are not significantly different at P < 0.05 based on LSD test.

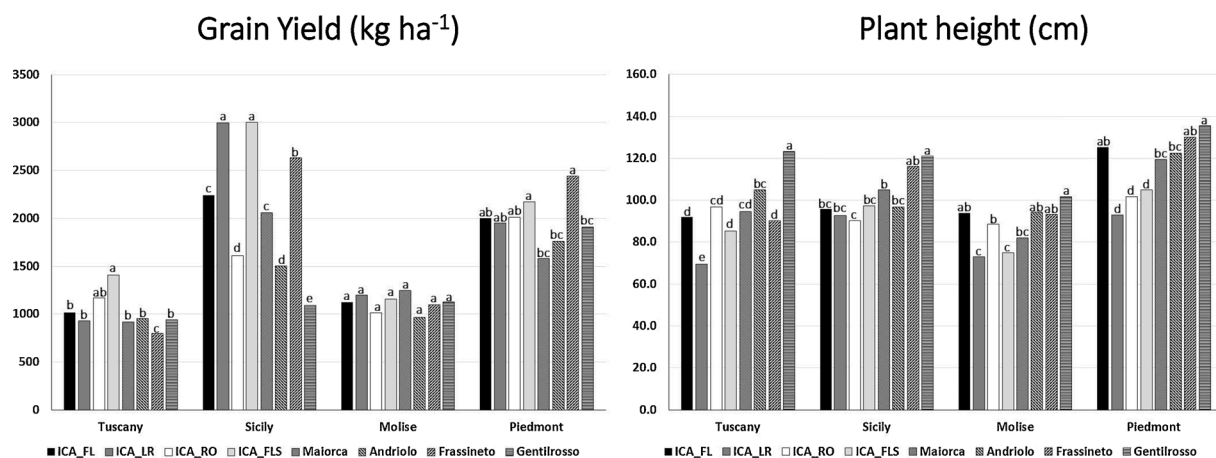


Fig. 4. Grain yields (kg ha<sup>-1</sup>) averaged over four years (three years in Piedmont) and plant height (cm) averaged over three years (only two years in Piedmont) of three evolutionary populations and four Italian landraces in four locations. Bars with letters in common are not significantly different at P < 0.05 based on LSD test.

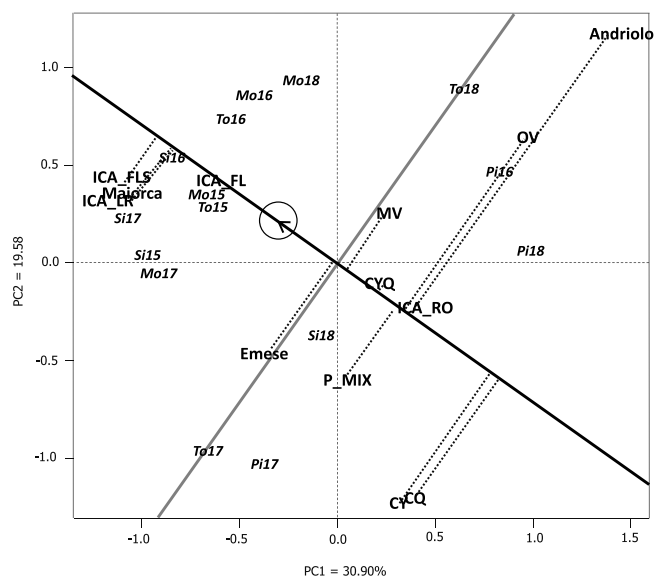
### 3.7. Yield stability

We omitted from all the stability analysis the two entries, Gentilrosso and Frassineto for which we only had two years data.

The graph resulting from the “mean and stability” options of GGE biplot (Fig. 5) confirms that the highest yielding entry was ICA\_FLS, which together with the landrace Maiorca, the EPs ICA\_LR and ICA\_FL, had the best combination of high yield and low GEI as indicated by their distance from the intersection between the mean environment axis and the biplot origin (PC1 = 0 and PC2 = 0), and their position near the mean environment axis. The CYQ and ICA\_RO populations were also very stable but at a lower yield level as they are placed at opposite sides of the intersection between the mean environment axis and the biplot origin (PC1 = 0 and PC2 = 0). Emese, with a mean yield close to the grand mean, was less stable than the ICA EPs, while Andriolo and OV\_Mix were the least stable of all the entries. The two static mixtures, OV\_Mix and MV\_Mix, had a similar grain yield, but associated with a higher stability in the case of the MV\_Mix. The dynamic mixture from Piedmont (P\_Mix) was as unstable as Emese and the MV\_Mix.

The other stability indices and superiority measures for grain yield (Table 7) show that ICA\_FLS has the lowest level of static stability (high value of EV), the highest cultivar superiority (lower value of CS), the lowest mean rank - indicating a consistent high yield in most environments as shown also by its low standard error -, and the highest yield

reliability. The EP from which it was selected, ICA\_FL, has a significantly lower EV, and ranks 3<sup>rd</sup>, 2<sup>nd</sup> and 6<sup>th</sup> for cultivar superiority, mean rank and reliability index, respectively. Of the other two ICA\_EPs, one (ICA\_LR) had a similar response to cultivation environments as ICA\_FLS, namely a high EV, a high cultivar superiority (lower value), a low mean rank and a high yield reliability. ICA\_RO is different, being the one with the highest static stability but also the one with the lowest cultivar superiority and the highest rank. As a group, the ICA\_EPs appear more stable than the three North European CCPs (CYQ, CY and CQ), which had a similar EV as ICA\_LR and ICA\_FLS but generally lower cultivar superiority, higher mean rank and lower yield reliability. The mixtures differed from each other, with the OV\_Mix having a low EV, similar to the two ICA\_EPs with the lowest EV, and a high yield reliability. The MV\_Mix had the second highest yield reliability index and also a significantly lower cultivar superiority index than the other two mixtures. The two landraces Maiorca and Andriolo were also markedly different from Andriolo being less stable, as indicated by the high mean rank and the low yield reliability index as already shown by the GGE biplot, and with the lowest cultivar superiority of all entries. The modern variety Emese had a low static stability similar to the MV\_Mix, a low cultivar superiority, a high rank with the largest standard error and a low yield reliability index.



**Fig. 5.** GGE biplot of standardized grain yields of 13 entries (seven evolutionary populations, three mixtures, two landraces and one modern variety) grown for four years in 4 locations (one location was missing in 2015). Abbreviation for entries are as in Table 1; To = Tuscany; Si = Sicily; Mo = Molise; Pi = Piedmont; 15, 16, 17 and 18 indicate 2015, 2016, 2017 and 2018, respectively (Gentilrosso and Frassineto were omitted because only two years data were available for these two entries).

### 3.8. Evidence and speed of divergent selection

The GGE biplot of the 2015 BLUEs of grain yield and plant height (Fig. 6) shows that the two populations which evolved in Tuscany (ICA\_FL by natural selection and ICA\_FLS by natural plus artificial selection) were similar both in yield level and in stability. They were also very close to Emese for both grain yield and plant height. However, there was already a clear divergence between ICA\_FL and ICA\_FLS on one side and ICA\_LR on the other. The latter had the highest expression of both yield and stability only in the location where it evolved (ICA\_LR and Si15 lie at the same side of the mean environmental axis). In Sicily, ICA\_LR was also significantly higher yielding (+24.9 %) than Emese.

