Marche Polytechnic University
Department of Agricultural, Food and Environmental Sciences
Scientific field: AGR/07 - Plant genetics

PhD School of Agricultural, Food and Environmental Sciences
XXXI cycle (2015-2018)

Experimental evolution approach to study crop adaptation

PhD supervisor:
Prof.ssa **Laura Nanni**

PhD School director:
Prof. **Bruno Mezzetti**

PhD co-supervisor:
Dott.ssa **Oriana Porfiri**

PhD candidate:
Antonia Mores

Table of contents

Dissertation Abstract

(English Version)

Landraces are defined as dynamic populations of a cultivated plant with an historical origin, distinct identity, often genetically diverse and locally adapted. The present industrialized agricultural system represents the most important threat to these genetic resources with the serious consequence of loss of biodiversity. For these reasons the study of the most important agricultural species and the evolution of populations over time is at the basis of modern conservation techniques. Moreover the use of these agricultural genetic resources is useful for the purpose of medium- and long-term genetic improvement. Given the potential of experimental evolution in the context of dynamic conservation ex-situ, in the present study we investigated the evolutionary forces involved in the adaptation process in Marche region of an important Sardinian barley landrace population. In collaboration with a local farmers', it was possible to study the effects of ex-situ dynamic conservation (conservation and selection) on diversity and adaptation. The genotyping of 252 individuals, spanning 15 years of cultivation, was used for the analysis of genetic diversity and population structure. We found that the action of the natural selection seems more likely acting. Results from a study over 15 years of experimental dynamic conservation are presented for the first time, showing that this dynamic conservation is a key element for the long-term conservation and use of agricultural biodiversity.

Dissertation Abstract

(Versione Italiana)

Le varietà locali sono definite come popolazioni dinamiche di una pianta coltivata con un'origine storica, un'identità distinta, spesso geneticamente diversificata e adattata localmente. L'attuale sistema agricolo industrializzato rappresenta la minaccia più importante per queste risorse genetiche con la grave conseguenza della perdita di biodiversità. Per questi motivi lo studio delle più importanti specie agricole e dell'evoluzione delle popolazioni nel tempo è alla base delle moderne tecniche di conservazione. Inoltre, l'uso di queste risorse genetiche agricole è utile ai fini del miglioramento genetico a medio e lungo termine. Dato il potenziale dell'evoluzione sperimentale nel contesto della conservazione dinamica ex-situ, nel presente studio abbiamo studiato le forze evolutive coinvolte nel processo di adattamento nella regione Marche di un'importante popolazione locale di orzo sardo. In collaborazione con un agricoltore locale, è stato possibile studiare gli effetti della conservazione dinamica ex-situ (conservazione e selezione) sulla diversità e l'adattamento. La genotipizzazione di 252 individui, che copre 15 anni di coltivazione, è stata utilizzata per l'analisi della diversità genetica e della struttura della popolazione. Suggeriamo che l'azione della selezione naturale abbia agito di più rispetto alle altre forze evolutive. I risultati di uno studio di oltre 15 anni di conservazione dinamica sperimentale sono presentati per la prima volta, dimostrando che questo tipo di conservazione dinamica è un elemento chiave per la conservazione e l'uso a lungo termine della biodiversità agricola.

Acknowledgments

First, I would like to thank all the members of the plant genetics group of the Polytechnic University of Marche, including Prof. Roberto Papa, Prof. Laura Nanni, Dr. Elisa Bellucci and Dr. Elena Bitocchi. Also my colleagues that I was fortunate to meet and share experiences with during my studies: Ana, Valerio and Eda and newcomers Giulia, Lorenzo and Gaia. I am grateful for the research doctorate program I have been a part of. Especially I thank Roberto and Laura for their supervision and guidance.

Also, I want to thank Dr. Oriana Porfiri, that co-financied my scholarship.

In addition, I thank Dr. Monica Rodriguez of the University of Sassari, for having her help in the data analysis part.

I would like to thank Prof. Andreas Graner for giving me the opportunity to spend six months at his laboratory at the IPK Institute, Gatersleben (Germany). Also Kerstin Neumann for her supervision during my activity at the IPK.

I take this opportunity to thank my family for their love, understanding and support in every single step and decision that I have needed to make. And Antonio for being close to me all the way.

I also would like to thank all the members of the E.B.I. group, which were fantastic adventure companions during my time in Germany. In particular Isa and Paride for giving me their wisdom and genuineness.

Thank you all so much!

CHAPTER 1: Review on In-situ conservation: exploiting genetic diversity and population structure of crop landraces

In-situ conservation: exploiting genetic diversity and population structure of crop landraces

Abstact

Landraces are dynamic populations that are cultivated by farmers as populations and subjected to the constant pressure of evolutionary forces. This means that they are a valuable source of genetic variability that can be used in plant genetic improvements. To be able to study the wide genetic diversity, through the use of molecular markers, phenotyping and/or QTL/GWA studies, it is necessary to understand the evolutionary forces that have acted and continue to act during the on-farm conservation process of landraces. This review explores all those works concerning the study of genetic diversity and population structure of local varieties collected directly by the farmers. In addition, new strategies for studying adaptation during the conservation process are proposed that will be useful in future genetic improvement programs.

Introduction

To cope with climate change and to meet the needs of new varieties for marginal areas, researchers and breeders are constantly looking for sources of genetic variability. A key component of agro-biodiversity is represented by landraces. The term landrace was introduced first by von Rünker (1908). There have been many attempts to define landraces. Harlan (1975) defined them as populations that had evolved in subsistence agricultural societies as result of 'millennia long', 'artificial' human selection pressures, mediated through human migration, seed exchange as well as natural selection. He believed that landraces have three basic characteristics: variability of genotypes, integrity and local adaptation; and this conception of a landrace is the one still most widely applied. Hawkes (1983) extended the term by adding the association with marginal environments and a lack of direct competition with highly bred cultivars. However, Zeven (1998) in a review of landrace definitions concluded that 'as a landrace has a complex and indefinable nature an all-embracing definition cannot be given'. Anyway, the definition of landrace is reflected with certain characteristics associated to them. Landraces represent an intermediate stage of domestication between the wild ancestor and modern varieties. Landraces reflect the cultural identity of peoples, for their adaptation to the environment in the area of cultivation, linked

to the uses, knowledge, habits, and traditions of the people who have developed and cultivated them in recent years (Polegri and Negri, 2010). As demonstrated by Negri (2005), landraces are adapted to specific agro-climatic conditions while maintaining considerable diversity between and within the population. Therefore, they constitute a reservoir of genetic diversity that is interesting for future breeding work, as well as for the development of new agricultural systems and new products.

With the occurrence of the green revolution (’40 -’60 of the XX century), it has been developed to develop elite cultivars adapted to areas where high-input agriculture is practiced. The advent of the modern varieties promised to satisfy a higher demand of food for a growing population, but they are influenced by environmental changes that reduce their fitness.

The cultivation of genetically similar individuals (pure lines, hybrids or clones) leads to the spread of disease outbreaks, causing large crop losses both in productive and economic terms (e.g. Southern Downy in the United States in the 70s, Ullstrup 1972). It also leads to dramatic consequences for crop improvement, as there is a loss of potentially useful characteristics, thus reducing genetic variability that can be subsequently used by plant breeders for crop improvement. Furthermore, agriculture facing challenges such as climate change drought, cold, heat and salinity, together with low soil fertility due to intensive agronomic practices, are the main cause of low productivity both in terms of yield and quality. Meanwhile they represent a threat to the survival of landraces that had been preserved for generations by local farmers and now in world genetic banks.

The investigation of the genetic diversity structure within landraces is well studied for most important crop varieties (barley, Hadado et al., 2009; beans, Sicard et al., 2005 ; rice, Pusadee et al., 2009; maize, Qi-lun et al., 2008; tomato, Mazzucato et al., 2008). Studies have also been conducted to understand how natural processes (selection, genetic drift and migration) and agricultural practices interact to shape and maintain diversity within and between landrace population (Tiranti and Negri, 2007; Pusadee et al., 2009; Bellucci et al., 2013).

Given their high genetic diversity, landraces are used in modern breeding programs to broaden the genetic diversity of genetic pools and to improve their adaptation, such as for abiotic or biotic stresses, performance characteristics, quality and resistance to pests (Fischbeck, 2003; Bradshaw and Ramsay, 2005; Reynolds et al., 2006; Sharma-Poudyal et al., 2013).

Evolutionary forces affecting genetic composition of landraces cultivated on-farm

The understanding of the nature, cadence and causes of genetic change during on-farm conservation is crucial. The forces responsible for changing the genetic structure of genetic resources are selection, genetic drift or population fragmentation (which is maintained through selection, isolation) and lack of migration and restrictions on outcrossing and genetic recombination. Another important aspect to understand is variation due to adaptation. This aspect includes indicators of the degree to which populations are adapted to their environment. Both biotic and abiotic aspects of the environment are involved.

Selection

Selection is the main evolutionary force that determines the levels and types of diversity between individuals within the population and between populations within a crop species. The first example of selection is the domestication process. Domestication is a genetic selection process made, consciously or unconsciously, by humans to adapt wild plants and animals to cultivation and herding, respectively. During domestication, plants and animals are subjected to natural as well as human selection. Darwin (1868) observed that "No doubt man selects varying individuals, sows their seeds, and again selects their varying offspring. Man therefore may be said to have been trying an experiment on a gigantic scale; and it is an experiment which nature during the long lapse of time has incessantly tried". During the selection process of domestication, crop plants acquired a set of new specific traits adaptive to new environments and increased their usefulness to farmers and consumers.

Selection usually modifies allele frequencies, favouring some alleles at the expense of others and gives rise to changes in the frequencies of genotypes in subsequent generations. Allard et al. (1972) have shown that there is extensive allelic variation within many of the accessions of a US barley collection, as well as pronounced differences in allelic frequencies among barleys from different ecogeographical regions of the world. In order to determine which evolutionary forces are responsible of these differences, they investigated a single-locus population dynamics of four esterase (co-dominant markers, which are useful for analysing the genetic structure of plant populations) loci in Composite Cross V (CCV, associates by intercrossing 30 barley varieties representing all of the major barley growing regions of the world). The results show that balancing selection with direction of selection different for different alleles is responsible for the changes in allelic frequencies which

occurred over generations in this population, and for the maintenance of variability within the population.

Moreover, in populations harbouring genetic variations, selection for adaptation will derive from the effect of environmental factors, as demonstrated by Pusadee et al. (2009). In this study was demonstrated a significant correlation between the genetic and geographical distances of the local populations of Thai rice. This correlation reflects an isolation by distance (IBD) population structure (Wright, 1946) where populations at greater distances are more genetically dissimilar than those populations that are geographically close. Genetic isolation by distance is a dynamic process in space and time that produces changing genetic composition. This result suggest that Thai rice local varieties represent a dynamic genetic system, that responds to natural and artificial selection, which promotes local adaptation of landraces populations. This dynamic nature of landraces makes them particularly important sources of germplasm for breeding programs.

Mutation

Mutation is an inheritable change in the nucleotide sequence and therefore a source of new genetic variation. Natural mutation rates are low and most mutations are neutral in relation to the fitness of the organism, or have only minor, generally deleterious, effects. However, domestication of crops has been the result of human selection for specific mutations conferring adaptation for cultivated environments and use. As demonstrated in a recent study on Jordanian barley landraces, it was found a relatively high rate of mutation. This rate is resulted in 107 unique alleles found in barley landraces two- and six-rowed, but not in the modern cultivars (Hasan et al., 2018). A similar result (5 private alleles) was found in a Tunisia barley landrace populations (Romdhane et al., 2017). The adaptation of the barley germplasm to stressed environments (mountains, frost) throughout thousands of years could explain the high number of private alleles in this region. Thus, the occurrence of so many unique alleles could be used as a potential pool of new alleles needed for plant breeding.

Recombination

As mutation, recombination can lead to new plant types that can be selected by farmers. This process generates new combinations of alleles at linked genes. The effects of recombination are strongly influenced by the breeding system. An outcrossing species as maize (*Zea mays*

L. ssp. *Mays*) exhibits high levels of recombination. In Bitocchi et al. (2009) was investigated the level of introgression from maize hybrids into the maize landraces. They found a low level of introgression, not excluding the occurrence of selection at specific loci, but just that the selection did not influence a large portion of the regions surrounding the loci under selection, probably due to the high effective recombination rate characteristic of maize, where LD (linkage disequilibrium, the non-random association of alleles at two or more loci in a population) between loci decreases rapidly. High recombination rate will restrict these areas to very small portions of the genome.

Whereas, in autogamous species the size of genome regions under recombination may be much larger. For example, barley landraces are the result of evolutionary processes that lead to heterogeneous populations rather than to few superior genotypes. In this populations the diversity is held together as blocks of genes in chromosomal regions with low frequency of recombination, may confer a specific adaptation to stress environments (Ceccarelli and Grando, 2000).

Migration

In population genetics, migration applies to the process of movement of individual plants, vegetative propagules, seeds or pollen among populations (Slatkin, 1987). Migration, also called gene flow, occurs both with the advancing front of a population when it is colonizing new areas, and when genes of two or more populations mix through pollen and seed dispersal. When the movement of this kind is accompanied by crossings between migrants and receiving population, it translates into genetic integration into a new introgressive population. When the sample of incoming gametes or seeds differs genetically from the receiving population, it is said that the gene flow will occur. An important study on the introgression from modern hybrid varieties in populations of landraces varieties of maize (an allogamous crop) was conducted by Bitocchi et al. (2009). They compared an ‘old’ collection, obtained during the 1950s, before the introduction of hybrids, and a recent collection of maize landraces. A total of 296 genotypes were analysed using 21 microsatellites (SSRs). The population structure, diversity and linkage disequilibrium analyses indicate a significant amount of introgression from hybrid varieties into the recent landrace populations. The level of introgression detected was very variable among recent landraces, with most of them showing a low level of introgression. As well as the gene flow between modern varieties and local varieties of barley (an important autogamous crop) was

evaluated (Bellucci et al., 2013). Using 12 SSR markers, they investigated the level of introgression occurred within two collection collected at a distance of 10 years. Both studies reported similar results, i.e. a low level of introgression probably due to the agronomic practices of individual farmers. This confirms that in situ conservation strategies based on the use of farmers can preserve the genetic identity of local varieties (Romdhane et al., 2017). The gene flow from wild relatives to cultivated species has been important in their evolution, gradually expanding their genetic base and increasing the amount of diversity available for selection (Jarvis and Hodgkin 1999).

Genetic drift

One of the major problems in the conservation of populations is the occurrence of genetic drift, that is, the random fluctuations of gene frequencies in a population from one generation to another. The extension of fluctuations of gene frequencies depends on the size of the population: as the population decreases, the effect is greater. Ultimately, genetic drift will lead to fixation of alleles in any locus, a process that can occur randomly, and in any populations or subpopulations different alleles can be fixed at a particular locus. The genetic drift in divided subpopulations can give rise to small changes among the populations; these small variations accumulate over generations and give rise to genetic differentiation. The two main general effects are the loss of allelic variants, particularly rare, and an increase in the level of consanguinity (or reduction of heterozygosity). As the population of large size decreases, the loss of allelic wealth or genetic erosion is first evident and the effects of inbreeding predominate in small populations. This theory is the basis of the so-called 50-500 rule proposed by Franklin (1980), that became a popular guiding principle in conservation genetics for assessing Minimum viable population (MVP, the minimum population size at which a population is likely to persist over some defined period of time with a given probability of extinction). Based on the estimated mutation rate of abdominal bristle number in *Drosophila*, he suggested that the minimum size of the vital population to avoid consanguinity is 50, and that to limit the genetic erosion is 500. Prevention of inbreeding depression it is a matter of immediate adaptation and suitability for current survival and allelic wealth is necessary for adaptability or survival to future environmental changes.

Studies of population genetic structure of landraces conserved in situ

An important aspect of genetic conservation point of view of landraces is the understanding of the amount and distribution of genetic diversity present in crops in farmers' fields. To this end, it is necessary to answer some questions: understand how genetically variable are the local varieties of a crop or the different populations of a local variety currently grown in a village or farm; how the varieties differ from each other in the type and quantity of variation they harbor; and if some varieties contain unique characteristics (singly or in combination) that are absent from other varieties.