In Sicily, the grain yield of ICA\_LR was significantly ( $P < 0.05$ ) higher than that of both ICA\_FL (+26.4 %) and ICA\_FLS (+34.4 %), while the grain yield of the three populations did not differ either in Tuscany or in Molise where they did not differ from Emese.

In the case of plant height, ICA\_LR was significantly ( $P < 0.05$ )

shorter than ICA\_FL both across locations and in Tuscany and Molise. In Sicily, it was taller, but not significantly different from either the two Tuscany populations.

### 3.9. Participants' preferences

In addition to being the highest yielding entry, the ICA\_FLS population was also one of the entries that participants liked most even though not with the high level of unanimity (shown by the high distance from the mean environmental axis) as was the case with ICA\_FL and Maiorca, which were the next best preferred entries (Fig. 7).

The entry with the highest preference, after excluding Gentilrosso and Frassineto from the analysis, was the mixture of old varieties (OV\_Mix), which was significantly lower yielding than ICA\_FLS and ICA\_LR, but had a significantly higher participants' score (Table 6) than the two EPs. The mixture of modern varieties (MV\_Mix) and the (P\_Mix) were the entries that participants, on average, liked the least.

There was no evidence of differences in preference associated with gender (Fig. 8), while preferences changed more substantially with the year within the same location. For example in Tuscany (Fig. 8A), the mixture of old varieties (OV\_Mix) received the highest score in 2017 and in 2018, while ICA\_FLS and Maiorca were the most liked entries in 2015 and 2016, respectively. In Tuscany, there was a large difference between the high score given to the population that evolved in that region (ICA\_FL) and the low score given to the population that evolved in Sicily (ICA\_LR). The reverse was only partially true in Sicily where the highest score were given to the population that evolved in Sicily (ICA\_LR) and to the selection made by the farmer in Tuscany. However, even the population that evolved in Tuscany received a score higher than the grand mean at least in 2017 and 2018. In Molise, there was a preference by both women and men for the ICA\_FLS and ICA\_FL in 2015 and 2016; in 2017 and 2018 the preference, by both women and men, changed for the mixture of old varieties (OV\_Mix). In Piedmont, where the experiment was conducted only for the last three years, Andriolo and ICA\_FLS received the highest score in 2016, while Gentilrosso and the mixture of old varieties (OV\_Mix), which were not included in the 2015 experiments, received the highest score in both 2017 and 2018. Piedmont was the only location where the ICA\_RO received a score higher than the grand mean.

## 4. Discussion

In this paper we reported the results of a multi environment experiment using different genetic material ranging from mixtures, evolutionary populations, landraces and one modern variety. The main objective was to test the hypothesis that natural selection in different

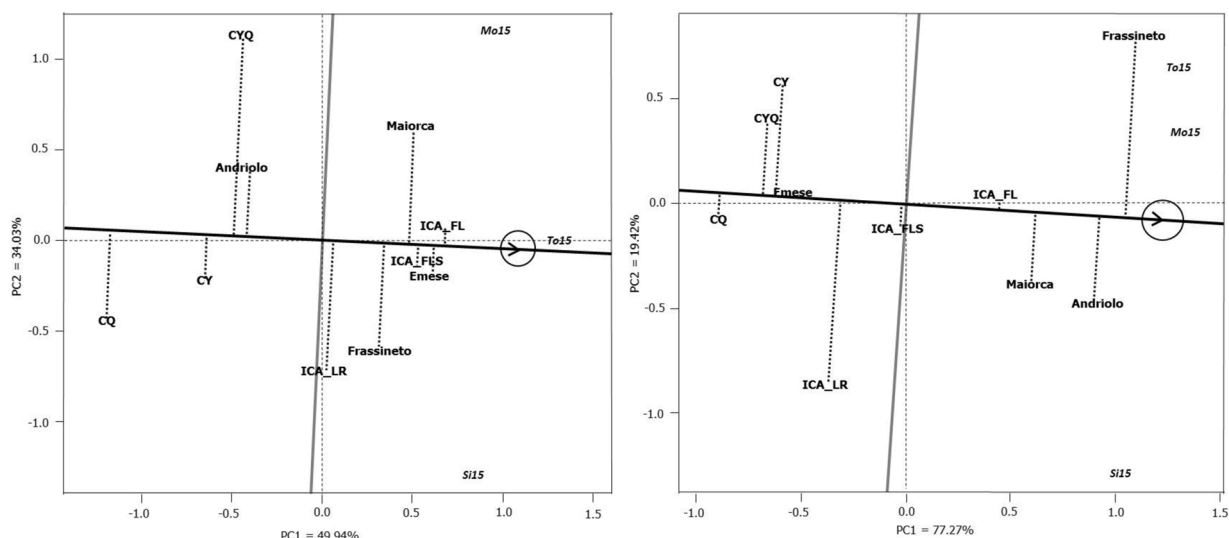
**Table 7**

Stability expressed as Environmental Variance (EV), Cultivar Superiority (CS), Mean Ranks (MR), index of reliability (IR), and their respective ranks, of grain yield of 13 entries (seven evolutionary populations, three mixtures, two landraces and one modern variety) grown for four years in 4 locations (one location was missing in 2015). Abbreviation for entries are as in Table 1. Values with letters in common are not significantly different at  $P < 0.05$  based on LSD test.

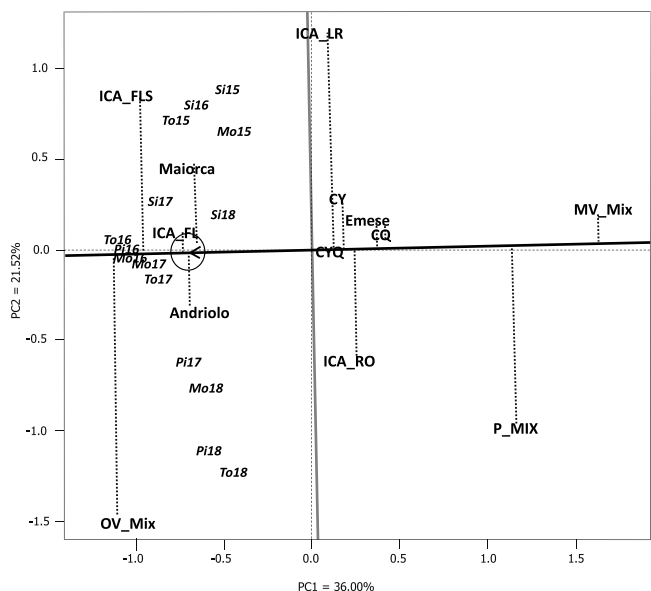
Entry <sup>a</sup>	Abbreviation	EV	Rank	CS	Rank	Mean Rank $\pm$ s.e.	Rank	IR	Rank
1	ICA_FL	97.13 a	5	36.00 c	3	6.13 $\pm$ 0.78	2	721	6
2	ICA_LR	140.01 b	12	27.37 b	2	6.27 $\pm$ 1.05	3	760	4
3	ICA_RO	69.06 a	1	66.28 h	12	8.58 $\pm$ 0.64	12	726	5
4	ICA_FLS	162.22 b	13	6.48 a	1	3.53 $\pm$ 0.81	1	883	1
5	CYQ	98.42 ab	6	54.05 ef	9	7.67 $\pm$ 0.94	9	617	8
6	CY	114.15 ab	9	43.49 d	4	7.53 $\pm$ 0.95	8	556	11
7	CQ	134.60 ab	11	50.29 e	8	8.93 $\pm$ 1.02	13	400	13
8	P_Mix <sup>b</sup>	131.48 ab	10	56.82 fg	10	6.92 $\pm$ 1.12	4	590	9
9	OV_Mix	73.39 a	2	60.91 g	11	7.25 $\pm$ 1.02	6	800	3
10	MV_Mix <sup>b</sup>	101.76 ab	7	46.03 de	5	6.92 $\pm$ 1.22	5	815	2
11	Maiorca	85.34 a	3	46.35 de	6	7.27 $\pm$ 1.07	7	656	7
12	Andriolo	91.54 ab	4	84.24 i	13	8.53 $\pm$ 1.30	11	445	12
15	Emese	111.29 ab	8	49.52 e	7	8.13 $\pm$ 1.45	10	571	10

<sup>a</sup> Entries 13 and 14 were omitted because only two years data were available.

<sup>b</sup> Only three years' data.



**Fig. 6.** Speed of evolution: GGE biplot of standardized grain yield (left) and plant height (right) of the entries evaluated in 2015 in 3 locations. Abbreviation for the entries are as in Table 1; environments are in italics: To = Tuscany; Si = Sicily; Mo = Molise, respectively.



**Fig. 7.** GGE biplot of standardized participant evaluation given to 13 entries (seven evolutionary populations, three mixtures, two landraces and one modern variety) grown for four years in 4 locations (three years in Piedmont). Abbreviation for entries are as in Table 1. Environments are in italics: To = Tuscany; Si = Sicily; Mo = Molise; Pi = Piedmont; 15, 16, 17 and 18 indicate 2015, 2016, 2017 and 2018, respectively.

locations led to a differentiation between populations for important agronomic traits such as grain yield and plant height. Additional objectives were to determine whether the evolution led to commercially competitive levels of grain yield and to verify the reaction of potential users.

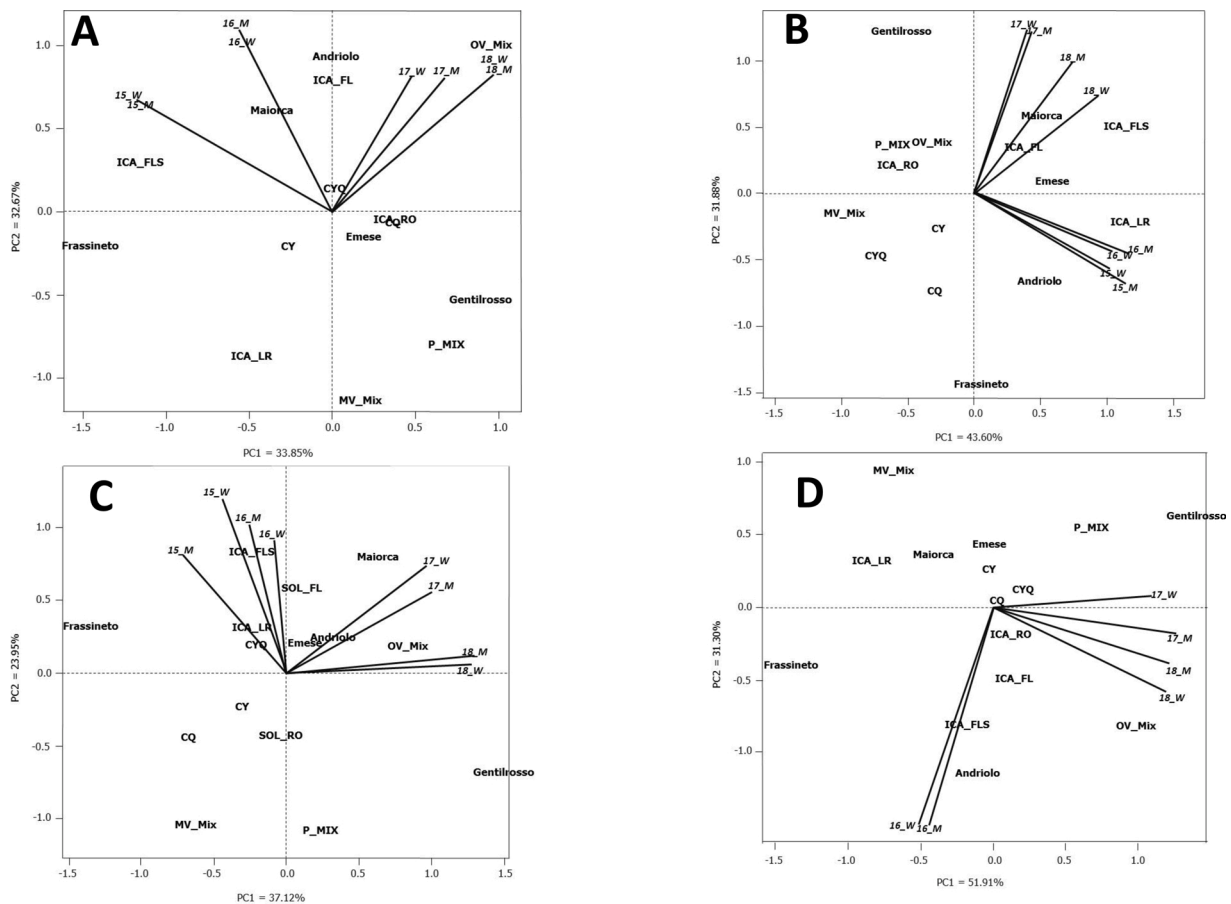
The very low average yield in Tuscany in 2015 was likely due to the delayed planting (January) because of bad weather and to poor soil fertility since wheat had been grown on the same field the previous year: the very low average yield in Piedmont in 2018 was possibly due to a prolonged period of wet and cold weather after planting, causing poor emergence and heavy weed infestation which created considerable problems at harvesting. This large variation in yield from one year to the next, associated with short term weather variation within the same cropping season is becoming more and more familiar to farmers and is

also reflected in the variation in plant height, particularly in Tuscany. It is a global issue as climate variability accounts for about one third of yield variability, with a maximum of 51 % in the case of wheat in Western Europe (Ray et al., 2015) and it is expected to become more frequent in Europe with climate change (Trnka et al., 2014). One possible way to cope with such increased variability is to exploit the population buffering of heterogeneous material (Allard, 1999).

While the significance of the interactions between entries and both years and locations is common in multi environment experiments in the case of grain yield, the significance of the mean square between entries ( $P = 0.031$ ) was not expected. In fact, selection theory tells us that the magnitude of directional selection, presumably different from location to location, is sufficient to produce rapid microevolutionary changes (Kingsolver and Diamond, 2011). Therefore, years of evolution in the same location should have resulted in increasing the level of specific adaptation, and consequently smoothing the overall difference between entries across locations. The significant difference in grain yield across locations might be explained by the unexpected performance of ICA\_FLS. This sub-population, perhaps because of its sharing similar phenology (data not shown) and plant height with the population which evolved in Sicily, or because it retained a higher genetic diversity than expected in material derived by artificial selection, was able to perform well in all locations.

However, although we do have information on precipitation, temperature, soil characteristics and agronomic management in the different locations, it is always difficult to understand precisely the causes of selection unless the experiment is designed specifically to address this question (Wade and Kalisz, 1991).