Monitoring genetic diversity with molecular markers

Analysis of population genetic structure focuses on the genetic characteristics of population, i.e. the genes and allelic variants and their frequencies, and how these vary within and among populations, in space and time. There are several methods to determine the extent and distribution of diversity within and between population. These include analysis of variation using morphological, quantitative and biochemical traits, and DNA markers (AFLP, RAPD, SSR, SNP). Sequencing techniques have developed rapidly over the last decade and it is now possible generate large amounts of DNA sequences of specific loci or complete sequence data for a large number of plants. Carović-Stanko et al. (2017) conducted a study on Croatian common bean (*Phaseolus vulgaris* L.) landraces accessions in order to assess the genetic diversity and structure of Croatian common bean germplasm. In Croatia the production of common bean is based on landraces which are adapted to the specific growing conditions and agro-environments which display high levels of morphological diversity. Landraces are traditionally grown in low-input production systems. However, in recent years, landraces are in danger of genetic erosion caused by complex socio-economic changes in rural communities (the low profitability of farms, their small size, and the advanced age of farmers, the replacement of traditional landraces with modern bean cultivars and/or other more profitable crops; FAO, 2008). By phaseolin (the major seed storage protein of common bean; Lioi et al., 1989; Kami et al., 1995) genotyping and microsatellite marker analysis of 183 genotypes, they found that 27.32% were of Mesoamerican origin, 68.31% of Andean, while 4.37% of accessions represented putative hybrids between gene pools.

Monitoring phenotypic traits

Despite the extraordinary development of DNA analysis capacities, it is worth remembering that much useful information can still be obtained from quantitative and qualitative data. Lazaridi et al. (2017) conducted a study to characterize, evaluate diversity and classify 23 local cowpea populations from Greece based on agro-morphological characters (seed yield, mineral and seed crude protein content as well as correlations between them). A relatively high phenotypic diversity was observed. In particular, a high level of diversity was found within the population (phenotypic diversity intra-population, Hs = 0.34) higher than that among populations (phenotypic diversity among populations, G_{ST} = 0.27). Principal component analysis classified the majority of the local populations in two groups (mainly according to populations' seed coat color and eye color), further divided into six subgroups independently of the geographical origin of the populations, presumably due to seed exchanges by farmers. Significant differences were also observed between populations studied for potassium and calcium seed content, as well as for their content in crude seed proteins (22.14-28.37%). The results showed appreciable levels of intra- and inter-phenotypic diversity in on- farm conserved cowpea populations, suggesting the existence of a valuable gene pool for future exploitation in breeding programs for cowpea.

Genome-wide association studies

Even the most comprehensive set of DNA data ultimately needs to be connected to the morphological, agronomic, and other useful traits in which farmers and producers are interested. Genetic mapping is highly valuable to localize economically important traits and to develop closely linked markers to these traits useful for marker-assisted selection. Allel et al. (2017) evaluated the genetic diversity among barley populations collected from North Africa by association of molecular marker data (RAPD) with morphogenetic characters. Molecular variance has shown that genetic variation was partitioned in prevalence within (60.31%) rather than between (39.69%) the populations. The data showed a considerable gene flow (Nm = 1.43, migrants exchanged between populations, on average, each generation) and a moderate genetic differentiation (inter-population genetic variation, G_{ST} = 0.26) with the highest diversity observed in Tunisian and Egyptian populations. No correlation was found between morphological and molecular similarity matrices suggesting that both approaches should be investigated as complementary tools for characterization of

such barley germplasm. They conclude that investigating the genetic structure of barley germplasm of some North African countries based on phenotypic and molecular variability remains of significant importance for genetic improvement and elite gene exploitation.

New strategies to retain diversity from landraces

Jensen et al. (2012), assert that the isolation of crop germplasm (ex-situ conservation) from the original conditions and their reintroduction could significantly lead to a fitness reduction. Parzies et al. (2000) compared the level of diversity of barley landraces accessions in Syria that were been conserved for different lengths of time in ex situ gene banks. They considered three storage time: 10, 40 and 72 years, in order to determine whether rejuvenation practice by multiplication in small parcels of soil of seed samples stored in cold stores has the potential to create bottlenecks in the population, leading to loss of genetic diversity and changes in gene frequencies at each cycle of rejuvenation. The average genetic diversity (H), the alleles per locus (A) and the percentage of polymorphic loci showed very significant decreases (criterion of significance, $P<0.01$) with the conservation time and the genetic differentiation (F_{ST}; Nei, 1977) between accessions increased over time. The results suggested that the differences observed in genetic diversity are caused by genetic drift. Efforts must be dedicated to conservation and use of landraces, especially in their place of origin (Jarvis et al., 2011).

Thus, in addition to static ex situ conservation, the building of appropriate dynamic in situ conservation is a very important strategy to develop and maintain a crop germplasm and to preserve its potential for adaptation. In fact, Bellucci et al. (2013) propose an in situ conservation method based on farmer use can preserve the genetic material of landraces and, at the same time, allow its adaptation to local environments.

Hence the importance of finding strategies of conservation of these pools of genetic diversity represented by landraces.

Experimental evolution approach to study crop adaptation

Experimental evolution concerns the study of populations in several generations under defined and reproducible conditions (Bennett, 2003; Garland, 2003; Swallow and Garland, 2005; Chippindale, 2006; Garland e Kelly, 2006). Experimental evolution studies have long been valued as a means of gaining insight into the rate and degree of response to selection

(Burke and Rose 2009). For many generations, a population is exposed to novel environments, while a parallel population is maintained within the ancestral environment, as experimental control. The environmental novelty may involve alteration of any aspects of the abiotic, biotic, or demographic condition of the ancestral population. Usually, only a single environmental variable is altered to keep the experiment as simple as possible. The novel experimental environment provides new selective conditions and hence promotes evolution. New genetic variants may be produced (through recombination, mutation, or other evolutionary processes). In altered conditions these new variants may be advantaged or disadvantaged, leading to an increase in the frequency of favourable genotypes (or allelic combinations) (Futuyma and Bennett, 2009).

After a sufficient number of generations, the novel experimental populations may be compared directly with the controls (or in some cases directly to their own ancestors), and a priori hypotheses concerning evolution may thereby be tested. The longer the individuals of the ancestral population have been maintained in the ancestral (control) condition, the more likely any differences observed within the novel experimental populations will be specific to the novel environmental alterations (Travisano, 2009).

Experimental evolution has been successfully used in microorganisms (Elena and Lenski, 2003), such as *Escherichia coli* exposed to high-temperature regimes (Tenaillon et al., 2012) and yeast (Kao and Sherlock, 2008) and also in multicellular organisms, including *Drosophila* population selected for longevity (Burke and Rose, 2009), and in crop species, such as barley (Allard et al., 1972; Weir et al., 1972; Clegg et al., 1972).

Although plant breeding populations are themselves highly successful long-term experimental evolution studies, there is a long-standing tradition of developing experimental populations to supplement traditional breeding approaches (Harlan and Martini, 1929; Suneson, 1956; Dudley and Lambert, 2004). These experimental populations explore both the potential for local adaptation (Allard et al., 1972) and the genetic basis of response to selection (Clegg et al., 1972; Weir et al., 1972).

According to this experimental evolutionary approach, a dynamic conservation study based on a barley landrace population has been realized in order to study the adaptation process.

CHAPTER 2: Study of adaptation of a Sardinia barley landrace population in an experimental dynamic conservation

Introduction

Barley: taxonomy and geographical distribution

Barley (*Hordeum vulgare* L., 2n = 2x = 14) is one of the oldest crops cultivated in the world. The crop was domesticated from its wild relative (*Hordeum spontaneum* C. Koch) (Hopf and Zohary, 1988; Harlan, 1995). Barley belongs to the grass family *Poaceae*, the tribe *Triticeae* and the genus *Hordeum* which is comprised of about 33 species and 45 taxa separated into four sections (von Bothmer, 1992). The division of the genus into four sections puts plants into groups that have similar morphological characteristics, life forms, similarities in ecology, and geographical area of origin (von Bothmer et al., 2003).

Barley as a whole is well adapted to marginal and stress-prone environments such as high soil salinity, and a more reliable crop than wheat or rice in regions which are colder or higher in altitude. Due to its wide adaptation to marginal and extreme conditions, barley is found in a wide range of geographical distributions throughout the world. Today, barley is grown in fertile as well as in marginal areas under extreme conditions, including altitudes of up to 5,500 m a.s.l. in the Himalayas, in seasonal flooded areas in south east Asia, and in arid regions of the Mediterranean (Angessa and Li, 2015; von Bothmer et al., 2003). There are different morphological forms of barley including two-rowed, six-rowed, hull/hull-less, and hooded barley. These different forms can be differentiated into spring or winter type based on the requirement of vernalization. The spring type requires no vernalization treatment as opposed to the winter one. This difference in both morphological and physiological characteristics is a reflection of the underlying genetic diversity which eases the adaptation of barley to different environments. Wild barley is predominantly a winter-type, the domesticated form is either spring or winter type.

The importance of barley

Barley together with other cereal grains such as emmer, einkorn and later modern wheat and rice were the staple food and probably the most important products of the world in earlier times. Both wheat and barley were grown in Turkestan in the third millennium B.C. and were also a basic food source for the Sumerian diet. Despite the primary use of barley as a staple food, probably as porridge or bread, it was used in making beer or alcoholic beverages called “barley wine” (Bishop, 1936; Harlan, 1978). In the earlier times, recipes from barley puls, an oily seasoned paste mixture was a popular food in Greece (Tannahill, 1988). Barley

was a common constituent of unleavened bread and porridge eaten by ancient Greeks and was also used as an energy food by the Roman gladiators who were called Hordearii or "barley men" (Percival, 1921; Ceccarelli et al., 2010).

Although barley was classified as an important food grain in ancient times, its use as a staple food source declined as other food grains such as wheat, rye and oat became abundant. Barley became relegated to the status of "poor man`s bread" (Zohary, 1988). However, consumer's interest in nutrition and the health benefits of barley has helped in the restoration of barley's status in the human diet. Barley currently ranks fourth after maize (*Zea mays* L.), rice (*Oryza sativa* L.) and wheat (*Triticum aestivum* L.) in terms of worldwide acreage cereal production (FAO 2015). European Union is the leading producer of barley (41.9%), followed by the former Soviet Union (21.9%), North America (9.3%), and the Middle East (8.5%). Today, barley is mostly used in the brewing and malting industry and also as an important source of animal feed in the developed world (Friedt and Ordon, 2013). Considered a staple food source in several regions of the world particularly in Asia and North Africa, including Ethiopia where barley is still used for bread porridge (Baik and Ullrich, 2008; von Bothmer et al., 1995), barley is rich in nutrients. Its nutritional components generally may differ greatly due to genotype, cultural practices and growing conditions. Starch, fiber, and proteins make up the largest portion of the kernel with a variation in one of the components directly influencing the amounts of the other two. Barley constitutes a rich source of dietary fiber with β-glucan being the most important in terms of human diet and health benefits (Van Hung, 2016). The high amount of β-glucan helps lowering cholesterol and blood glucose levels.

Barley domestication

Archaeological evidence regarding barley, along with einkorn and emmer wheat, indicates that the movement of mankind from hunter-gathering to cultivation and agriculture occurred in the Fertile Crescent around 12,000 to 9,500 years ago with the domestication process lasting several centuries (Tanno and Willcox, 2006; Weiss et al., 2006; Willcox et al., 2008). As revealed by Badr et al. (2000), with allelic frequencies at 400 AFLP polymorphic loci studied in 317 wild and 57 cultivated barley, wild populations from Israel-Jordan were molecularly more similar than to the cultivated gene pool leading to the early hypothesis of Israel-Jordan to be the main region of domestication. This hypothesis was later supported by the diagnostic allele I of the homeobox gene BKn-3 (a gene involved in awn suppression in

barley), rarely but exclusively found in Israel ssp. spontaneum. However, recent studies were conducted on wild and landrace (primitive domesticates) barley collections (Morrell and Clegg, 2007; Saisho and Purugganan, 2007). With evidence of independent origins of important domestication-related traits such as the brittle rachis (controlled by *btr1* and *btr2*) (Pourkheirandish et al., 2015), supported the hypothesis of at least two independent domestication events followed by some degree of admixture amongst domesticates from distinct portions of wild barley (Fuller et al., 2011; Kilian et al., 2007). Morrell and Clegg (2007) showed that two tough rachis mutants were associated with genetically distinct groups of domesticated barley, one associated with West and the other with the East, suggesting that a core centre of origin does not apply to barley. The two genes, non-brittle rachis 1 (*btr1*) and non-brittle rachis 2 (*btr2*), control spike disarticulation in barley with a mutation in either of the two genes converting the brittle rachis (wild) to a non-brittle (domesticated) type (Pourkheirandish et al., 2015). These authors reported two distinct regions where early farmers must have independently selected for mutations of the barley brittle rachis phenotype. Jones et al. (2013) provided additional evidence, indicating the possibility of at least two domestication events. Such evidence has been gained from analysis of European barley, where at least two different wild barleys were adopted into central and northern Europe. Poets et al. (2015) while examining 6152 single nucleotide polymorphisms (SNPs) have reported further evidence of more than one centric origin of domestication. Here, several regional groups of domesticated barley were discerned relating to Central Europe, Coastal Mediterranean, Asia and East Africa. Data from chloroplast DNA clearly indicate that wild barley as it is found today in the Fertile Crescent might not be the progenitor of barley cultivated in Eritrea/Ethiopia, indicating that an independent domestication might have taken place at the Horn of Africa (Ethiopia) (Orabi et al., 2007). Barley exist as hulled or hull-less, with early reports pointing to a single gene (Nud) controlling the hulled phenotype (Taketa et al. 2008) which also led to earlier suggestion of a single origin of domesticated hull-less barley. A recent study has reported more than one origin of domesticated hull-less barley, with a suggestion of Tibetan hull-less barley having an independent origin of domestication (Yu et al., 2016).

The genepool of barley

Barley is a diploid predominantly self-pollinating grain crop with a basic chromosome number of n=7 (2n=14) and a large genome size of 5.1 Gb with approximately 84% of its

genome mostly comprising of mobile elements or other repetitive structures (Dolezel et al., 1998; Mayer et al., 2012; Wicker et al., 2008). As showed in Figure 1, barley and its related species have been classified into genepools as primary, secondary, and tertiary based on cross-ability, hybrid viability, and meiotic chromosome pairing (Harlan and Wet, 1971).

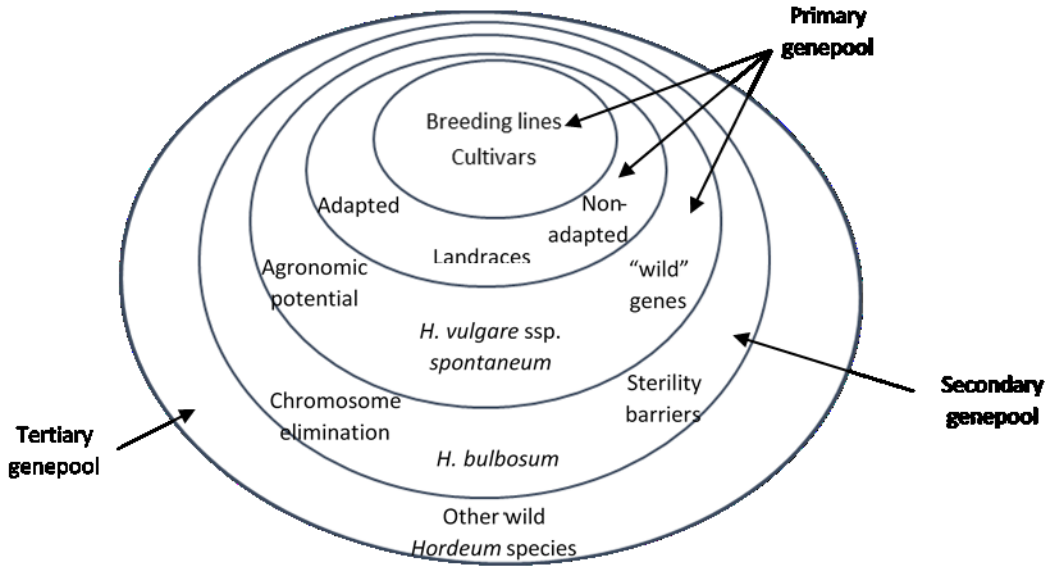

Figure 1. Genepools in cultivated barley by Harlan (*Hordeum vulgare*) (von Bothmer et al. 1995).

The primary genepool of barley includes domesticated barley (*H. vulgare* ssp. *vulgare*) and the wild form (*H. vulgare* ssp. *spontaneum*) which are both annual, diploid and predominantly inbreeding. Both members of the primary genepool are cross-compatible producing vigorous, viable, and fertile hybrids, facilitating the transfer of desired traits from the wild subspecies to cultivated barley. Within the primary genepool, wild and cultivated barley from several geographic areas are genetically highly diverse (Petersen et al., 1994). The secondary genepool includes only a single species *H. bulbosum* L. that shares the H genome with the primary genepool and consists of both diploid and tetraploid forms which are mainly self-incompatible perennials. Crosses between the two genepools are difficult. Diploid and tetraploid *H. bulbosum* from the secondary genepool is widely used to produce haploid barley through chromosome elimination (Kasha and Kao, 1970; Pickering, 1984; Subrahmanyam and von Bothmer, 1987). Several genes from *H. bulbosum* have been transferred to cultivated barley, providing a new source for breeding (Pickering, 2000). The tertiary genepool is the largest and includes all the remaining species of *Hordeum.* Members

belonging to the tertiary genepool are either diploid, tetraploid or hexaploid and consist of both annuals and perennials (von Bothmer et al., 1995; von Bothmer et al., 2003).

Barley landraces

Barley landraces are heterogeneous populations developed by natural or farmer directed selection with a high local adaptation (Poets et al., 2015). Barley landrace populations are comprised of inbred lines and hybrids generated by a low level of outcrossing (Nevo and Shewry, 1992). Landraces harbour a rich source of genetic diversity which has been exploited by The International Centre for Agriculture Research in Dry Areas (ICARDA) to improve yield and yield stability in dry areas. Similar to wild barley, landraces have high within-population diversity. In selected gene bank collections, at least 50-60% of the total genetic variation captured resides within populations, the remainder being accounted for by differences between landraces (Endresen et al., 2011; Poets et al., 2015). Genetic diversity of landrace populations collected from Sardinia revealed that only 11% of the diversity detected by RAPD markers occurred between populations (Papa et al. 1998). Also, high level of genetic diversity was reported in a collection of 1485 barley landraces originating from 41 countries genotyped with a set of 42 SSR markers (Pasam et al., 2014). The utilization of landrace genetic diversity as a source of crop improvement has been successful in many areas. Some examples include: the introgression of dwarfing alleles (*Rht1* and *Rht2*) derived from the Japanese wheat landrace "Shiro Daruma"(Kihara, 1983); powdery mildew resistance allele *mlo11* derived from an Ethiopian barley landrace (Piffanelli et al., 2004); the barley yellow mosaic resistance gene, *rym4* (Graner and Bauer, 1993); the boron-toxicity tolerances in barley obtained from the Algerian landrace "Sahara" (Sutton et al., 2007); and improvement of yield and abiotic stress adaptation (Dwivedi et al., 2016). Exploiting genetic diversity in the wild form and in landraces for crop improvement will help improve barley production in the future.

Barley landraces in Sardinia

In Italy, barley cultivated area amounted to 260,051 hectares, with a production of 1,06 million tonnes. In many countries in the developing world, cultivation of barley is an important food for humans as a source of carbohydrates and proteins. Instead, in most developed countries barley production is mainly destinated in the animal feed livestock (85-

90%) and secondarily in malt industry (10-15%). The Italian region where the barley represented a significant food for animals source is Sardinia. Today barley represents the most cultivated cereal after wheat (13,335 ha in 2018; http://agri.istat.it). Attene et al., (1996) described a widespread cultivation of a polistic barley landrace known as "S'orgiu sardu". This barley is used both for the production of fresh forage through the direct grazing (sheep) and for the production of grain. Local accessions of barley have long been cultivated in different environments at different latitudes, soil types and with different cultural practices and this may have allowed the development of locally adapted populations. In 1990, twenty-seven populations of barley were collected in different growing areas of Sardinia (Attene et al., 1996). Farms that for at least 30 years had used seeds of own production were selected. Of 27 local populations collected in Sardinia, 240 lines of 12 populations were analyzed. The existence of high genetic variability for many characters such as date of ear emergence, plant height, kernel weight, grain yield, biomass total and harvest index were shown (Papa et al., 1994). The collection of populations of Sardinian barley was repeated in 1999, visiting the same areas sampled in 1990 (Rau, 2002). The comparison of the collections of 1990 and 1999 using microsatellite molecular markers (SSR, Simple Sequence Repeats) and based on retrotransposons (SSAP, Simple Sequence Amplified Polymorphism) not showed neither changes in the levels or in the structure of genetic diversity. This indicated that farmers implement a system of conservation "in situ" able of preserving the genetic identity of the populations, even in the absence of specific conservation projects for these plant materials (Rodriguez et al., 2012; Bellucci et al., 2013). The analysis of the spatial structure of the genetic variability showed a level of genetic divergence of populations moderate-low, measured by a Wright's fixation index (F_{ST}; 1931) of 0.18 (Rodriguez et al., 2012). It was observed that 18% of molecular variance (the allele frequencies) was explained by the differences between populations and 82% of the variance within the population. This finding is in line with that obtained with other types of molecular marker such as isozymes and RAPD (Papa et al., 1998) and SSR (Bellucci et al., 2013) with F_{ST} values of 0.16, 0.11 and 0.14, respectively. The high proportion of variability within this population is also confirmed by the fact that, with a relatively small number of molecular marker (134 SSAP or 12 SSR; Rodriguez et al., 2012; Bellucci et al., 2013) all individuals analyzed (over 350) present a unique genetic profile, two individuals are still genetically distinguishable (Rodriguez et al., 2012; Bellucci et al., 2013). Rodriguez et al. (2012) found four well-defined genetic groups of populations, with and a clear structure according to the geographical site of origin. Moreover, they found a similarity between individuals geographically close, indicating the

existence of exchange of seed among farmers. The result also suggests the existence of diversifying selection.

In this regard, Merilä and Crnokrak (2001) proposed a model that can be particularly informative. This model compare the degree of divergence of populations for neutral molecular markers (F_{ST}) and quantitative traits (Q_{ST}). In fact, according to the theoretical model is expected that in a survey neutral $F_{ST} = Q_{ST}$; in a survey in which balancing selection (Dobzhansky and Dobzhansky, 1970) $Q_{ST} < F_{ST}$; and in a survey in which divergent selection $Q_{ST} > F_{ST}$. For the Sardinian materials Q_{ST} equal 0.40 was calculated, the degree of divergence medium for quantitative traits was 2-3 times higher than the F_{ST}, supporting the hypothesis of selection for local adaptation in different environments. Finally, the analysis of linkage disequilibrium showed rather low levels (13% of pairs of locus, with $P < 0.01$) and a decay with the distance of the map. The analysis of multilocus LD showed that genetic drift and founder effect have played an important role in shaping the genetic diversity present in this landrace, although the epistatic selection may also have played a role in this system. Overall, the results suggest that these populations could be used profitably for gene mapping by association mapping (Rodriguez et al., 2012).

Research objectives

The adaptation study of important agricultural species such as barley and the evolution of populations over time is at the basis of modern conservation techniques and the use of agricultural genetic resources also for the purpose of medium- and long-term genetic improvement. Given the potential of experimental evolution in the context of dynamic conservation "*ex-situ*" of a barley landrace population, the main objectives of this study were:

- to understand the amount and distribution of genetic diversity within a landrace population;
- to understand how to preserve the level of genetic diversity through conservation (ex-situ or in-situ);
- to evaluate the adaptation, i.e. the evolutionary forces acting during this process (selection and/or genetic drift);
- to identify interesting genes/traits under selection for future plant breeding programs.

Materials and methods

Plant material

This study evaluated a Sardinian six-rowed landrace population of barley, called Erdas from the name of the farmer that cultivated this population at time of sampling. This landrace belong to landraces population known as ‘‘S’orgiu sardu’’ by local farmers (Attene et al., 1996) that are characterized by a wide genetic diversity, adaptation to the local environment, and relevant interest as a source of germplasm for plant breeding (Papa et al., 1998; Rodriguez et al., 2008; Rodriguez et al., 2012).

This population was sampled in 1999 in the Sinis Penisula in the province of Oristano (Sardinia, Italy) (Figure 2).

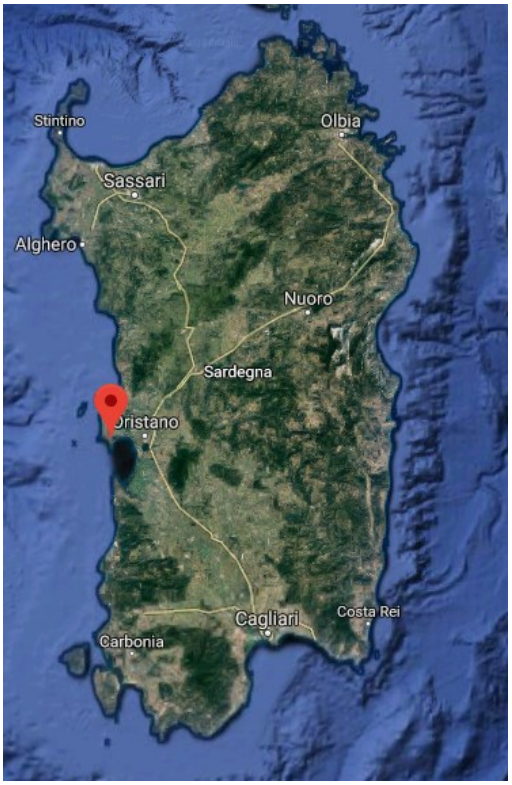

Figure 2. Satellite view of the region of Sardinia. Red arrow the red arrow indicates the Sinis Peninsula (Oristano).

Field experiment

In winter of the 1999, seeds collected were used to sow about 1.500 m² of barley in the experimental fields of Polytechnic University of Marche. In the spring of 2000, at harvesting

time, the field was subdivided in three plots of different size, 50 m^2 , 500 m^2 and 5000 m^2, respectively, with two replicates (A and B), for a total of six plots. In winter of the 2000, seeds collected from the plots obtained from the subdivided starting plot of 1500 m^2 were sown according to the scheme shown in Figure 3. Around each plot were sown rows of durum wheat as edges to keep the plots separated in order to avoid cross contamination. Before harvesting, for each plot about 100 spikes were sampled and stored. The next generation was grown in 2001, and so on, from random samples of seeds taken from the harvest of the previous year. According to the experimental design described, the experiment was repeated until 2015. The plots was managed according to usual organic agricultural practice and no conscious selection was practiced at any time. However, it should be noted that, in the 15 years that the subpopulations (plots of different size) have been grown at Polverigi, Ancona (Italy), temperature, rainfall and many other factors of the environment have fluctuated sharply from year to year, and they have also fluctuated in longer cycles.

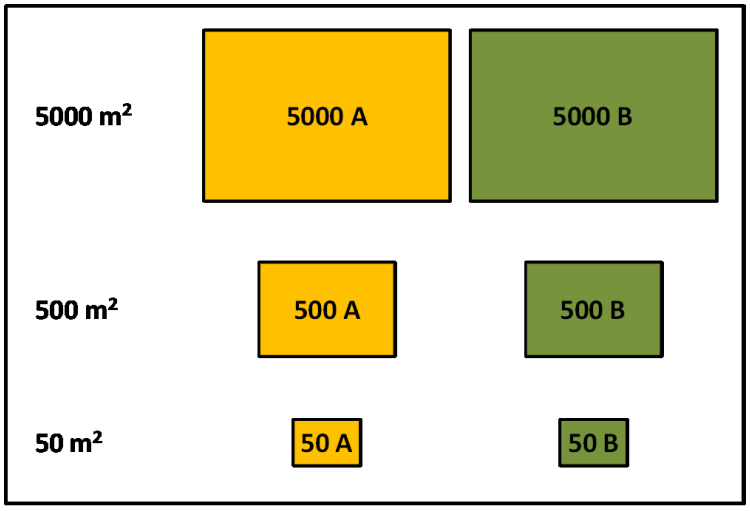

Figure 3. Experimental design

Sample collection

Fifteen years of experiment three different generations of cultivation were considered in the present study: 1999, 2007-08 and 2014-15, i.e. the beginning, the middle and the last year of the experiment, respectively. For each generation 84 individuals were analyzed (14 individuals for each plot of different size), for a total of 252 individuals (Table 1). Each individual derived from a single seed randomly chosen from the 100 spikes sampled in the field before harvest and stored as described above.

Year of collection	Sample	Subpopulation	Sample size subpop	Plot size
1999	84	1999_50_A	14	50 m^2
		1999_50_B	14	50 m^2
		1999_500_A	14	500 m^2
		1999_500_B	14	500 m^2
		1999_5000_A	14	5000 m^2
		1999_5000_B	14	5000 m^2
2008	84	2008_50_A	14	50 m^2
		2008_50_B	14	50 m^2
		2008_500_A	14	500 m^2
		2008_500_B	14	500 m^2
		2008_5000_A	14	5000 m^2
		2008_5000_B	14	5000 m^2
2015	84	2015_50_A	14	50 m^2
		2015_50_B	14	50 m^2
		2015_500_A	14	500 m^2
		2015_500_B	14	500 m^2
		2015_5000_A	14	5000 m^2
		2015_5000_B	14	5000 m^2
All	252			

Table 1. Number of individuals for each year of cultivation and plot of different size considered.

DNA extraction

A single seed of the 252 selected individuals were pre-germinated in Petri plate (24h at 4°C and than at room temperature) and transferred in pots with universal organic soil. For genomic DNA extraction, about 5 mg of fresh young leaves material was harvested. DNA was extracted from each individual by using the DNeasy® 96 Pant Kit (Qiagen) following the protocol provided. The quality and the quantity of extracted DNA was evaluated. A volume of 20 µl (concentration of about 60 ng/µl) for each sample were sent for genotyping to TraitGenetics GmbH (Gatersleben, Germany).

Molecular data

Molecular Markers

Genome analysis based on the detection of molecular markers is a considerable branch of interest and great potential opened up by genetic biotechnologies. This section allows an

ever better knowledge of nucleic acids structure, organization and function. Molecular markers are able to co-segregate with the genes and inherited according to the Mendelian models, allowing to build saturated genetic maps of many species of agrarian interest.
A molecular marker is a DNA fragment that can be ascribable to a genomic locus and with its presence unequivocally distinguishes the chromosomal fragment with which it is identified. It is not related to a specific genes activity, but is able to directly detect polymorphisms in the nucleotide sequence, differences due to insertions, deletions, translocations, point mutations.
Molecular markers are not influenced by the environment and plant phenology, as they are based on differences in DNA, and are abundant in the genome, also covering introns and regulatory regions. Among the many advantages, they do not have epistatic or pleiotropic effects, and are also inherited according to the Mendelian models, allowing in some cases to distinguish an heterozygous genotype from homozygous one. Molecular markers can be divided into two main groups: markers based on molecular hybridization (RFLP and VNTR) and PCR-based markers (RAPD, AFLP, STS, SSR and SNP).

SNPs markers

SNP markers (Single Nucleotide Polymorphisms), allow to highlight polymorphisms attributable to differences for single nucleotides (Barcaccia et al., 2000). For their detection it is necessary to amplify specific regions of genes of interest with optimal primers in different genotypes, sequencing the amplification products and align them in order to highlight any differences due to point mutations (insertions, deletions, substitutions). It is possible to conduct an in silico research of SNPs, comparing a large amount of EST (Expressed Sequence Tag) sequences and/or genomic clones from different varieties and deposited in public databases. Given the frequency with which point mutations occur, SNPs are considered to be the markers with the highest detectable genomic polymorphism potential. Moreover, thanks to other advantages such as codominance and high repeatability, these markers are considered the most important class of markers with different possible applications in the vegetable area, listed below:
- Construction of genetic maps
- Genetic mapping and assisted selection
- Characterization and dissection of genetic variability
- Typographic and varietal identification

- Identification of genetically modified organisms
- Analysis of gene expression and cloning of gene sequences

Barley 50k iSelect SNP Array

High-density SNP genotyping arrays are widely used to detect marker-trait associations in QTL (Quantitative Trait Locus) mapping experiments and genome-wide association studies (GWAS) (Cook et al., 2012; Jia et al. , 2013; Tian et al., 2011; Zhao et al., 2011). Advances in new generation sequencing have facilitated the discovery of SNPs throughout the genome (Berkman et al., 2012; Chia et al., 2012; Xu et al., 2012), transcriptome (Allen et al., 2011; Cavanagh et al., 2013; Oliver et al., 2013) or reduced sequence representation in different populations of individuals (Elshire et al., 2011; Poland et al., 2012; Saintenac et al., 2013; Van Poecke et al. , 2013). SNP arrays may be subject to errors of evaluation due to SNPs preselection in small populations (Albrechtsen et al., 2010), reduced computational requirements for downstream data processing, high call frequency, low percentage of error and ease of use of SNP-based platforms on a genotyping tool. The SNP genotyping arrays for the characterization of the germplasm are now available for many important species of cultivated plants, for example wheat (Winfield et al., 2016), rice (Chen et al., 2014), maize (Ganal et al., 2011), potato (Vos et al., 2015), rapeseed (Clarke et al., 2016; Mason et al., 2017), apple (Bianco et al., 2014, 2016), tomato (Sim et al., 2012), and others reviewed in Ganal et al. (2012).