The evolution of plant height in the ICA\_EPs is of particular interest. In fact, one critical point that has been expressed about evolutionary plant breeding is based on the evidence that natural selection favours tall plants (Knapp et al., 2020), thus shifting the population in a direction opposite to the reduction of plant height, which has been one of the main objectives of modern plant breeding (Denison et al., 2003). Our results confirm that this is the case for the four ICA\_EPs, which were significantly taller than the modern variety (with few exceptions in some locations) and of the mixture of modern varieties; however, they were significantly shorter than the studied landraces, which are commercially grown in Italy, mainly under organic conditions, and of the mixture of old varieties. This might explain the spontaneous adoption of EPs which emerged alongside this experimentation. Indeed, under organic no-herbicide conditions, a relatively tall crop may be of advantage (Knapp et al., 2020) because of its ability to suppress the development of



**Fig. 8.** GGE biplot of standardized women and men evaluation given to 15 entries (seven evolutionary populations, three mixtures, four landraces and one modern variety) grown for four years in 4 locations (three years in Piedmont). Abbreviation for entries are as in Table 1. A = Tuscany; B = Sicily; C = Molise; D = Piedmont; 15, 16, 17 and 18 indicate 2015, 2016, 2017 and 2018, respectively; W = women evaluation, M = men evaluation.

weeds as often informally reported by the several Italian farmers currently growing these evolutionary populations as commercial crops.

In this experiment we did not find evidence for a significant weed-suppressive ability nor for a higher ground cover associated with plant height. This is likely to be due, particularly in the case of weed density, to the unsuitability of small plots or to the use of too few replications to achieve a sufficient precision in the measure of this trait since weeds are most often unevenly distributed.

#### 4.1. Adaptation to location: the three ICA populations versus each other

Several field studies of natural populations of plants and animals have shown abundant evidence for directional selection on morphology and life history traits (number, size and sex ratio of offspring, the timing of reproduction, age and size at maturity and growth pattern, longevity) (Kingsolver et al., 2001). These studies have shown that temporal fluctuations in magnitude and direction of directional selection are common (Siepielski et al., 2009) and that the magnitude of directional selection is sufficient to produce rapid microevolutionary changes in many populations. Directional natural selection is forecasted to increase with more frequent droughts and rising temperatures (Exposito-Alonso et al., 2019).

In the experiment described in this paper we found evidence of directional selection on two important quantitative traits such as grain yield and plant height as shown by the significant difference between two evolutionary populations which evolved for four years prior to the experiment in two different Italian regions.

In one of the regions where the ICA\_EPs evolved longer, namely Sicily, there was evidence of significant divergent evolution for grain

yield. This was significant only in Sicily, likely because in this region there was a more consistent pattern of both temperature and rainfall sufficient to apply directional selection to the population evolving there. However, as mentioned earlier, unless selection experiments are specifically designed to detect the causes of selection, we can only speculate what the causes could be. One additional complication in interpreting the results of selection experiments is that the fitness of an individual, on which natural selection is acting, is the result of the interaction between the phenotype and the environment and is not an intrinsic feature of either one (Wade and Kalisz, 1991). We may speculate that with the more variable climatic conditions in the first four years of cultivation in Tuscany (2010–2013, Fig. 1A), the relationships between phenotype and fitness may have changed, and therefore in that location, directional selection may have been alternating in time, thus reducing the cumulative effect of selection (Siepielski et al., 2009).

In the case of plant height, the evidence for divergent selection was even stronger with ICA\_FL becoming significantly taller than ICA\_LR and expressing the height advantage in three (Tuscany, Molise and Piedmont) of the four experimental sites. This is likely to be associated with the high heritability of the trait (see for example Reynolds et al., 2017).

This paper provides important additional experimental evidence of populations' capacity to evolve, thus confirming the results of past research on evolutionary populations and dynamic mixtures, which showed that they do evolve and become higher yielding (Suneson, 1956; Allard and Hansche, 1964; Rasmusson et al., 1967; Patel et al., 1987; Soliman and Allard, 1991). Our research has provided the additional evidence that artificial selection, in our case farmer's' selection within a population, which is evolving on their farm, can speed up the adaptive evolution generating a selected evolutionary population higher yielding



that the one which evolved in the same location under the sole effect of natural selection. Farmers' selection was also effective in counteracting the effect of natural selection and in reducing the plant height of ICA\_FLS to the level of ICA\_LR.

The effectiveness of farmers' selection in this experiment points to a possible use of evolutionary populations as source material to select improved subpopulations, as in this case, or pure lines as suggested by Döring et al. (2015). Indeed, this is informally already happening, with other farmers requesting seed and cultivating it (hence, most likely, adapting it) in their own fields.

The efficiency of selection for grain yield within an evolutionary population was also shown by Suneson (1956): working with four barley composite crosses, a line was identified after 20 generations that out yielded Atlas 46 (a widely grown variety at the time of the experiment) by 37 %, and in a later generation (F<sub>24</sub>), three top selections showed a 56 % greater yield than Atlas 46. These represent genetic gains similar or even greater than those commonly reported (Cobb et al., 2019). Suneson's results as well as those obtained with the ICA\_FLS EP for grain yield and plant height, confirm that in a self-pollinated crop, selection within evolutionary populations is more efficient when delayed until homozygosity is high and hence, heritability is also high (Allard, 1999). In our case, the original population was a mixture of three different segregating populations (in unknown proportions), and therefore the homozygosity was higher than what the number of cropping seasons previous to the selection operated by the Tuscan farmer would suggest. However, although this might explain the response to selection both in grain yield and plant height, it does not explain the higher stability of ICA\_FLS.

#### 4.2. The three ICA populations vs. the modern variety Emese

This contrast is interesting because it responds to the question of whether the EPs are able to evolve and to reach yield levels that are commercially relevant. The yield advantages of the best EPs over the improved variety observed in this study ranged from 34.0 % to 6.3 % (not significant) across locations and years. The highest yield advantage was obtained with the combination of natural and artificial selection (ICA\_FLS) and reached its maximum expression (48.8 %) in the region in which selection was practiced, namely in Tuscany, but also in Sicily (33.8 %) and to a lesser extent in Molise (24.3 %) and Piedmont (12.4 %). On the other hand, the population which evolved in Sicily as a result of natural selection only, showed the maximum yield advantage over Emese (33.7 %) in Sicily, yield advantage which decreased to 28.8 % in Molise and disappeared in the other two regions.

In a classical experiment with a composite cross of barley in Canada, Patel et al. (1987) concluded that "maximum response to selection is expected when the material is advanced under directional selection pressure in a single, unchanging environment" indicating that this is the most efficient way to exploit the evolutionary potential of evolutionary populations. Our data show this clearly: the positive effect of adaptive evolution using Emese, unable to evolve, as a benchmark, disappeared when the EPs were compared to Emese in the two regions (Molise and Piedmont) where the two EPs did not evolve prior to the experiment. However, one implicit assumption in Patel's conclusions is that in a single environment selection pressure is consistent over time: in the case of yearly-variable environmental conditions, as have occurred in the first few years of evolution in Tuscany, or as may occur by moving the evolutionary population from one location to another, stabilizing selection may become more important than directional selection, thus reducing the cumulative effects of selection (Siepielski et al., 2009).