The majority of SNPs on the array have been extracted from variants called in exome capture data of a wide range of European barley germplasm. Were used the recently published barley pseudomolecule assembly to map the exome capture data, which allowed to generate markers with accurate physical positions and detailed gene annotation.

The development of high throughput genotyping platforms in barley has taken place since 2006 with two Illumina GoldenGate assays (Fan et al., 2003) that featured 1,572 SNP markers each (Close et al., 2009), followed in 2009 by 9k Illumina Infinium iSelect Custom Genotyping BeadChip (Comadran et al., 2012) included 2,832 markers developed for the previous GoldenGate assays and 5,010 additional markers based on SNP marker discovery in Illumina RNAseq data from 10 UK elite cultivars. The continuous drop in sequencing costs and ease of identifying new markers, prompted to develop a new genotyping platform which provides more in-depth coverage of the complex 5.1 Gbp barley genome. The availability of a chromosomal-scale assembly with detailed genetic annotation allows

accurate information on the position of the markers and interpreted in the context of the associated genes. Bayer et al., 2017 developed a new 50k Illumina Infinium iSelect SNP genotyping array that includes most of the markers from the previous barley genotyping platforms and features 42,316 new SNPs derived from exome capture (EC) data of 170 spring and winter barley accessions carefully selected.

Barley whole genome EC data was made available for each of the 170 lines for the purposes of SNP discovery and iSelect development. Of a total of 519,742 robust SNPs were submitted to Illumina's Assay Design Tool to prioritise candidates for inclusion on the array, 295,642 SNPs passed the criteria. After filtering by including only SNPs with ≤ 20% missing samples and a design score of ≥ 0.6, and removing duplicates based on association with genes (according to SnpEff annotation), 16,957 SNPs, representing single genes, ensured maximum possible genic representation. Further SNPs were then added taking account of the relationship between the genetic and physical map in barley (Comadran et al., 2012; The International Barley Genome Sequencing Consortium, 2012).

The genetic map positions of the 9k SNPs were plotted on a Morex × Barke RIL population against the calculated physical positions of these SNPs (1H e 5H). This second set totalled 26,091 SNPs.

6,951 functional markers of the existing 9k SNP chip were included. Combining these with the 6,951 9k SNPs produced a total of 49,267 SNPs in the final design, which was subsequently submitted to Illumina for chip fabrication.

A free online resource is available (https://ics.hutton.ac.uk/50k) where are stored data such as the SNP source sequence, locus name, sequence orientation, SNP effects data and any additional information known on each locus.

Genotyping

All the 252 individuals of the sample collection were genotyped with Illumina Barley iSelect 50k SNP chip containing 49,267 single nucleotide polymorphisms (SNPs). SNP filtering and quality control resulted in a final markers set containing 44,040 loci.

The dataset was filtered for monomorphic SNPs using the TASSEL 5.2.48 software (Bradbury et al., 2007). SNPs with minor allele frequency (MAF) lower than 0.03 and genotypes with more than 25% missing were excluded from the analysis. Genotypes with heterozygous proportion lower than 1% were retained. Of 49,267 SNPs, 40,040 were scorable, of which 18,054 were polymorphic.

Statistical analysis

Summary statistics

Basic descriptive statistics for the eighteen subpopulations represented by plot of different size of the three years of experiment considered were calculated.
The minor allele frequency for each locus was calculated using the *minorAllele* function in the R package *adegenet* (Jombart et al., 2018).
The polymorphic information content (PIC) of the individual markers was calculated as explained by Botstein et al. (1980):

$$PIC = 1 - (\sum_{i=1}^{n} p_i^2) - \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} 2p_i^2 p_j^2$$

where *pi* and *pj* are allele frequencies at alleles *i* and *j*, respectively, and *n* is the number of alleles. The higher this value is, the more useful a marker is for distinguishing individuals and understanding relationships among them.

Observed heterozygosity

The observed heterozygosity (*Ho*) per locus for each individual was calculated, based on markers which passed the quality control, using the *basic.stats* function in the R package *hierfstat* (Goudet and Jombart, 2015) according to Nei (1978) statistics:

$$Ho = 1 - \sum_{k} \sum_{i} \frac{Pkii}{np},$$

where *Pkii* represents the proportion of homozygote *i* in sample *k* and *np* the number of samples.

Expected heterozygosity

The genetic diversity, or expected heterozygosity (H), was investigated within the whole collection and within the subpopulations for each chromosome, by calculation of the estimator of genetic diversity *Hi* (Nei, 1973).

$$Hi = 1 - \sum p_i^2$$

where p_i represents the allelic frequencies of allele *i.*

The mean expected heterozygosity within the subpopulations (H_S) was calculated the formula below:

$$H_S = 1 - \sum_{i=1}^{k} p_{i,s}^2$$

where $p_{i,s}$ represents the allelic frequencies of allele i of subpopulation s.
The mean expected heterozygosity (H_T) among the subpopulations was calculated as:

$$H_T = 1 - \sum_{i=1}^{k} \bar{p}_i^2$$

where p_i represents the allelic frequencies of allele i.
To measure the relative loss of genetic diversity in the plot A and B of subpopulation, according to Vigoroux et al. (2002), was used the parameter:

$$\Delta H = 1 - \left(\frac{Hi_{(A)}}{Hi_{(B)}}\right)$$

where $Hi_{(A)}$ and $Hi_{(B)}$ are the genetic diversity for the plot A and B, respectively.
The effect of subdivision of the population is measured with the Fixation index (F_{IS}; Wright, 1978):

$$F_{IS} = \frac{H_T - \bar{H}_S}{H_T}$$

that represent the mean reduction in heterozygosity of a subpopulation (relative to the total population) due to random genetic drift among subpopulations, i.e. a measure of the extent of genetic differentiation among subpopulations. The values of F_{IS} can range from 0.0 (no differentiation) to 1.0 (complete differentiation – subpopulations fixed for different alleles).

Pairwise genetic differentiation (F_{ST})

The average pairwise genetic differentiation among all pairs of subpopulations was calculated in R (http://www.r-project.org) following Weir and Cockerham (1984).

Population structure

Analysis of MOlecular VAriance (AMOVA)

Partitioning of the total genetic variance was assessed by analysis of molecular variance (AMOVA) using *Pegas* (Paradis, 2010), an R package (http://www.r-project.org). This function performs a hierarchical analysis of molecular variance as described in Excoffier et al. (1992). AMOVA is the analysis of variance for molecular markers and one uses a matrix of genetic distances among sampled individuals as a starting point. The output of AMOVA is a table that details the degrees of freedom (Df), sum of squared differences (Sum Sq) and resulting variation for comparisons within and among subpopulations. To test the significance of partitioned variance components a series of statistics called phi-statistics (Φ), which summarize the degree of differentiation between population divisions and are analogous to Φ-statistics (Excoffier, et al. 1992; Excoffier, 2001) was calculated. To test if populations are significantly different, we perform a randomization test using the function *rand.test()* from the *ade4* package using 1,000 permutations. This will randomly permute the sample matrices as described in (Excoffier et al., 1992).

Principal Component Analysis (PCA)

In order to determine population structure was used principal component analysis (PCA). PCA uses orthogonal transformation to convert a set of correlated variables into a set of linearly uncorrelated variables called principal components. PCA was performed by linear transformation of SNP data (all 18.054 SNPs) into a new coordinate system. The analysis was performed with compressed mixed linear model (CMLM; Zhang et al., 2010) implemented in the GAPIT R package (Lipka et al., 2012) for genomic prediction. The CMLM produces predictions at the group level. Then a group prediction for the associated individuals can be used (i.e., individuals in the same group share the same prediction). The greatest variance lies on the first coordinate (first principal component, PC1), the second greatest variance on the second coordinate (PC2) and so on.

Structure

A model-based clustering method was applied on SNPs data to infer genetic structure and define the number of clusters in the dataset using the software STRUCTURE ver. 2.3.4

(Pritchard et al., 2000). STRUCTURE uses a Bayesian approach to allocate proportions (Q) of K presumed ancestral populations contributing to each sample. Individuals may be assigned as the descendant of one population or having an admixture of two or more presumed ancestral populations. K is an important parameter that shapes the subsequent inferred sub-division of genetic structure, although it is a summary and convenience rather than a strictly defined hypothesis about the origin of a given set of individuals. Twenty runs per each cluster (K) ranging from 1 to 10 were carried. Each run consisted of a burn-in period of 1.000 steps followed by 1.000 MCMC (Monte Carlo Markov Chain) replicates assuming admixture model and correlated allele frequencies. No prior information was used to define the clusters. The model choice criterion in STRUCTURE to determine the true K is an estimate of the posterior probability of the data for a given K, Pr(X | K) (Pritchard et al. 2000), and is called LnP(D) or L(K). The most likely number of sub-groups K was identified by calculating an ad hoc statistic ΔK, based on the rate of change in the log probability of data between successive K-values as described by Evanno et al. (2005):

$$\Delta K = m\,(|L\,(K + 1)\text{-}2L\,(K) + L\,(K\text{-}1)|)/S\,[L\,(K)],$$

K denotes the assumed number of subgroups, L refers to the average of LnP(D) for the ten replications of the Kth STRUCTURE run, and S denotes the average standard deviation of the ten replications of the Kth STRUCTURE run.

The web-based program STRUCTURE HARVESTER v0.6.94 was used to process the STRUCTURE results files (Earl and vonHoldt, 2012).

Linkage disequilibrium

Linkage disequilibrium is the non-random association of alleles at two or more loci in a population. The statistics of linkage disequilibrium (LD) between genome wide SNP pairs was calculated as an r^2 (squared correlation) value among all pairs of SNP markers. The extent of LD was measured using the following equation:

$$r^2 = \frac{D^2}{freq(A1) * freq(A2) * freq(B1) * freq(B2)}$$

where *D = freq(A1_B1)*freq(A2_B2)–freq(A1_B2)*freq(A2_B1); freq(A1_B2)* is the frequency of the *A1_B1* haplotype, *freq(A2_B2)* is a *A2_B2* haplotype frequency and *freq(A1_B2)*, *freq (A2_B1)* is a frequency of haplotype *A1_B2* and *A2_B1* in the population,

respectively. The *freq(A1)*, *freq(A2)*, *freq(B1)*, and *freq(B2)* were observed frequencies of A1, A2, B1 and B2 alleles, respectively.

LD is a measure that has to be taken cautiously because of its variability across genetic backgrounds. It is therefore highly population-specific, and the fact that since LD varies across the genome, it is usually considered in terms of average values. Nonetheless, LD patterns can be a useful tool for understanding recombination, breeding and selection history.

LD was estimated by calculating the square value of correlation coefficient (r^2) between all pairs of markers with the software package TASSEL 5. The rate at which LD (r^2) decays with genetic distance was investigated for each of the 3 populations. *P* values for each r^2 estimate were calculated using 1000 permutations in TASSEL. The decay of LD across the entire genome was investigated by plotting pairwise r^2 values against the distance (Mbp) between markers overall chromosomes. A second degree smoothing curve (LOESS curve) was fitted to the plot and a critical r^2 value was derived from the distribution of r^2 values of unlinked markers (markers which were > 50 cM apart) by square root transformation of those r^2 values to obtain a normally distributed random variable. The parametric 95th percentile of that distribution was then taken as the population-specific critical value of r^2, beyond which LD was likely to be as a result of genetic linkage. The point where the intersection of the loess curve fit to the syntenic r^2 was then considered as an estimate of average LD decay (Breseghello and Sorrells 2006)

Signal of selection and determination of genomic regions under selection

To identify loci under selection between populations pairs was used the Weir and Cockerham (1984) measure of allele frequency differentiation, F_{ST}, as implemented in the R package *hierfstat* (Goudet and Jombart, 2015). F_{ST} was calculated for the following partitions of the data: 1) 1999 versus 2015, 2) 1999 versus 2008, and 3) 2008 versus 2015. An empirical genome-wide F_{ST} threshold value of 0.15 was used to identify SNPs with large differences in frequency relative to the genome-wide average. To identify the degree of differentiation that each comparison has with respect to other were reported the F_{ST} values for all pairwise comparisons.

Following recommendations to utilize multiple methods to detect loci under putative directional selection (Benestan et al., 2016; Rellstab et al., 2015), another approaches was used on the SNP dataset: BayeScan v.2.1 (Foll and Gaggiotti, 2008).

For BayeScan, individuals were grouped into subpopulations by plot of different size per

year of cultivation.

BayeScan uses a Bayesian approach to apply linear regression to decompose F_{ST} coefficients into population- and locus-specific components and estimates the posterior probability of a locus showing deviation from Hardy–Weinberg proportions (Foll and Gaggiotti, 2008). BayeScan analysis was based on 1:100 prior odds, with 100,000 iterations, a burn-in length of 50,000, a false discovery rate (FDR) of 0.05, and default parameters. Results were visualized in R. Outlier loci were functionally annotated through BARLEYMAP (http://floresta.eead.csic.es/barleymap) (Cantalapiedra et al., 2015) and Barley 50K SNP Platform - James Hutton Institute (http://bioinf.hutton.ac.uk/iselect/app/). SNPs identified as outliers in both two approaches were classified as putative adaptive loci.

Results

The original dataset included 252 individuals representative of the 3 years of experiment considered spanning 15 years of cultivation. The individuals were genotyped with Illumina Barley iSelect 50k SNP chip (Bayer et al., 2017). Of 44,040 SNPs scorable, 25,986 monomorphic SNPs and SNPs with ≥ 25% missing data in the full panel were removed. Genotypes with heterozygosity < 0.1% across SNPs are retained. After quality control, our data set consisted of 252 barley individuals and 18,054 SNPs.

The analysis of molecular distribution and Polymorphic Information Content (PIC) indicate high genetic diversity in the population sampled. Moreover, the genetic variability was maintained during several years of experiment. For each of the 18 subpopulations, the observed and expected heterozygosity, and the Fixation index (F_{IS}; Wright, 1978) were estimated. Pairwise comparisons between pairs of populations of the 3 years, and between pairs of subpopulations using the F_{ST} index (Weir and Cockerham, 1984) revealed a small significant differentiation for all pair of populations. AMOVA values pointed to within-population differentiation. Principal component analysis (PCA) also showed dispersed genotypes among different components suggesting very diverse genetic backgrounds. Additionally, the analysis of population structure using Bayesian method characterized three most probable clusters. The average LD decay for all chromosomes was estimated at approximately 2 Mbp, when the value of the cut off for r^2 was set to 0.2. Since, barley is a self-pollinated crop, a greater extent of LD is expected as compared to out-crossed crops.

Genetic diversity

Marker distribution and polymorphic information content

Out of all 18.054 SNPs, 17.983 were informative mapped SNPs considered for the 252 genotypes. Only 71 were unmapped. Those with MAF < 0.03 where excluded from the further analysis. A high percentage of SNP markers (~43% of total markers) had MAF within the range of 0.10 - 0.19 (Figure 4). The number of markers, average MAF and PIC for each chromosome are shown in Table 2. All seven chromosomes had mean MAF ≥ 0.24 and mean PIC values ≥ 0.26. This is an indication that the markers were highly diverse and hence very informative. Genome-wide distribution of SNP markers across the seven chromosomes of barley was not uniform. Chromosome 4H had the minimum number of markers (1763 SNPs) with an average of 15.42 markers per cM while chromosome 5H (3642 SNPs) had the

highest number of markers with an average of 21.49 markers per cM. The shortest chromosome was chromosome 4H with a total length of 114.3 cM while chromosome 5H was the longest having a genetic length of 169.5 cM. In Figure 5, it is possible to notice that the number of SNPs per chromosome is closely related to the length of the chromosome itself, as shown in Table 2, indicating that all loci were highly informative.