By deploying the same evolutionary population with a large genetic diversity in a range of locations representing the target population of environments, we can expect the development of as many distinct populations as there are locations, given that they are characterized by different types of directional selection. This is a cost-effective way of exploiting repeatable GL interactions with a beneficial increase of agrobiodiversity and the possibility, as mentioned earlier, of selecting

improved populations or pure lines each specifically adapted to the location in which the population evolved or the selection is conducted.

#### 4.3. The three ICA populations vs. the three mixtures

Research has shown that the evolutionary potential of EPs is higher than that of mixtures (Patel et al., 1987) and therefore this contrast was considered of interest even though it could be possibly biased by the fact that the three mixtures were added only in 2016. One additional reason of interest for this contrast is that mixtures are easier to construct independently by farmers who either cannot make crosses or cannot count on the cooperation of scientists to make the crosses necessary to produce an evolutionary population. The P\_Mix was the only dynamic mixture, and with the caution imposed by the availability of only three years data and despite the small number of components (seven), this mixture yielded as much as the more complex EPs, but only in Piedmont. In that region, it out yielded significantly the static MV\_Mix but not the other static mixture OV\_Mix made with three components, all in common with P\_Mix.

The MV\_Mix (static mixture) contains some cultivars widely used in organic agriculture, such as Bologna, Emese, also used in the experiment as individual entry, and the high yielding variety A416. It was therefore of interest to find that the yield of some of the EPs varied from being similar to being slightly superior to this particular mixture which, being a static mixture, can be considered as an additional "modern control". Taken together with the results of the comparison with Emese, this points once again to evolutionary plant breeding as an efficient breeding method, which combines the increase in grain yield with an increase in agrobiodiversity due to divergent evolution. In addition, it is a cost-effective way of increasing grain yield because all it requires is planting, harvesting and storing part of the harvested seed for the following cropping season.

#### 4.4. The three ICA populations vs. the landraces

The recently increased interest in cultivating bread wheat landraces and old varieties, and in the commercialization of products derived from them in several regions in Italy and elsewhere in Europe, could represent a challenge for the spread and appreciation of EPs. It was therefore of interest to compare the EPs with potential commercial competitors and to find that even in Piedmont, where one of the popular old varieties was the top yielder, some of the EPs were at the same yield levels. Therefore, as in the case of Emese and of the mixture of modern varieties, some of the EPs have a yield level comparable or higher than that of other commercially successful varieties.

#### 4.5. Yield stability

Even before the experiment, farmers reported that one of the best attributes of the bread wheat EP was its stability over time. Obviously, temporal stability is the type of stability that an individual farmer can easily detect. In the experiment, the two ICA\_EPs which evolved continuously in the respective regions appear to carry this property, which was confirmed using different estimates of stability, including a reliability index. This was expected because continued natural selection in the same locations builds local (specific) adaptation, which translates into high temporal adaptation and low spatial adaptation. The experiment revealed that one cycle of artificial selection not only significantly increased grain yield above the level reached with natural selection only, but changed the adaptation from typically specific to much wider as the ICA\_FLS yielded more than ICA\_FL also in Sicily where it did not differ significantly from the locally adapted ICA\_LR. ICA\_FLS ranked first for three of the four estimates of stability indicating a higher temporal and spatial stability. A study about stability and adaptation of wheat populations developed by farmers in France through participatory plant breeding found that these populations, on which both natural and

artificial selection were applied, had more stable yields over time and space than commercial varieties (van Frank et al., 2020). Döring et al. (2015) found an advantage of CCPs and mixtures over the mean of their components in terms of cultivar superiority.

The superiority of ICA\_FLS and CYQ is interesting because, although of very different origin, they both derive from selection, in the first case from within an evolving EP with a large genetic diversity, the second from the selection of the parents of a CCP. The CYQ is also the most diverse of the three North European CCPs and was also the highest yielding and the most stable of the three CCPs.

A clear difference in stability between populations and mixtures was not evident as it was in the experiment of Allard (Allard, 1961); in our case, stability (or lack of it) was probably associated with the level of specific adaptation.

#### 4.6. Evidence and speed of divergent selection

Results discussed so far, albeit showing evidence of divergent evolution, do not allow to estimate how quickly the divergence occurred, an issue of particular interest to farmers. In theory, the speed of divergent evolution depends on the magnitude of genetic diversity available in the initial population for adaptive traits, the heritability of those traits and the magnitude and consistency of directional selection (Kingsolver et al., 2001).

The results obtained during the first year of the experiment (2015) suggest that sometime during the first four years of its cultivation, the Sicilian population (ICA\_LR) evolved to give, in the year our experiment started, a significantly ( $P < 0.05$ ) higher yield than the two Tuscan populations in Sicily, while in Tuscany the differences were not significant nor was the difference between ICA\_FLS and the other entries. Therefore the superiority of ICA\_FLS discussed earlier, must have evolved later in the course of the experiment. Similarly, at the beginning of the experiment, ICA\_LR was already higher yielding ( $P < 0.05$ ) than Emese, but, also in this case, only in Sicily.

Contrary to grain yield, the divergence for plant height was detectable only in Tuscany and only in the population which evolved through natural selection, which was significantly taller than the population which evolved in Sicily and the Tuscan population selected by the farmer. The latter was likely due to the preference by the Tuscan farmer for a slightly shorter plant during their selection in 2012.

#### 4.7. Participants' preferences

The importance of gender and social targeting in plant breeding has been recently discussed in two CGIAR workshops in the framework of the Gender & Breeding Initiative (<http://www.rtb.cgiar.org/gender-breeding-initiative/>), which was translated in a working paper (Tufan et al., 2018) compiling a set of case studies that consider gender in various steps of a plant breeding cycle. This is in recognition that considering gender is important in developing varieties or testing new material to lead to equitable benefits (Tufan et al., 2018). What is particularly compelling about these cases is that they not only provide evidence that men and women may have different trait preferences, but they also illustrate steps taken by breeding programs to address these issues. Although most of the evidence comes from developing countries, issues of gender inequality and (dis)empowerment may arise also in developed countries (Wilbur, 2014). Given that in plant breeding a product profile is a "package of traits needed to replace a specific reference or target variety" (Cobb et al., 2019), the inclusion of a gender-responsive participatory component in the evaluation of breeding materials is a pre-requisite for the development of a gender-responsive product profile.

Although in our experiments we were not able to detect gender related differences in the preference for the different types of material, such lack of differences cannot be taken for granted. It is therefore important, given the role of women in agriculture, that breeders use a

gender-responsive approach because this may lead to greater adoption of varieties (Beuchelt, 2016).

Our results about farmers' preferences indicate a reasonably consistent evaluation of some of the entries, but a more environment (year)-dependent evaluation for others. This is suggested by the low probability levels (just below the significance level) of the interactions between entries and years, although the same interaction was highly significant in the case of grain yield. Therefore, grain yield was not the main criteria used in evaluating the different entries, confirming the phenomenon well described in literature, by which adoption (hence preference) is not necessarily driven by yield (Sall et al., 2000; Jalleta, 2004; Aw-Hassan et al., 2008).

In other experiments using farmers' evaluation of large numbers of breeding materials in contrasting locations, we found that the location was an important factor in driving the evaluation, but not necessarily for all traits (Ceccarelli et al., 2000).

## 5. Conclusions

Overall, the experiment showed that one evolutionary population grown continuously in different locations evolved into locally adapted populations with significant differences in important quantitative traits such as grain yield and plant height. Some of the populations resulting from natural selection yielded as well as or better than commercial varieties in the location in which they evolved, while the population resulting from a combination of natural and artificial selection showed both spatial and temporal stability.

Our findings allow to draw the following conclusions of relevance for plant breeding. Firstly, the evolutionary populations represent a dynamic response to climate change, a complex breeding objective for both conventional and molecular breeding. Indeed, 1) changes in temperature and rainfall are likely to vary from location to location; 2) climate change is not only about temperature and rainfall, but those changes affect the distribution and outbreak of pests (Heeb et al., 2019), and in particular affect the spectrum of insects (Zavala et al., 2008; Deutsch et al., 2018), diseases (Newton et al., 2011; Pautasso et al., 2012) and weeds (Ziska and Dukes, 2010; Colautti and Barrett, 2013; Matzrafi et al., 2016); 3) extreme weather events can influence the interactions between crops and pests in an unpredictable way (Rosenzweig et al., 2001). All this evidence points at climate change as an extremely complex and evolving problem, which requires an evolving solution. Evolutionary populations, with their capacity to evolve in response to both biotic and abiotic stresses, provided they possess sufficient initial genetic diversity, is likely to be a quick, cost-effective and adaptive solution to such a complex problem with the additional advantage of increasing yield gains resulting from a combination of natural and artificial selection. Secondly, the inclusion of participatory evaluation when testing novel genetic material such as evolutionary populations make easier to develop innovations more likely to be adopted; this is an important advantage considering that lack of adoption had been one of the major problems in conventional plant breeding (Ceccarelli, 2015).

An additional advantage of evolutionary populations, not studied in this experiment but emerging from a large body of research, is their greater tolerance to diseases, thanks to their intrinsic diversity which avoids a too narrow selection pressure which favors the evolution of resistance (Bøhn and Millstone, 2019). This makes them particularly suited to organic systems because they represent an ecological solution to disease control.

One of the consequences of the ability of EPs to adapt relatively quickly to different environments was the rapid spreading of the bread wheat EP to nearly every region in Italy, where individual farmers, cooperatives and associations started using it directly to either sell the flour or use it in local bakeries, invariably experiencing unexpected commercial success. This was greatly facilitated by the Commission Implementing Decision of 18 March 2014 pursuant to Council Directive 66/402/EEC, which made it possible, in Europe, to experimentally

market heterogeneous material of different cereals including wheat up to 29 February 2021. This opening allowed for the official registration of these populations, creating the opportunity for farmers to legally buy and sell seed of heterogeneous materials, but most importantly to become the producers of their own seed throughout the years.

### CRedit authorship contribution statement

**Riccardo Bocci:** Funding acquisition, Investigation, Data curation, Supervision, Project administration, Writing - review & editing. **Bettina Bussi:** Funding acquisition, Investigation, Data curation, Supervision, Project administration, Writing - review & editing. **Matteo Petitti:** Funding acquisition, Investigation, Data curation, Supervision, Writing - review & editing. **Riccardo Franciolini:** Funding acquisition, Investigation, Data curation, Supervision. **Virginia Altavilla:** Funding acquisition, Investigation, Data curation, Supervision. **Gea Galluzzi:** Funding acquisition, Investigation, Data curation, Supervision, Writing - review & editing. **Paolo Di Luzio:** Funding acquisition, Investigation, Data curation, Supervision. **Paola Migliorini:** Funding acquisition, Investigation, Data curation, Supervision, Writing - review & editing. **Sandra Spagnolo:** Funding acquisition, Investigation, Data curation, Supervision. **Rosario Floriddia:** Resources. **Giuseppe Li Rosi:** Resources. **Modesto Petacciato:** Resources, Writing - review & editing. **Vincenzo Battezzato:** Resources. **Andrea Albino:** Resources. **Giovanni Faggio:** Resources. **Carlo Arcostanzo:** Resources. **Salvatore Ceccarelli:** Investigation, Data curation, Supervision, Conceptualization, Methodology, Formal analysis, Software, Validation, Visualization, Writing - original draft, Writing - review & editing.

### Declaration of Competing Interest

The authors report no declarations of interest.

### Acknowledgments

The authors acknowledge ICARDA' bread wheat breeder for making available his breeding material, the financial support of the European Union, first through the SOLIBAM project, which facilitated the initial spreading of the evolutionary populations and later through the project DIVERSIFOOD (grant agreement no. 633572), which supported the execution of the experiments. Thanks are due to the University of Florence, particularly to prof. S. Benedettelli, Ms. B. Trucchi and their PhD students, for assistance in planting, harvesting and data collection. Thanks to the National University of Ireland, Galway, in the persons of Prof. Charles Spillane and Dr. Peter McKeown for supporting Matteo Petitti's CCAFS MSc final project on this topic.

### References

Allard, R.W., 1961. Relationship between genetic diversity and consistency of performance in different environments. *Crop Sci.* 1 (2), 127–133.

Allard, R.W., 1999. *Principles of Plant Breeding*, 2nd edition. John Wiley & Sons, New York, p. 254.

Allard, R.W., Hansche, P.E., 1964. Some parameters of population variability and their implications in plant breeding. *Adv. Agron.* 16, 281–325.

Annicchiarico, P., 2002. Genotype X Environment Interactions: Challenges and Opportunities for Plant Breeding and Cultivar Recommendations, Paper 174. Food and Agricultural Organization, Rome.

Arief, V.N., DeLacy, I.H., Basford, K.E., 2020. Design and analysis of multi-year field trials for annual crops. In: Kang, M.S. (Ed.), *Quantitative Genetics, Genomics and Plant Breeding*, pp. 178–193. Wallingford, UK.

Aw-Hassan, A., Mazid, A., Salahieh, H., 2008. The role of informal farmer-to-farmer seed distribution in diffusion of new barley varieties in Syria. *Exp. Agric.* 44, 413–431.

Baethgen, W.E., 2010. Climate risk management for adaptation to climate variability and change. *Crop Sci.* 50, S-70–S-76.

Bedó, Z., Szűcs, L., Láng, L., Szűcs, L., Veisz, O., Karsai, I., Juhász, A., Rakszegi, M., Vida, Gy., Szűcs, P., 2001. Six, early maturing, Martonvásár-bred winter wheat varieties were registered in 2000. *Ann. Wheat Newsletter* 47, 56–59.

Bertholdsson, N.O., Weedon, O., Brumlop, S., Finckh, M.R., 2016. Evolutionary changes of weed competitive traits in winter wheat composite cross populations in organic and conventional farming systems. *Eur. J. Agron.* 79, 23–30.

Beuchelt, T.D., 2016. Gender, social equity and innovations in smallholder farming systems: pitfalls and pathways. In: Gatzweiler, F., von Braun, J. (Eds.), *Technological and Institutional Innovations for Marginalized Smallholders in Agricultural Development*. Springer International, p. 442.

Bocci, R., 2009. Seed Legislation and agrobiodiversity: conservation varieties. *J. Agric. Environ. Int. Dev.* 103 (1/2), 31–49.

Böhn, T., Millstone, E., 2019. The introduction of thousands of tonnes of glyphosate in the food chain—An evaluation of glyphosate tolerant soybeans. *Foods* 8 (669).

Ceccarelli, S., 2015. Efficiency of plant breeding. *Crop Sci.* 55, 87–97.

Ceccarelli, S., Grando, S., Tutwiler, R., Baha, J., Martini, A.M., Salahieh, H., Goodchild, A., Michael, M., 2000. A methodological study on participatory barley breeding. I. Selection phase. *Euphytica* 111, 91–104.