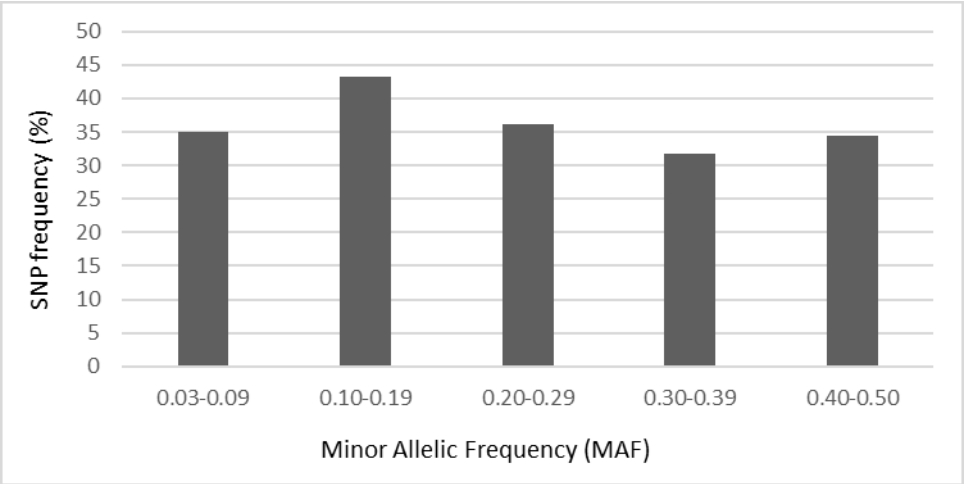

Figure 4. Distribution of SNP frequencies within the population barley landraces considered for analysis.

Chromosome	Length cM	N° of markers	Coverage (cM)	Mean MAF	Mean PIC
1H	132.8	1956	14.73	0.24	0.26
2H	149.4	3033	20.30	0.23	0.26
3H	154.9	3026	19.54	0.21	0.24
4H	114.3	1763	15.42	0.23	0.25
5H	169.5	3642	21.49	0.27	0.28
6H	126.6	2253	17.80	0.25	0.27
7H	140.7	2310	16.42	0.24	0.27

Table 2. Summary table of marker distribution, mean MAF and mean PIC values of 18.054 SNPs markers across the seven chromosomes of barley.

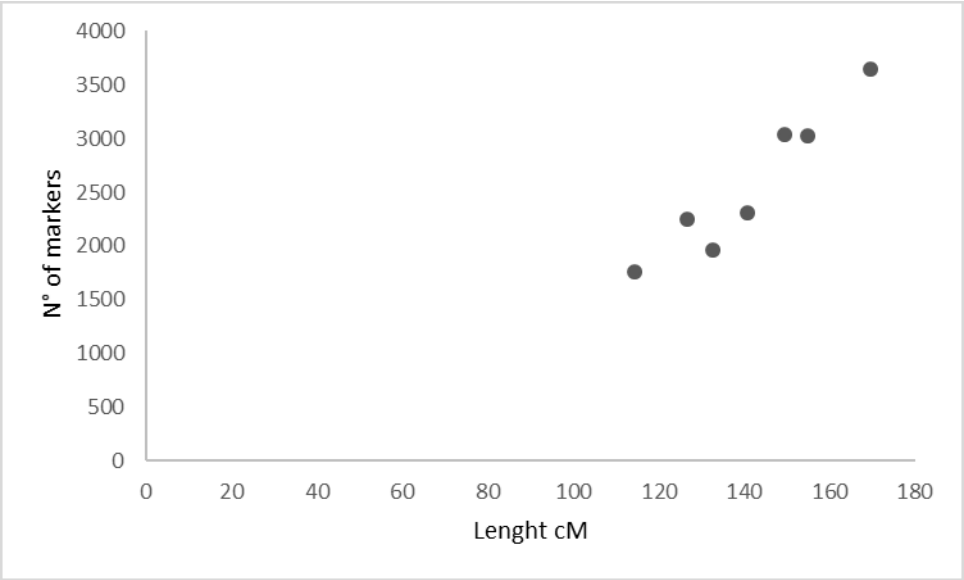

Figure 5. Correlation between length of chromosome to number of SNPs per chromosome

Observed heterozygosity

For all the eighteen subpopulations (Table 1) constituted by the plot of different size for the three years of experiment considered, observed heterozygosity (*Ho*) was calculated (Table 3). The *Ho* values ranged from 0.003 (2015_50_B) to 0.042 (2015_500_B).

As shown in Figure 6, observed heterozygosity was lower than the expected showing a departure from Hardy–Weinberg equilibrium (HWE). Low levels of heterozygosity reflect a low rate of outcrossing for barley species, which is expected for a crop with an inbreeding mating system. According to Abdel-Ghani et al. (2004), barley is a strictly autogamous species with an outcrossing rate of 0.59%.

Population	S	Ho
1999_50_A	14	0.010
1999_50_B	14	0.006
1999_500_A	14	0.009
1999_500_B	14	0.006
1999_5000_A	14	0.032
1999_5000_B	14	0.008
2008_50_A	14	0.006
2008_50_B	14	0.011
2008_500_A	14	0.011
2008_500_B	14	0.009
2008_5000_A	14	0.007
2008_5000_B	14	0.008
2015_50_A	14	0.009
2015_50_B	14	0.003
2015_500_A	14	0.010
2015_500_B	14	0.042
2015_5000_A	14	0.034
2015_5000_B	14	0.006
All	252	

Table 3. Mean value of observed heterozygosity (Ho) for each subpopulation. S = number of individuals.

Expected heterozygosity

Expected heterozygosity (*Hi*) varied from 0.285 to 0.308. Total genetic diversity of the Erdas landrace was apportioned into three components (Table 4); genetic diversity among individuals within plot of different size (*Hi* = from 0.285 to 0.308), genetic diversity within plot of same size A and B (H_S = from 0.293 to 0.319) and genetic diversity among plot of same size A and B (H_T = from 0.305 to 0.330). In details, it is possible to see that the proportion of heterozygosity (H_T) from the first year to the last year of experiment has not decreased but has fluctuated about apparent equilibrium values. F_{IS}, a measure of heterozygote deviation from Hardy–Weinberg equilibrium, ranged 0.031 to 0.087 (Table 4), indicating that individuals within each subpopulations are mostly homozygous, as would be expected for this inbreeding plant. These results show how the larger plots are able to maintain the genetic variability among the subpopulations, minimizing the action of the force of genetic drift.

Subpopulation	S	Hi	H_S	H_T	ΔH	F_{IS}
1999_50_A	14	0.302	0.302	0.311	-0.033	0.031
1999_50_B	14	0.302				
1999_500_A	14	0.285	0.293	0.302	0.068	0.030
1999_500_B	14	0.301				
1999_5000_A	14	0.306	0.302	0.313	-0.053	0.038
1999_5000_B	14	0.297				
2008_50_A	14	0.308	0.319	0.330	0.038	0.087
2008_50_B	14	0.330				
2008_500_A	14	0.288	0.294	0.305	0.030	0.039
2008_500_B	14	0.299				
2008_5000_A	14	0.304	0.304	0.315	0.004	0.041
2008_5000_B	14	0.305				
2015_50_A	14	0.297	0.301	0.314	0.012	0.043
2015_50_B	14	0.304				
2015_500_A	14	0.285	0.295	0.305	0.043	0.034
2015_500_B	14	0.305				
2015_5000_A	14	0.287	0.293	0.308	0.029	0.049
2015_5000_B	14	0.299				

Table 4. Mean value of expected heterozygosity within individuals *(Hi)* in each of 18 subpopulations; mean expected heterozygosity within plot of the same size (H_S); mean expected heterozygosity between plot of the same size (H_T); loss of genetic diversity *(ΔH);* fixation index *(FIS).*

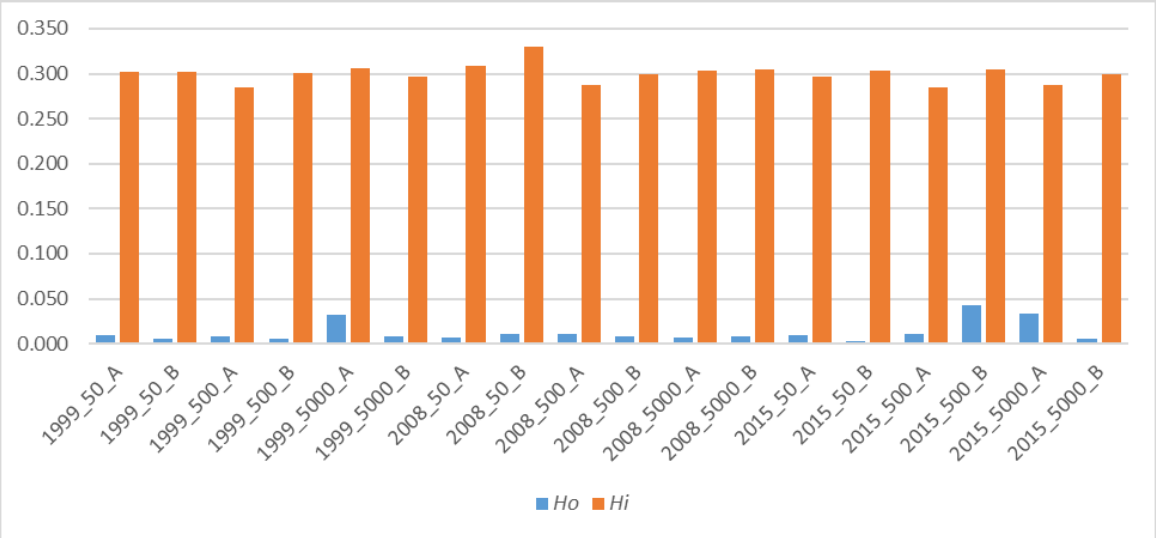

Figure 6. Comparison of observed (*Ho*, blue) and expected (*Hi*, orange) heterozygosity value within the 18 subpopulation of Erdas barley landraces of the three years of experiment considered.

Pairwise differentiation indices (F_{ST})

SNP data showed a relatively high level of genetic differentiation between populations. We computed two kinds of F_{ST} values between individuals: one using the comparison between pair of populations represented by three years of experiment; the other using comparison between pairs of subpopulations represented by eighteen plot of different size.

First case: the lowest levels of genetic differentiation (F_{ST} = 0.007) was detected between 2008 and 2015 populations (middle and last year of experiment, respectively), while the highest (F_{ST} = 0.015) was detected between 1999 and 2015 populations pair (first and last year of experiment, respectively) (Table 5).

Second case: the lowest levels of genetic differentiation (F_{ST} = 0.001) was detected for several pair of comparisons: plot 1999_50A and 1999_500B, 1999_50B and 2008_50B, 2008_50B and 2008_5000B, 2008_5000B and 2015_50B, 1999_500B and 2015_500A, 2015_50B and 2015_500B. While the highest (F_{ST} = 0.074) was detected between 1999_500A and 2015_5000A (Table 6).

F_{ST} values were roughly higher between more distant years. All subpopulations (plots of different size) weren't significantly differentiated from each other, except 1999_500A and 2015_5000A, relative to the others (Table 6).

The pattern of F_{ST} values based on subpopulations comparisons was similar to that one based on three populations. These results indicate that during the 15 years of cultivation there was an adaptation process, as we expected. Although this is a small differentiation, this is positive because it means that the effect of the genetic drift (i.e. random factors) is minimized. This difference is due to the natural selection (a positive signal of adaptation process).

Pairwise F_{ST}	**1999**	**2008**	**2015**
1999	**0**		
2008	**0.011**	**0**	
2015	**0.015**	**0.007**	**0**

Table 5. Pairwise population genetic differentiation (pairwise F_{ST} estimates) between pair of population (3 years of experiment) based on 18.054 SNP markers.

Pairwise F_{ST}	1999_50_A	1999_50_B	1999_500_A	1999_500_B	1999_5000_A	1999_5000_B	2008_50_A	2008_50_B	2008_500_A	2008_500_B	2008_5000_A	2008_5000_B	2015_50_A	2015_50_B	2015_500_A	2015_500_B	2015_5000_A	2015_5000_B
1999_50_A	0																	
1999_50_B	-0.012	0																
1999_500_A	0.013	-0.003	0															
1999_500_B	-0.008	-0.012	-0.015	0														
1999_5000_A	-0.012	-0.021	-0.001	-0.013	0													
1999_5000_B	0.008	0.001	0.005	-0.006	0.003	0												
2008_50_A	0.014	0.024	0.043	0.015	0.011	0.043	0											
2008_50_B	-0.003	0.001	0.021	-0.005	-0.007	0.023	-0.004	0										
2008_500_A	0.013	0.017	0.026	0.003	0.013	0.025	0.024	0.024	0									
2008_500_B	0.007	0.007	0.015	-0.006	0.017	-0.004	0.038	0.016	0.000	0								
2008_5000_A	0.006	-0.001	0.017	-0.006	-0.001	0.002	0.012	-0.010	0.029	0.005	0							
2008_5000_B	0.005	0.004	0.029	0.023	0.013	0.031	0.006	0.001	0.028	0.015	-0.008	0						
2015_50_A	0.009	0.002	0.011	-0.014	0.006	0.003	0.016	0.007	-0.003	-0.010	0.005	0.021	0					
2015_50_B	0.026	0.028	0.034	0.010	0.021	0.030	0.002	-0.003	0.035	0.014	-0.003	0.001	0.012	0				
2015_500_A	0.025	0.023	0.018	0.001	0.010	0.015	0.012	0.023	-0.008	0.016	0.024	0.036	-0.004	0.028	0			
2015_500_B	0.009	0.013	0.013	-0.011	0.008	0.010	0.010	0.003	-0.008	-0.003	0.003	0.020	-0.011	0.001	-0.006	0		
2015_5000_A	0.046	0.056	0.074	0.047	0.059	0.070	0.016	0.018	0.051	0.039	0.020	0.020	0.032	-0.010	0.055	0.013	0	
2015_5000_B	0.010	-0.001	0.006	-0.009	-0.003	0.014	-0.002	-0.002	-0.002	0.000	0.000	-0.006	-0.008	-0.004	-0.012	-0.010	0.024	0

Table 6. Pairwise population genetic differentiation (pairwise F_{ST} estimates) between pair of subpopulation (18) based on 18.054 SNP markers.

Population structure

AMOVA

A hierarchical analysis of molecular variance (AMOVA) was performed using different levels: between years of cultivation, between plot of different size within year, between individuals within plots, and within all individuals. The AMOVA analysis assigned most of the molecular variance to individuals within subpopulations (95.2%) and only 3.5% and 0.9% within individuals and between the three years, respectively (Table 7). Also, AMOVA showed that among populations (years), between individuals within subpopulation and within individuals differences are significant (Φ values equal to 0.009, 0.965 and 0.965, respectively; $P \leq 0.001$). Between subpopulations within year difference is not significant. The high level of differentiation among subpopulations appears to be due to differential changes in allele frequencies in different subpopulations. This result indicates that a much larger proportion of variation is due to within subpopulations diversity than among subpopulations diversity. As expected, the within individuals component was very low (due to the level of heterozygosity).

	Df	Sum Sq	Mean Sq	Variance	% Variation	Φ	*P*
Between year	2	41334.28	20667.1	52.1	0.9	0.009	*
Between plot Within year	15	178729	11915.3	26.5	0.5	0.005	ns
Between samples Within plot	234	2614265	11172.1	5486.3	95.2	0.965	*
Within samples	252	50279.55	199.5	199.5	3.5	0.965	*
Total	**503**	**2884607**	**5735.0**	**5764.4**	**100**		

Table 7. Analysis of the molecular variance using SNP markers, performed at different levels. (ns, not significant; *P<0.001)

Principal component analysis

A principal component analysis (PCA) for all 252 genotypes based on 18,054 SNPs to study population structure was computed. Of the 252 genotypes considered for analysis, 84 were collected in 1999, 84 in 2008 and 84 in 2015. The first two principal components explained about 15% of the total variation with the first PC explaining 9.39% of the variation and the second PC explaining 5.65% of the total genetic variation (Figure 7a). As expected, PCA showed that all the individuals share most of genetic variation, but also showed dispersed

individuals among different components suggesting very diverse genetic backgrounds. For a better visualization of the differentiation occurred in 15 years of cultivation, the mean of the two PCs for the all individuals within the 3 years of experiment considered (Figure 7b) and mean of the first two PCs for all individuals within the 18 subpopulations (Figure 7c) were plotted.