Cobb, J.N., Juma, R.U., Biswas, P.S., Arbelaez, J.D., Rutkoski, J., Atlin, G., Hagen, T., Quinn, M., Ng, E.H., 2019. Enhancing the rate of genetic gain in public-sector plant breeding programs: lessons from the breeder's equation. *Theor. Appl. Genet.* 132, 627–645.

Colautti, R.I., Barrett, S.C.H., 2013. Rapid adaptation to climate facilitates range expansion of an invasive plant. *Science* 342, 364–366.

Coomes, N.E., 2009. *Design Search T.*

Costanzo, A., Barberi, P., 2016. Field scale functional agrobiodiversity in organic wheat: effects on weed reduction, disease susceptibility and yield. *Eur. J. Agron.* 76, 1–16.

Cullis, B.R., Smith, A.B., Coomes, N.E., 2006. On the design of early generation variety trials with correlated data. *J. Agric. Biol. Environ. Stat.* 11, 381–393.

Denison, R.F., Kiers, E.T., West, S.A., 2003. Darwinian agriculture: when can humans find solutions beyond the reach of natural selection? *Q. Rev. Biol.* 78 (2), 148–168.

Deutsch, C.A., Tewksbury, J.J., Tigchelaar, M., Battisti, D.S., Merrill, S.C., Huey, R.B., Naylor, R.L., 2018. Increase in crop losses to insect pests in a warming climate. *Science* 361 (6405), 916–919.

Döring, T.D., Knapp, S., Kovacs, G., Murphy, K., Wolfe, M.S., 2011. Evolutionary plant breeding in cereals—Into a new era. *Sustainability* 3, 1944–1971.

Döring, T.F., Annicchiarico, P., Clarke, S., Haigh, Z., Jones, H.E., Pearce, H., Snape, J., Zhang, J., Wolfe, M.S., 2015. Comparative analysis of performance and stability among composite cross populations, variety mixtures and pure lines of winter wheat in organic and conventional cropping systems. *Field Crops Res.* 183, 235–245.

Ekbohm, G., 1981. A test for the equality of variances in the paired case with incomplete data. *Biom. J.* 23 (3), 261–265.

Exposito-Alonso, M., 500 Genomes Field Experiment Team, Burbano, H.A., Bosse, O., Nielsen, R., Weigel, D., 2019. Natural selection on the Arabidopsis thaliana genome in present and future climates. *Nature* 573 (7772), 126–129.

FAO, 2019. The state of the world's biodiversity for food and agriculture, J. Bélanger and. In: Pilling, D. (Ed.), *FAO Commission on Genetic Resources for Food and Agriculture Assessments*, p. 572. Rome.

Finckh, M., Gacek, E., Goyeau, H., Lannou, C., Merz, U., Mundt, C.C., Munk, L., Nadziak, J., Newton, A., de Vallavieille-Pope, C., Wolfe, M.S., 2000. Cereal variety and species mixtures in practice, with emphasis on disease resistance. *Agronomie* 20 (7), 813–837.

Frankel, O.H., 1950. The development and maintenance of superior genetic stocks. *Heredity* 4, 89–102.

Frutos, E., Galindo, M.P., Leiva, V., 2014. An interactive biplot implementation in R for modeling genotype-by-environment interaction. *Stoch. Environ. Res. Risk Assess.* 28, 1629–1641.

Goldringer, I., Prouin, C., Rousset, M., Galic, N., Bonnin, I., 2006. Rapid differentiation of experimental populations of wheat for heading time in response to local climatic conditions. *Ann. Bot.* 98, 805–817.

Harlan, H.V., Martini, M.L., 1929. A composite hybrid mixture. *J. Amer. Soc. Agron.* 21, 487–490.

Heeb, L., Jenner, E., Cock, M.J.W., 2019. Climate-smart pest management: building resilience of farms and landscapes to changing pest threats. *J. Pest Sci.* 92, 951–969.

Hufford, M.B., Berny Mier y Teran, J.C., Gepts, P., 2019. Crop biodiversity: an unfinished magnum opus of nature. *Annu. Rev. Plant Biol.* 70, 727–751.

Jalletta, T., 2004. Participatory evaluation of the performance of some improved bread wheat (*Triticum aestivum*) varieties in the Jijiga Plains of eastern Ethiopia. *Exp. Agric.* 40, 89–97.

Jaramillo, F., Destouni, G., 2015. Comment on “planetary boundaries: guiding human development on a changing planet”. *Science* 348, 1217–1218.

Kahiluoto, H., Kaseva, J., Balek, J., Olesen, J.E., Ruiz-Ramos, M., Gobin, A., Kersebaum, K.C., Takáč, J., Ruget, F., Ferrise, R., Bezak, P., Capellades, G., et al., 2019. Decline in climate resilience of European wheat. *Proc. Natl. Acad. Sci. U. S. A.* 116 (1), 123–128.

Kataoka, S., 1963. A stochastic programming model. *Econometrica* 31 (1/2), 181–196.

Kingsolver, J.G., Diamond, S.E., 2011. Phenotypic selection in natural populations: what limits directional selection? *Am. Nat.* 177, 346–357.

Kingsolver, J.G., Hoekstra, H.E., Hoekstra, J.M., Berrigan, D., Vignieri, S.N., Hill, C.E., Hoang, A., Gibert, P., Beerli, P., 2001. The strength of phenotypic selection in natural populations. *Am. Nat.* 157, 245–261.

Knapp, S., Döring, T.F., Jones, H.E., Snape, J., Wingen, L.U., Wolfe, M.S., Leverington-Waite, M., Griffiths, S., 2020. Natural selection towards wild-type in composite cross populations of winter wheat. *Front. Plant Sci.* 10 (1757).

Lin, C.S., Binns, M.R., 1988. A superiority measure of cultivar performance for cultivar x location data. *Can. J. Plant Sci.* 68, 193–198.

Matzrafi, M., Seiwert, B., Reemtsma, T., Rubin, B., Peleg, Z., 2016. Climate change increases the risk of herbicide-resistant weeds due to enhanced detoxification. *Planta* 244, 1217–1227.

McMaster, G.S., Wilhelm, W., 1997. Growing degree-days: one equation, two interpretations. *Agric. For. Meteorol.* 87 (4), 291–300.