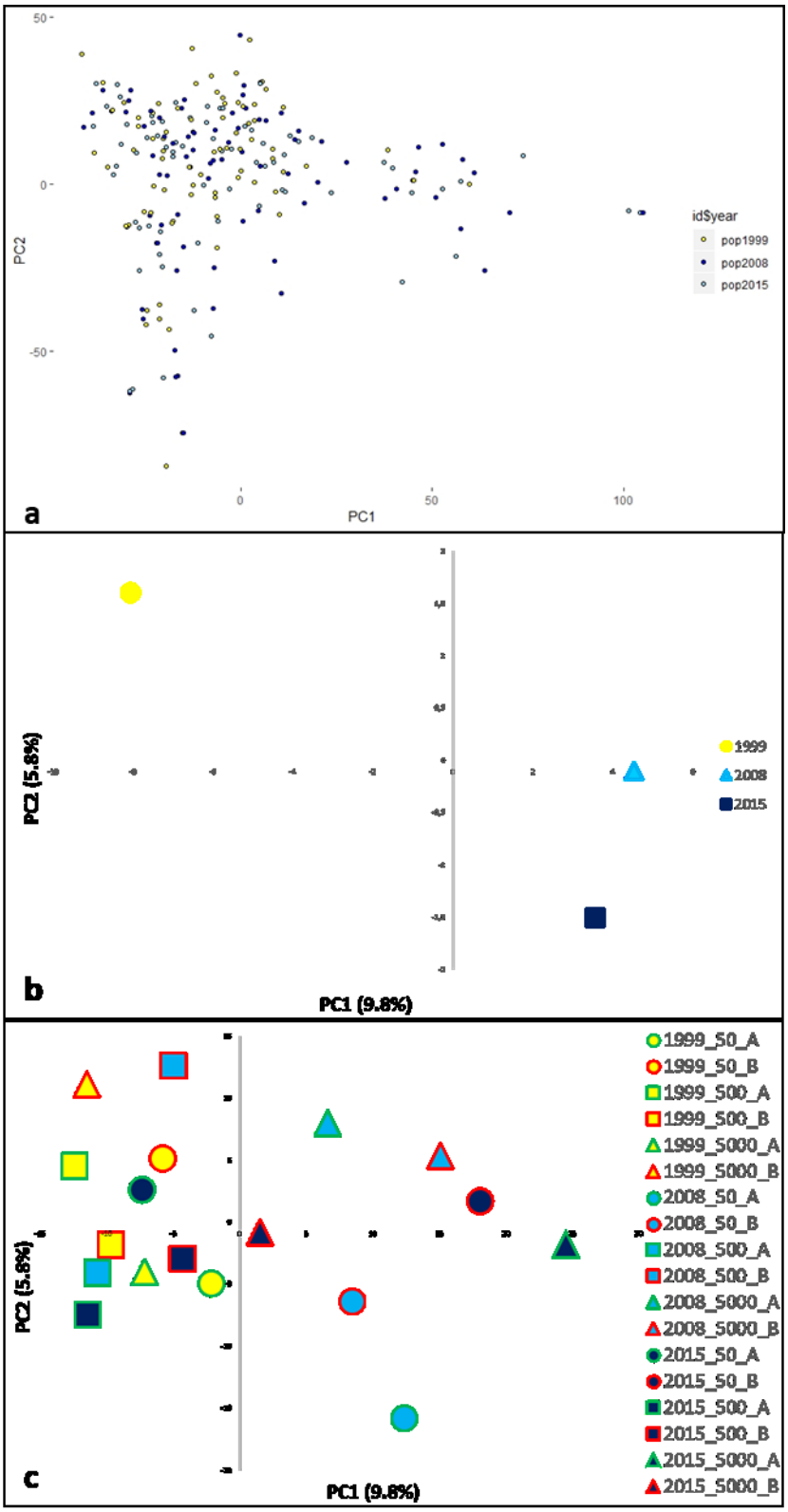

Figure 7. Population structure analysis of 252 individuals based on principal component analysis (PCA): first 2 PCs of all 252 individuals (a); mean of first 2 PCs considering 3 populations (b); mean of first 2 PCs considering 18 subpopulations.

Structure

To determine the maximum number of sub groups within the association panel, population structure was analyzed on all 252 individuals using the software package STRUCTURE. The analysis was performed for the three years of experiment considered (1999, 2008 and 2015). The program was tested from K=1 to 10 with 20 iterations per K. An individual was assigned to a K group if ≥ 60% of the genomic information was estimated to belong to this group. Delta K (ΔK, the mean log probability of the likelihood that an accession will belong to a group) was plotted against the number of subgroups.

For the first year of experiment (year 1999, consisting of 84 individuals), a maximum increase in ΔK was observed at K=3 (Figure 8a). At K=5 and K=7, ΔK reached its second and third maximum before flattening until K=10. This significant increase in delta K at K=3 indicates that the individuals could be separated into three major subgroups. The three main groups could be further divided into 5 and 7 subgroups.

For the middle year of experiment (year 2007-08, consisting of 84 individuals), a maximum increase in ΔK was observed at K=2 (Figure 8b). At K=5, ΔK reached its second maximum before flattening until K=10. This significant increase in ΔK at K=2 indicates that the individuals could be separated into two major subgroups. The two main groups could be further divided into 5 subgroups.

For the last year of experiment (year 2014-15, consisting of 84 individuals), a maximum increase in ΔK was observed at K=3 (Figure 8c), followed a flattening until K=10. This significant increase in delta K at K=3 indicates that the individuals could be separated into three major subgroups.

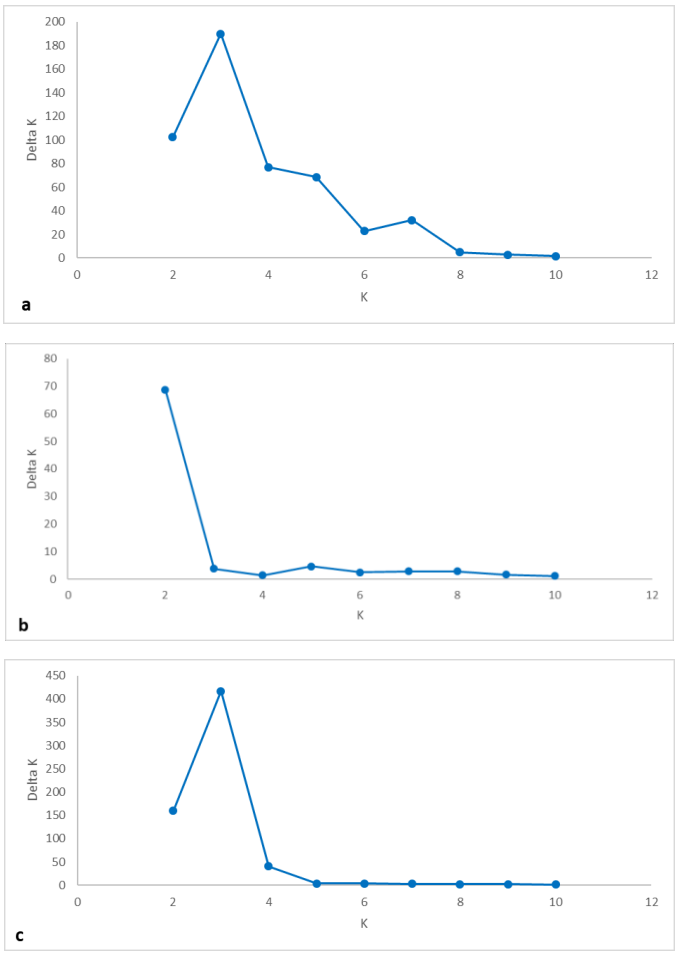

Figure 8. Plot of mean likelihood of delta K against the number of K groups. a) 1999, the highest peak observed at K=3 signifies the grouping of individuals into three groups while the small peak at K=5 and K=7 signifies further grouping of individuals into 5 and 7 groups. b) 2008, the highest peak observed at K=2 signifies the grouping of individuals into two groups while the small peak at K=5 signifies further grouping of individuals into 5 groups. c) 2015, the highest peak observed at K=3 signifies the grouping of accessions into three groups.

The analysis showed at the beginning of experiment (1999) a clear maximum for ΔK, K= 3 and as result all accessions were grouped into three different subgroups. STRUCTURE showed difficulty in assigning all individuals to each cluster, indicating some degree of admixture (Figure 9a). Subpopulations could be discriminated to a great extent on the basis of the geographical origin (Rodriguez et al., 2012).

In the first half of the experiment the best K=2, indicating a reduction in the number of clusters but an increase in the composition of cluster 1 (Figure 9b).

Subsequently, the last year of experiment the best K=3, indicating as a restabilization of the number of clusters with a different degree of admixture (Figure 9c).

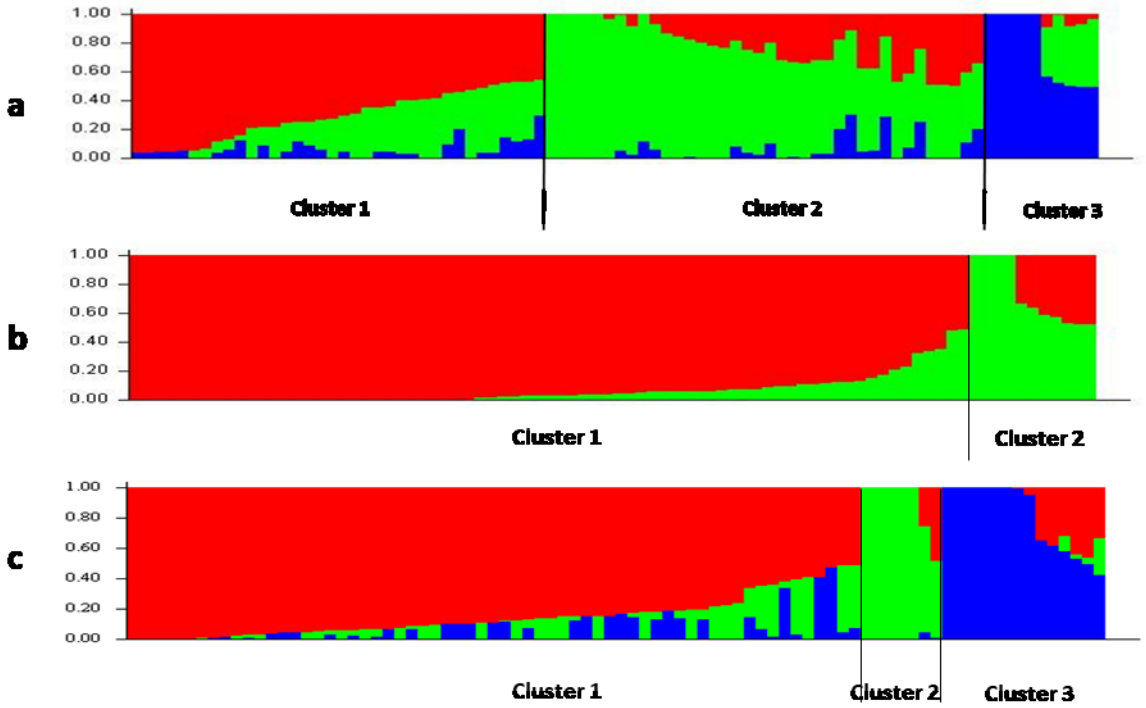

Figure 9. Q-matrix plot of STRUCTURE analysis: 1999 (a); 2008 (b); 2015 (c).

Linkage disequilibrium (LD) decay

Another important factor to consider is linkage disequilibrium (LD), which is the non-random association of alleles between two loci and shows the correlation between SNPs polymorphisms and their history of mutations and recombination (Flint-Garcia et al. 2003). LD is important since the rate of its decay in a given species determines the number and density of the molecular markers needed to perform GWAS (Rafalski, 2002). In many self-pollinated species such as barley where LD extends over long chromosomal distances (Malysheva-Otto et al. 2006), fewer markers are needed to cover the whole genome, whereas a higher marker density is needed when LD decays very rapidly in species such as maize where it declines to nominal levels within 1.5 kb (Remington et al. 2001).

To study the extent of LD decay within experimental panel of 252 individuals, pairwise correlations between all 18.054 markers were generated for each year of experiment considered separately and results plotted against the genetic distance between marker pairs. LD varied along each chromosome with regions of high LD interspersed with regions of low LD. Pair-wise correlation r^2 was found to decay rapidly with genetic distance along each chromosome. To estimate the average genome-wide LD decay, a critical r^2 value (0.2) was calculated for all unlinked loci pairs (> 50 cM). Beyond this threshold value, LD was assumed to be caused by genetic linkage. Average genome-wide LD decay was determined

by fitting a Loess curve and the point of interception between the Loess curve and the critical r^2 value (average genome-wide LD decay) was estimated to be around 2 Mbp (Figure 10).

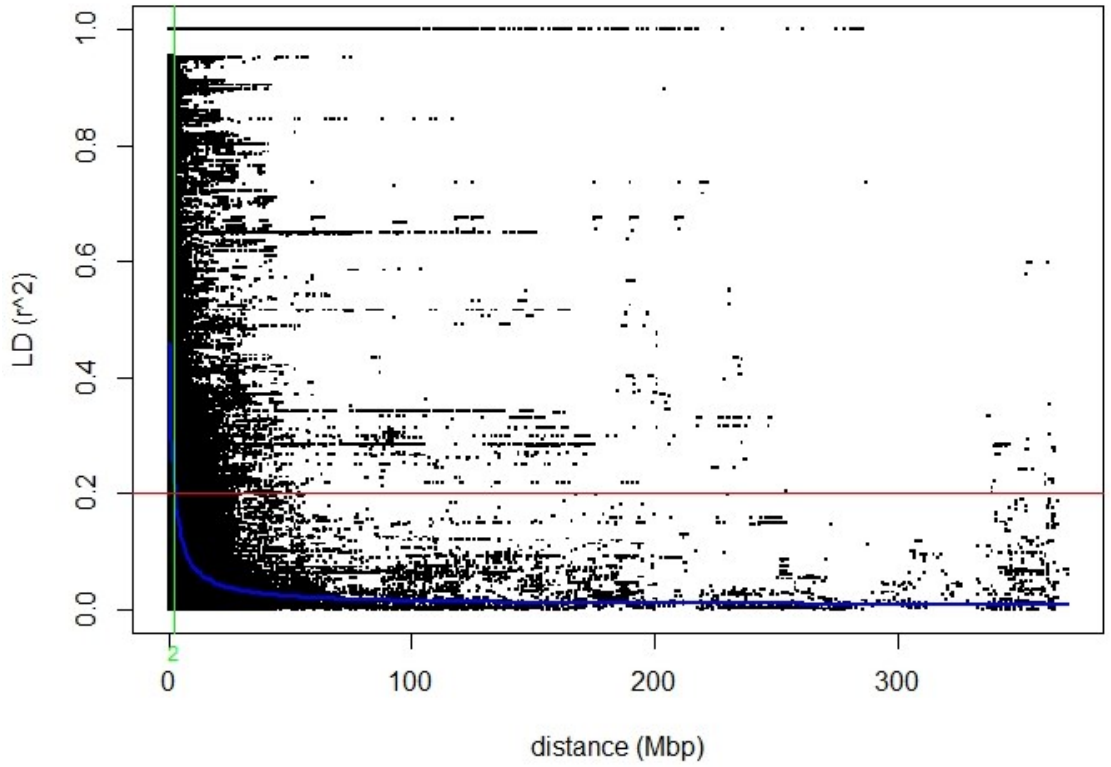

Figure 10. Average genome wide linkage disequilibrium (r^2) decay over genetic distance (Mbp) for the seven barley chromosomes. The blue line shows the non-linear regression line of the LD decay. Red line shows the intersection of an r^2 value of 0.2 with the regression line. The green line shows the average LD decay as a function of genetic distance in Mbp.

Loci under selection

To identify putative targets of long-term selection involved in the populations were adopted two theoretical approaches: the Weir and Cockerham (1984) measure of allele frequency differentiation, F_{ST} and BayeScan (see section Materials and Methods). F_{ST} was calculated for the following partitions of the data: 1) 1999 versus 2015, 2) 1999 versus 2008, and 3) 2008 versus 2015. Using *qqman* R package it was possible to display the genome-wide distribution of F_{ST} values (Figure 11). To identify the degree of differentiation that the comparison 1999 versus 2015 has with respect to other (1999 vs 2008 and 2008 vs 2015) were reported the F_{ST} values for all pairwise comparisons. An empirical genome-wide F_{ST} threshold value of 0.15 was used to identify SNPs with large differences in frequency relative to the genome-wide average. Other loci were detected as outliers by both tests, but, since

they only occurred in one pairwise comparison, they were not considered to be reliable outliers, but rather due to chance alone (Campbell and Bernatchez, 2004; Storz, 2005; Vasemägi et al., 2005). In total, 39 SNPs were considered to be under selection and for all was annotated the genes and their function that are located in the same positions (Table 8). When we conducted the pairwise population analysis using the BAYESCAN method, a total of 30 SNPs were indicated to be under selection (Figure 11), four of which (JHI_HV50K_2016_265281, JHI_HV50K_2016_44346, JHI_HV50K_2016_81439, JHI_HV50K_2016_461256) were detected potentially under directional selection when all populations were considered. Of these, one was located on the chromosome 1, one on chromosome 2, one on chromosome 4 and one on chromosome 7 (Table 8).

Based on known functions, many of these genes are crucial in adaptation to harsh environments and/or environmental changes. On chromosome 1 was found Ty1. The Ty1-copia type retrotransposons (long terminal repeat retrotransposons) are ubiquitous in plant kingdom (Flavell et al., 1992; Voytas et al., 1992; Hirochika and Hirochika, 1993). The activation of retrotransposons has been detected in specific developmental stages, with stress treatments, and during defense reactions against phytopathogens (Grandbastien, 1998).

On chromosome 2 was found a DNA repair protein, that could contribute to light-dependent and dark DNA repair pathways for the efficient elimination of cyclobutane pyrimidine dimers (CPDs) from the genomic DNA of barley leaf seedlings (Manova et al., 2016).

Another important gene found on chromosome 4 was that one coding for ZIFL1 - zinc induced facilitator-like 1. Remi et al., (2013) report that modulates root auxin-related processes in *Arabidopsis thaliana*. Involved in auxin efflux and acts as a positive regulator of shootward transport at the root apex and may mediate proton efflux from the vacuolar compartment. Also mediates drought stress tolerance by regulating stomatal closure.

On chromosome 7 was found FAR1 (FAR-RED-IMPAIRED RESPONSE1), a transposase-derived transcription factor (DNA sequences that can change their positions, transpose, within the genome) was discovered by Barbara McClintock (1984). It was initially identified as crucial components of phytochrome A (phyA)-mediated far-red (FR) light signaling in *Arabidopsis thaliana.* This components are the founding members of the FAR1-related sequence (FRS) family of transcription factors present in most angiosperms. This plant-specific gene family was termed the FRS gene family and is now emerging as a model system to investigate various crucial themes in the molecular domestication of transposable elements, including its mode of action and far-reaching ramifications in genome evolution and speciation. More recent studies revealed multifaceted roles of FAR1 in diverse

physiological and developmental processes such as UV-B signalling (Huang et al., 2012), circadian clock entrainment, flowering (Allen et al., 2006; Lin and Wang., 2004; Li et al., 2011), chloroplast biogenesis (Ouyang et al., 2011 ; Gao et al., 2013), chlorophyll biosynthesis (Tang et al., 2012), programmed cell death, ROS homeostasis, branching and plant architecture (Stirnberg et al., 2012).

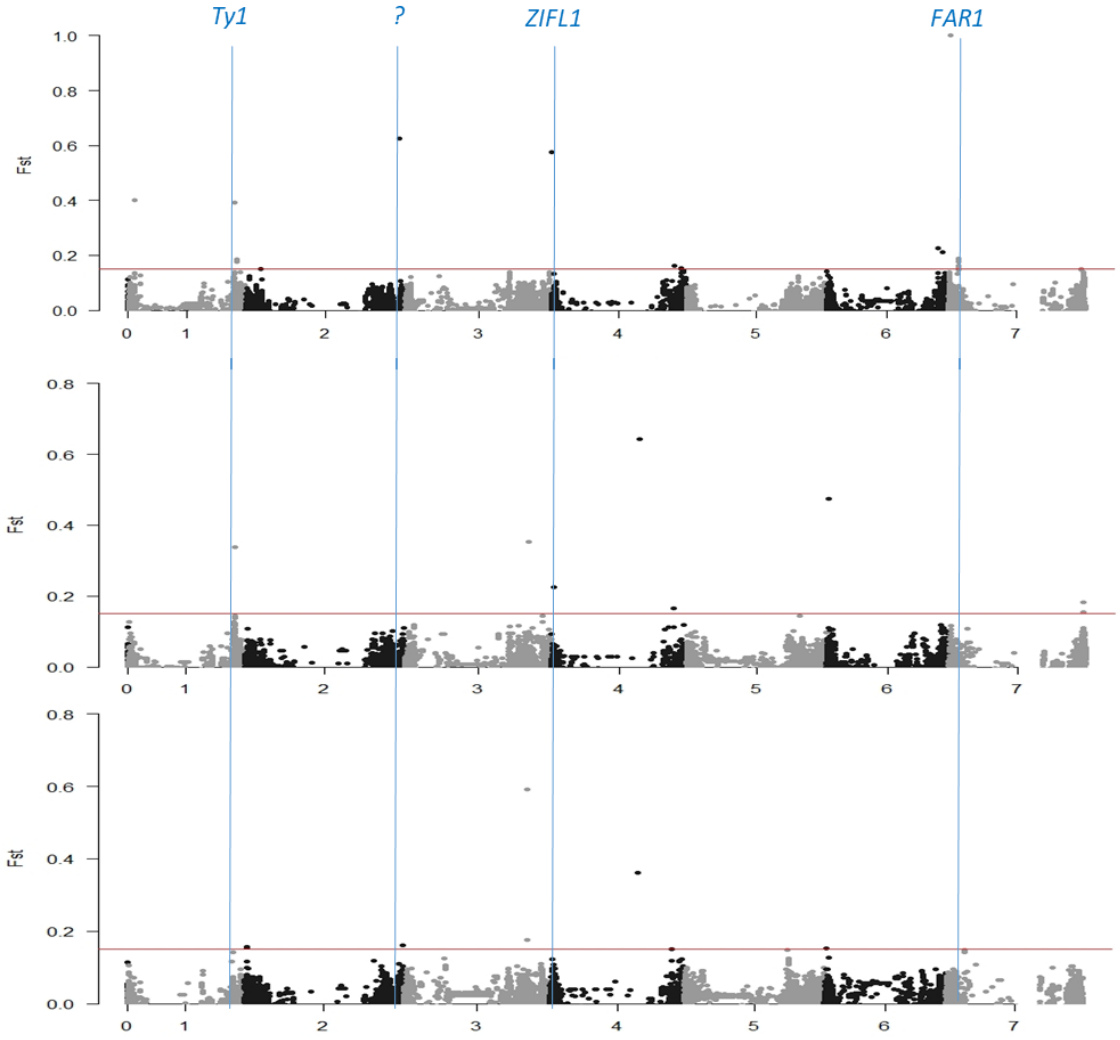

Figure 11. Genetic differentiation between populations 1999, 2008 and 2015. (A) Genetic differentiation measured by F_{ST} for populations 1999 and 2015 (A); population 1999 against populations 2008 (B); and population 2008 against populations 2015(C).

SNP	Chr	Position (Mb)	Map with genes	Gene class	Description	Fst (1999-2015)	Fst (1999-2008)	Fst (2008-2015)	BayeScan
JHI_HV50K_2016_17441	1	31.65	HORVU1Hr1G012670	HC_G	Protein FAR1-RELATED SEQUENCE 3	0.402	-0.061	-0.109	-
SCRI_RS_188218	1	510.60	HORVU1Hr1G075270	HC_G	Transcription elongation factor Spt5	0.392	0.339	-0.650	-
JHI_HV50K_2016_44342	1	518.80	HORVU1Hr1G077860	HC_G	Single-stranded DNA-binding protein	0.176	0.030	0.011	-
JHI_HV50K_2016_44343	1	518.80	HORVU1Hr1G077860	HC_G	Single-stranded DNA-binding protein	0.176	0.030	0.011	-
JHI_HV50K_2016_44344	1	518.80	HORVU1Hr1G077861	HC_G	Single-stranded DNA-binding protein	0.180	0.030	0.013	-
JHI_HV50K_2016_44345	1	518.80	HORVU1Hr1G077862	HC_G	Single-stranded DNA-binding protein	0.181	0.018	0.016	-
JHI_HV50K_2016_44346	1	518.80	HORVU2Hr1G024960	LC_TE	Transposon Ty1-OL Gag-Pol polyprotein	0.186	0.030	0.020	*
JHI_HV50K_2016_81439	2	76.56	HORVU2Hr1G024970	-	DNA repair protein-related	0.150	0.083	0.041	*
JHI_HV50K_2016_134176	2	739.15	HORVU2Hr1G116880 HORVU2Hr1G116910	HC_G HC_G	Protein kinase superfamily protein diacylglycerol kinase 5	0.625	1.000	-0.477	-
JHI_HV50K_2016_226108	4	0.64	HORVU4Hr1G000280	HC_G	Disease resistance protein (CC-NBS-LRR class) family	0.575	-0.562	0.480	-
JHI_HV50K_2016_259049	4	586.71	HORVU4Hr1G072900	HC_G	Exosome component 10	0.161	0.113	0.119	-
JHI_HV50K_2016_265281	4	619.70	HORVU4Hr1G081570	HC_G	zinc induced facilitator-like 1	0.154	0.010	0.122	*
JHI_HV50K_2016_265467	4	620.14	MLOC_81710 HORVU4Hr1G081750	HC HC_G	Similarity to late embryogenesis abundant protein late embryogenesis abundant protein-related / LEA protein-related	0.152	0.024	0.119	-
JHI_HV50K_2016_416262	6	536.22	0	0	0	0.227	0.052	0.096	-
SCRI_RS_43891	6	557.89	HORVU6Hr1G085170	HC_G	Photosystem II 10 kDa polypeptide	0.211	0.078	0.067	-
JHI_HV50K_2016_445507	7	10.16	HORVU7Hr1G007800	HC_G	undescribed protein	1.000	0.099	0.030	-
JHI_HV50K_2016_461205	7	46.83	HORVU7Hr1G026840 MLOC_13033	HC_G	Histone-lysine N-methyltransferase E(z) Histone-lysine N-methyltransferase MEDEA	0.182	0.012	0.083	-
JHI_HV50K_2016_461209	7	46.83	HORVU7Hr1G026840 MLOC_13033	HC_G	Histone-lysine N-methyltransferase E(z) Histone-lysine N-methyltransferase MEDEA	0.182	0.012	0.083	-
JHI_HV50K_2016_461210	7	46.83	HORVU7Hr1G026840 MLOC_13034	HC_G	Histone-lysine N-methyltransferase E(z) Histone-lysine N-methyltransferase MEDEA	0.182	0.012	0.083	-
JHI_HV50K_2016_461212	7	46.83	HORVU7Hr1G026840 MLOC_13035	HC_G	Histone-lysine N-methyltransferase E(z) Histone-lysine N-methyltransferase MEDEA	0.182	0.012	0.083	-
JHI_HV50K_2016_461213	7	46.83	HORVU7Hr1G026840 MLOC_13036	HC_G	Histone-lysine N-methyltransferase E(z) Histone-lysine N-methyltransferase MEDEA	0.182	0.012	0.083	-
JHI_HV50K_2016_461214	7	46.83	HORVU7Hr1G026840 MLOC_13037	HC_G	Histone-lysine N-methyltransferase E(z) Histone-lysine N-methyltransferase MEDEA	0.182	0.012	0.083	-
JHI_HV50K_2016_461215	7	46.83	HORVU7Hr1G026840 MLOC_13038	HC_G	Histone-lysine N-methyltransferase E(z) Histone-lysine N-methyltransferase MEDEA	0.182	0.012	0.083	-
JHI_HV50K_2016_461220	7	46.83	HORVU7Hr1G026840 MLOC_13039	HC_G	Histone-lysine N-methyltransferase E(z) Histone-lysine N-methyltransferase MEDEA	0.182	0.012	0.083	-
JHI_HV50K_2016_461223	7	46.83	HORVU7Hr1G026840 MLOC_13040	HC_G	Histone-lysine N-methyltransferase E(z) Histone-lysine N-methyltransferase MEDEA	0.182	0.012	0.083	-
JHI_HV50K_2016_461228	7	46.83	MLOC_13034	HC	FAR1-related sequence 5 LENGTH=788	0.152	0.005	0.069	-
JHI_HV50K_2016_461230	7	46.84	HORVU7Hr1G026850	LC_TE?	FAR1-related sequence 9	0.189	0.012	0.088	-
JHI_HV50K_2016_461238	7	46.84	MLOC_13034 HORVU7Hr1G026850	HC LC_TE?	FAR1-related sequence 5 LENGTH=788 FAR1-related sequence 9	0.182	0.012	0.083	-
JHI_HV50K_2016_461243	7	46.84	MLOC_13034 HORVU7Hr1G026850	HC LC_TE?	FAR1-related sequence 5 LENGTH=788 FAR1-related sequence 10	0.189	0.012	0.088	-
JHI_HV50K_2016_461244	7	46.84	MLOC_13034 HORVU7Hr1G026850	HC LC_TE?	FAR1-related sequence 5 LENGTH=788 FAR1-related sequence 11	0.182	0.012	0.083	-
JHI_HV50K_2016_461255	7	46.84	HORVU7Hr1G026860	HC_U	FAR1-related sequence 5 LENGTH=788 FAR1-related sequence 12	0.182	0.012	0.083	-
JHI_HV50K_2016_461256	7	46.84	HORVU7Hr1G026860	HC_U	FAR1-related sequence 5 LENGTH=788 FAR1-related sequence 13	0.182	0.012	0.083	*
JHI_HV50K_2016_461259	7	46.84	HORVU7Hr1G026860	HC_U	FAR1-related sequence 5 LENGTH=788 FAR1-related sequence 14	0.182	0.012	0.083	-
JHI_HV50K_2016_461261	7	46.84	HORVU7Hr1G026860	HC_U	FAR1-related sequence 5 LENGTH=788 FAR1-related sequence 15	0.182	0.012	0.083	-
JHI_HV50K_2016_461269	7	47.04	HORVU7Hr1G026860	HC_G	P-loop containing nucleoside triphosphate hydrolases superfamily protein	0.163	0.007	0.072	-
JHI_HV50K_2016_461278	7	47.04	HORVU7Hr1G026860	HC_G	P-loop containing nucleoside triphosphate hydrolases superfamily protein	0.182	0.012	0.083	-
JHI_HV50K_2016_461286	7	47.04	HORVU7Hr1G026880	HC_G	P-loop containing nucleoside triphosphate hydrolases superfamily protein	0.176	0.000	0.075	-
JHI_HV50K_2016_461293	7	47.04	HORVU7Hr1G026880	HC_G	P-loop containing nucleoside triphosphate hydrolases superfamily protein	0.182	0.012	0.083	-
JHI_HV50K_2016_509661	7	638.20	HORVU7Hr1G113720	LC_TE	Retrotransposon protein, putative, unclassified	0.151	0.100	-0.005	-

Table 8. Annotation of putative loci under selection.