- Mundt, C.C., 2002. Use of multiline cultivars and cultivar mixtures for disease management. *Annu. Rev. Phytopathol.* 40, 381–410.
- Nassar, R., Hühn, M., 1987. Studies on estimation of phenotypic stability: tests of significance for nonparametric measures of phenotypic stability. *Biometrics* 43 (1), 45–53.
- Newton, A.C., Johnson, S.N., Gregory, P.J., 2011. Implications of climate change for diseases, crop yields and food security. *Euphytica* 179, 3–18.
- Patel, J.D., Reinbergs, E., Mather, D.E., Choo, T.M., Sterling, J.D., 1987. Natural selection in a double-haploid mixture and a composite cross of barley. *Crop Sci.* 27, 474–479.
- Pautasso, M., Döring, T.F., Garbelotto, M., Pellis, L., Jeger, M.J., 2012. Impacts of climate change on plant diseases—Opinions and trends. *Eur. J. Plant Pathol.* 133 (1), 295–313.
- Piepho, H.-P., 2019. Recent claim of declining climate resilience in European wheat is not supported by the statistics used. *Proc. Natl. Acad. Sci.* 116 (22), 10625–10626.
- R Core Team (2019). **R: A language and environment for statistical computing.** R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>. R Development Core Team. 2015. **R Language Definition Version 3.2.3 DRAFT.**
- Raggi, L., Ciancaleoni, S., Torricelli, R., Terzi, V., Ceccarelli, S., Negri, V., 2017. Evolutionary breeding for sustainable agriculture: selection and multi-environment evaluation of barley populations and lines. *Field Crops Res.* 204, 76–88.
- Rasmusson, D.C., Beard, B.H., Johnson, F.K., 1967. Effect of natural selection on performance of a barley population. *Crop Sci.* 7, 543–543.
- Ray, D.K., Gerber, J.S., MacDonald, G.K., West, P.C., 2015. Climate variation explains a third of global crop yield variability. *Nat. Commun.* 6, 5989.
- Reiss, E.R., Drinkwater, L.E., 2018. Cultivar mixtures: a meta-analysis of the effect of intraspecific diversity on crop yield. *Ecol. Appl.* 28, 62–77.
- Reynolds, M.P., Pask, A.J.D., Joshi, A.K., 2017. Strategic crossing of biomass and harvest index—Source and sink—Achieves genetic gains in wheat. *Euphytica* 213 (257).
- Rockström, J., Steffen, W., Noone, K., Persson, Å., Chapin III, F.S., Lambin, E.F., et al., 2009. A safe operating space for humanity. *Nature* 461, 472–475.
- Rollins, J.A., Drosse, B., Mulki, M.A., Grando, S., Baum, M., Singh, M., Ceccarelli, S., von Korff, M., 2013. Variation at the vernalisation genes *Vrn-H1* and *Vrn-H2* determines growth and yield stability in barley (*Hordeum vulgare*) grown under dryland conditions in Syria. *Theor. Appl. Genet.* 126, 2803–2824.
- Rosenzweig, C., Iglesias, A., Yang, A.B., Epstein, P.R., Chivian, E., 2001. Climate change and extreme weather events – implications for food production, plant diseases, and pests. *Glob. Change Hum. Health* 2, 90–104.
- Sall, S., Norman, D., Featherstone, A.M., 2000. Quantitative assessment of improved rice variety adoption: the farmer's perspective. *Agric. Syst.* 66, 129–144.
- Siepielski, A.M., Di Battista, J.D., Carlson, S.M., 2009. It's about time: the temporal dynamics of phenotypic selection in the wild. *Ecol. Lett.* 12 (11), 1261–1276.
- Singh, M., Malhotra, R.S., Ceccarelli, S., Sarker, A., Grando, S., Erskine, W., 2003. Spatial variability models to improve dryland field trials. *Exp. Agric.* 39, 1–10.
- Smithson, J.B., Lenné, J.M., 1996. Varietal mixtures: a viable strategy for sustainable productivity in subsistence agriculture. *Ann. Appl. Biol.* 128 (1), 127–158.
- Soliman, K.M., Allard, R.W., 1991. Grain yield of composite cross populations of barley: effects of natural selection. *Crop Sci.* 31, 705–708.
- Springmann, M., Clark, M., Mason-D'Croz, D., Wiebe, K., Bodirsky, B.L., Lassaletta, L., et al., 2018. Options for keeping the food system within environmental limits. *Nature* 562, 519–525.
- Steffen, W., Richardson, K., Rockstrom, J., Cornell, S.E., Fetzer, I., Bennett, E.M., et al., 2015. Planetary boundaries: guiding human development on a changing planet. *Science* 347, 1259855.
- Suneson, C.A., 1956. An evolutionary plant breeding method. *Agron. J.* 48, 188–191.
- Thornton, P.K., Ericksen, P.J., Herrero, M., Challinor, A.J., 2014. Climate variability and vulnerability to climate change: a review. *Glob. Change Biol.* 20, 3313–3328.
- Trnka, M., Rötter, R., Ruiz-Ramos, M., Ruiz-Ramos, M., Kersebaum, K.C., Olesen, J.E., Žalud, Z., Semenov, M.A., 2014. Adverse weather conditions for European wheat production will become more frequent with climate change. *Nat. Clim. Change* 4, 637–643.
- State of the Knowledge for Gender in Breeding: Case Studies for Practitioners. Lima (Peru). In: Tufan, H.A., Grando, S., Meola, C. (Eds.), 2018. CGIAR Gender and Breeding Initiative. Working Paper. No. 3.
- van der Wouw, M., Kik, C., van Hintum, T., van Treuren, R., Visser, B., 2010. Genetic erosion in crops: concept, research results and challenges. *Plant Genet. Resour. Charact. Util.* 8 (1), 1–15.
- Van Frank, G., Rivière, P., Pin, S., Baltassat, R., Berthelot, J.-F., Caizergues, F., Dalmasso, et al., 2020. Genetic diversity and stability of performance of wheat population varieties developed by participatory breeding. *Sustainability* 12 (384).
- Vermeulen, S.J., Campbell, B.M., Ingram, J.S.I., 2012. Climate change and food systems. *Annu. Rev. Environ. Resour.* 37, 195–222.
- VSN International, 2019. A Guide to Genstat® Release 20th. Hemel Hempstead, UK.
- Wade, M.J., Kalisz, S., 1991. The causes of natural selection. *Evolution* 44, 1947–1955.
- Wilbur, A., 2014. Back-to-the-house? Gender, domesticity and (dis)empowerment among back-to-the-land migrants in Northern Italy. *J. Rural Stud.* 35, 1–10.
- Wolfe, M.S., Ceccarelli, S., 2020. The need to use more diversity in cereal cropping requires more descriptive precision. *J. Sci. Food Agric.* 100, 4119–4123.
- Wolfe, M.S., Brändle, U., Koller, B., Limpert, E., McDermott, J.M., Müller, K., Schaffner, D., 1992. Barley mildew in Europe: population biology and host resistance. *Euphytica* 63 (1–2), 125–139.
- Yan, W., Hunt, L.A., 2020. Biplot analysis of multi-environment trial data. In: Kang, M.S. (Ed.), *Quantitative Genetics, Genomics and Plant Breeding*, 2nd edition. CABI, Wallingford, UK, pp. 162–177.
- Yan, W., Rajcan, I., 2002. Biplot analysis of test sites and trait relations of soybean in Ontario. *Crop Sci.* 42, 11–20.
- Yan, W., Hunt, L.A., Sheng, Q., Szlavnic, Z., 2000. Cultivar evaluation and mega environment investigation based on the GGE biplot. *Crop Sci.* 40, 597–605.
- Zavala, J.A., Casteel, C.L., DeLucia, E.H., Berenbaum, M.R., 2008. Anthropogenic increase in carbon dioxide compromises plant defense against invasive insects. *Proc. Natl. Acad. Sci.* 105 (13), 5129–5133.
- Zhu, Y., Chen, H., Fan, J., Wang, Y., Li, Y., Chen, J., Fan, J.-X., Yang, S., Hu, L., Leung, H., Mew, T.W., Teng, T.W., et al., 2000. Genetic diversity and disease control in rice. *Nature* 406, 718–722.
- Ziska, L.H., Dukes, J.S., 2010. *Weed Biology and Climate Change*. Blackwell Publishing Ltd, Ames, Iowa 50014-58300, USA.