Discussion

Landraces are defined as dynamic populations of a cultivated plant with an historical origin, distinct identity, often genetically diverse and locally adapted, and associated with a set of farmers' practices of seed selection and field management as well as with a farmers' knowledge base (Villa et al, 2005). They can be defined as a traditional variety with a high capacity to tolerate biotic and abiotic stresses, resulting in high yield stability and an intermediate yield level under a low-input agricultural system (Zeven, 1998). Farmers sowing, harvesting and saving a proportion of seed for subsequent sowing over millennia have enriched the genetic pool of crops by promoting intraspecific diversity (Frankel et al. 1998). Landraces are repository of gene pool that enrich biodiversity and maintain and stabilize ecosystem in a sustainable way to make it functional. The current industrial agriculture system may be the single most important threat to biodiversity (Dwivedi et al., 2016). Also, a serious consequence of biodiversity loss is the displacement of locally adapted landraces with adaptation traits to future climates by monocropping with genetically uniform hybrids and improved cultivars (Sarker and Erskine, 2003). In others words, modern agriculture has contributed to decreasing agricultural biodiversity. For this reason conservation of all gene pools is a high priority for sustaining food security and coping with current and future climate change effects.

This study focuses on the investigation of adaptation process of an important Sardinian barley landrace population. The study of the most important agricultural species and the evolution of populations over time is at the basis of modern conservation techniques. Moreover the use of these agricultural genetic resources is useful for the purpose of medium- and long-term genetic improvement.

The availability of large numbers of SNPs uniformly distributed across the genome has provided opportunities to refine the analysis of population structuring of genetic diversity and permit to obtain precise estimates of genetic differentiation based on very small sample sizes of random individuals (Willing et al., 2012).

The initial population was subdivided in eighteen subpopulations of different size (plots of 50, 500 and 5000 m^2) and each plot was kept isolated from the others in order to avoid cross contaminations. The experimental design was repeated for 15 years. This experiment was conducted in order to assess which forces acts during this type of conservation and if the dimension of population is enough to maintain the genetic variability that characterized these genetic resources.

When a population is divided into isolated subpopulations, there is less heterozygosity than there would be if the population was undivided. Founder effects acting on different demes (semi-isolated subpopulations) generally lead to subpopulations with allele frequencies that are different from the larger population. Also, these demes are smaller in size than the larger population; since allele frequency in each generation represents a sample of the previous generation's allele frequency, there will be greater sampling error in these small groups than there would be in a larger undifferentiated population. Hence, genetic drift will push these smaller demes toward different allele frequencies and allele fixation more quickly than would take place in a larger undifferentiated population.

Differences among subpopulations were detected regarding the degree of variation in the SNPs loci.

For all loci the expected heterozygosity was higher than the observed heterozygosity (Bartlett test, p-value < 2.2e-16) (Figure 5). The proportion of mean expected heterozygosity (H_T) from the first year (1999) to the last year (2015) of experiment has not decreased but has fluctuated about apparent equilibrium values. F_{IS}, a measure of heterozygote deviation from Hardy–Weinberg equilibrium, ranged 0.031 to 0.087 (Table 4), indicating that individuals within each subpopulations are mostly homozygous, as would be expected for this inbreeding plant. Ours results show how the larger plots are able to maintain the genetic variability among the subpopulations, minimizing the action of the force of genetic drift. The quantification of population differentiation based on genetic markers can highlight the relative role of evolutionary processes such as natural selection, genetic drift and gene flow for patterns of local adaptation (Brommer, 2011; Leinonen et al., 2013).

Fixation index (F_{ST}) is widely used to estimate genetic differentiation by analysing variance in allele frequency (Wright, 1949). F_{ST} is a fundamental measure of genetic differentiation and population structure currently defined for subdivided populations. F_{ST} in practice typically assumes the “island model”, where subpopulations have evolved independently from their last common ancestral population. The pairwise genetic differentiation comparison shows a small but significant higher level of genetic differentiation between 1999 and 2015 subpopulations (F_{ST} = 0.015) then a lower value between 2008 and 2015 subpopulations (F_{ST} = 0.007).

The structure of this experiment allows it to say that the observed genetic changes in plots of larger size (500 and 5000 m^2) were not due to genetic drift. In our study the action of the natural selection seems more likely acting. AMOVA analysis shows that most of the variation is due to differences between individuals within subpopulation (95.2%) and to a

small extent between years of cultivation (0.9%), while a rather limited amount of genetic variation (3.5%) is due to within all individuals. These results was also consistent with the F_{ST} values, thereby confirming that the populations analyzed are not genetically homogeneous, but show a moderate differentiation level according to the findings of Papa et al. (1998), Rodriguez et al. (2012), Bellucci et al. (2013) that this plant materials are highly variable.

Bayesian analysis and PCA pointed to similar results. Three admixed genetic clusters were identified, with a composition of cluster 1 increasing during the years. A particular case came out analyzing the population structure taking into consideration the 3 years of experiment studied separately. In the first year it was seen structure of genetic variation had three main genetic groups with most of the individuals were apparently admixed i.e. derived from a hybrid between different genetic groups. Our findings are in accordance with the work of Rodriguez et al. (2012), in which also this population was investigated. In the following years there was a reduction in the number of clusters (K=2) with an increase in the cluster 1 component, and then an increase in the last year (K=3). Despite the recent nature of the experimental dynamic conservation program, with just 15 generations of cultivation compared to the scale of natural evolution, a highly significant evolution was observed for major adaptive genes. Local adaptation is one of the expected mechanisms driving the divergence of subpopulations of the experimental program, but its expected effect in the long term is a local reduction in genetic diversity. While this reduction is expected to be reinforced by the high homozygosity of selfing populations, no drastic decrease in variability was observed within the experimental population. This is probably due to the low number of generations of evolution, as well as to the residual outcrossing rate of barley (< 1%, Briggs, 1978; Abdel-Ghani et al., 2004) and to mutations, which contributed to the maintenance of intra-population diversity.

More than 27 years ago, Richard Lenski and colleagues (1991) set up his Long-Term Evolution Experiment (LTEE), to study the adaptation of 12 replicate populations of the workhorse bacterium *Escherichia coli* to a new environment (glucose as the limiting resource). As Lenski, we studied the adaptation of a landrace population in a new environment (from Sardinia to Marche, Italy). The dynamics are interesting, and sometimes surprising, in several aspects. During the first 7 generations or so, the effect sizes of beneficial mutations were large and produced fitness trajectories with step-like dynamics. Over longer periods, the rate of improvement slowed substantially. This trend might suggest that fitness is approaching some upper bound.

Conclusions

The present study addresses the effects of a strategy of *ex-situ* conservation based on dynamic management on population evolution, adaptation and diversity. Results from a study over 15 years of experimental dynamic conservation of a Sardinia barley landrace are presented for the first time. An experimental context is coupled with on-farm management in this study. In collaboration with a farmers', it was possible to study the effects of *ex-situ* dynamic conservation (conservation and selection) on diversity and adaptation. Results show that this dynamic conservation is a key element for the long-term conservation and use of agricultural biodiversity. In concluding, the *ex-situ* on-farm conservation needs more investigation and deserves more attention in industrialised countries.

Bibliography

Abdel-Ghani A.H., Parzies H.K., Omary A., Geiger H.H. (2004). Estimating the outcrossing rate of barley landraces and wild barley populations collected from ecologically different regions of Jordan. *Theor Appl Genet* 109: 588–595.

Albrechtsen A., Nielsen F.C. and Nielsen R. (2010). Ascertainment biases in SNP chips affect measures of population divergence. *Mol. Biol. Evol.* 27, 2534–2547.

Allard R.W., Kahler A.L., and Weir B.S. (1972). The effect of selection on esterase allozymes in a barley population. *Genetics.* 72(3), 489-503

Allel D., Ben-Amar A., Lamine M. and Abdelly C. (2017). Relationships and genetic structure of North African barley (*Hordeum vulgare* L.) germplasm revealed by morphological and molecular markers: Biogeographical considerations. *South African Journal of Botany*, *112* (September), 1–10. https://doi.org/10.1016/j.sajb.2017.05.005.

Allen A.M., Barker G.L.A., Berry S.T., Coghill J.A., Gwilliam R., Kirby S., Robinson P., Brenchley R.C., D'Amore R., McKenzie N., Waite D., Hall A., Bevan M., Hall N. and Edwards K.J. (2011). Transcript-specific, single-nucleotide polymorphism discovery and linkage analysis in hexaploid bread wheat (*Triticum aestivum* L.). *Plant Biotechnol. J.* 9, 1086–1099.

Allen T., Koustenis A., Theodorou G., Somers D.E., Kay S.A., Whitelam G.C., and Devlin P.F. (2006). *Arabidopsis FHY3* specifically gates phytochrome signaling to the circadian clock. *The Plant Cell*, 18(10), 2506-2516.

Angessa T.T. and Li C. (2015). Exploration and Utilization of Genetic Diversity Exotic Germplasm for Barley Improvement. Exploration, Identification and Utilization of Barley Germplasm:223.

Attene G., Ceccarelli S. and Papa R. (1996). The barley (*Hordeum vulgare* L.) of Sardinia, Italy. *Genetic Resources and Crop Evolution*, 43(5), 385- 393.

Badr A., Sch R., El Rabey H., Effgen S., Ibrahim H., Pozzi C., Rohde W. and Salamini F. (2000). On the origin and domestication history of barley (*Hordeum vulgare*). *Mol Biol Evol,* 17:499-510.

Baik B.K. and Ullrich S.E. (2008). Barley for food: Characteristics, improvement, and renewed interest. *J Cereal Sci*, 48:233-242.

Barcaccia G., Lucchin M. and Parrini P. (2000). Analisi del genoma mediante marcatori molecolari: I. fondamenti metodologici. Sementi Elette, 46(5), 5-16.

Bayer M.M., Rapazote-Flores P., Ganal M., Hedley P.E., Macaulay M., Plieske J., Ramsay L., Russell J., Shaw P.D., Thomas W. and Waugh R. (2017). Development and Evaluation of a Barley 50k iSelect SNP Array. *Front. Plant Sci*. 8:1792. doi: 10.3389/fpls.2017.01792.

Bellucci E., Rau D., Nanni L., Ferradini N., Giardini A., Rodriguez M., Attene G. and Papa

R. (2013). Population structure of Barley landrace populations and gene-flow with modern varieties. *PLoS ONE* 8:e83891. doi: 10.1371/journal.pone.0083891

Benestan L.M., Ferchaud A.L., Hohenlohe P.A., Garner B.A., Naylor G.J., Baums I.B., Schwartz M.K., Kelley J.L. and Luikart G. (2016). Conservation genomics of natural and managed populations: building a conceptual and practical framework. *Molecular ecology*, 25(13), 2967-2977.

Bennett A.F. (2003). Experimental Evolution and the Krogh Principle: Generating Biological Novelty for Functional and Genetic Analyses1.*Physiological and Biochemical Zoology*, 76(1), 1-11.

Berkman P.J., Lai K., Lorenc M.T. and Edwards D. (2012). Next-generation sequencing applications for wheat crop improvement. *Am. J. Bot.* 99, 365–371.

Bianco L., Cestaro A., Linsmith G., Muranty H., Denancé C., Théron A., Poncet C., Micheletti D., Kerschbamer E., Di Pierro E.A., Larger S., Pindo M., Van de Weg E., Davassi A., Laurens F., Velasco R., Durel C. and Troggio M. (2016). Development and validation of the Axiom®Apple480K SNP genotyping array. *Plant J*, 86: 62-74. doi:10.1111/tpj.13145

Bishop C.W. (1936). Origin and early diffusion of traction-plough. *Antiquity* 10: 261-278

Bitocchi E., Nanni L., Rossi M., Rau D., Bellucci E., Giardini A., Buonamici A., Vendramin G.G. and Papa R. (2009). Introgression from modern hybrid varieties into landrace populations of maize (*Zea mays* ssp. *mays* L.) in central Italy. *Molecular Ecology*, *18*(4), 603–621. https://doi.org/10.1111/j.1365-294X.2008.04064.x

Botstein M., White R.L., Skolnick M. and Davis R.W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* 32, 314–331.

Bradshaw J. E. and Ramsay G. (2005). Utilisation of the Commonwealth Potato Collection in potato breeding. *Euphytica*, *146*(1–2), 9–19. https://doi.org/10.1007/s10681-005-3881-4

Briggs D.E. (1978). Barley. London: Chapman and Hall.

Bradbury P.J., Zhang Z., Kroon D.E., Casstevens T.M., Ramdoss Y. and Buckler E.S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*, 23(19), 2633-2635.

Breseghello F. and Sorrells M.E. (2006). Association analysis as a strategy for improvement of quantitative traits in plants. *Crop Science*, 46(3), 1323-1330.

Brommer J.E. (2011). Whither PST? The approximation of QST by PST in evolutionary and conservation biology. *Journal of evolutionary biology*, 24(6), 1160-1168.

Brown A.H. (1999). The genetic structure of crop landraces and the challenge to conserve them in situ on farms. In *Genes in the Field* (pp. 36-55). CRC Press.

Brush S.B. (1995). In situ conservation of landraces in centers of crop diversity. *Crop*

science, 35(2), 346-354.

Burke M.K. and Rose M.R. (2009). Experimental evolution with *Drosophila*. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 296(6), R1847-R1854.

Campbell D. and Bernatchez L. (2004). Generic scan using AFLP markers as a means to assess the role of directional selection in the divergence of sympatric whitefish ecotypes. *Molecular Biology and Evolution*, 21(5), 945-956.

Cantalapiedra C.P., Boudiar R., Casas A.M., Igartua E., Contreras-Moreira B. (2015). BARLEYMAP: physical and genetic mapping of nucleotide sequences and annotation of surrounding loci in barley. *Mol Breeding* 35:13 DOI 10.1007/s11032-015-0253-1

Carović-Stanko K., Liber Z., Vidak M., Barešić A., Grdiša M., Lazarević B., and Šatović Z. (2017). Genetic Diversity of Croatian Common Bean Landraces. *Frontiers in Plant Science*, *8* (April), 1–8. https://doi.org/10.3389/fpls.2017.00604

Cavanagh C.R., Chao S., Wang S., Huang B.E., Stephen S., Kiani S., Forrest K., Saintenac C., Brown-Guedira G.L., Akhunova A., See D., Bai G., Pumphrey M., Tomar L., Wong D., Kong S., Reynolds M., da Silva M.L., Bockelman H., Talbert L., Anderson J.A., Dreisigacker S., Baenziger S., Carter A., Korzun V., Morrell P.L., Dubcovsky J., Morell M.K., Sorrells M.E., Hayden M.J. and Akhunov E. (2013). Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. *Proc. Natl Acad. Sci*. USA, 110, 8057–8062.

Ceccarelli S., Grando S., Maatougui M., Michael M., Slash M., Haghparast R., Rahmanian M., Taheri A., Al-Yassin A., Benbelkacem A., Labdi M., Mimoun H., Nachit M. (2010). Plant breeding and climate changes. *J Agr Sci* 148:627-637.

Ceccarelli S. and Grando S. (2000). "Barley Landraces from the Fertile Crescent: A lesson for Plant Breeders", in S.B.Brush (ed.), *Genes in the Field*: On-Farm Conservation of Crop Diversity, International Development Research Centre, Boca Raton, Florida, pp.51-76.

Chen H., Xie W., He H., Yu H., Chen W., Li J., Yu R., Yao Y., Zhang W., He Y., Tang X., Zhou F., Deng XW., Zhang Q. (2013). A high-density SNP genotyping array for rice biology and molecular breeding. Mol Plant. 7:541–553. DOI: 10.1093/mp/sst135.

Chia J.-M., Song C., Bradbury P.J., Costich D., de Leon N., Doebley J., Elshire R.J., Gaut B., Geller L., Glaubitz J.C., Gore M., Guill K.E., Holland J., Hufford M.B., Lai J., Li M., Liu X., Lu Y., McCombie R., Nelson R., Poland J., Prasanna B.M., Pyhäjärvi T., Rong T., Sekhon R.S., Sun Q., Tenaillon M.I., Tian F., Wang J., Xu X., Zhang Z., Kaeppler S.M., Ross-Ibarra J., McMullen M.D., Buckler E.S., Zhang G., Xu Y. and Ware D. (2012). Maize HapMap2 identifies extant variation from a genome in flux. *Nat. Genet.* 44, 803–807.

Chippindale A.K. (2006). Experimental evolution. Evolutionary genetics: concepts and case

studies, 482-501.

Clarke W.E., Higgins E.E., Plieske J., Wieseke R., Sidebottom C., Khedikar Y., Batley J., Edwards D., Meng J., Li R., Lawley CT., Pauquet J., Laga B., Cheung W., Iniguez-Luy F., Dyrszka E., Rae S., Stich B., Snowdon R.J., Sharpe A.G., Ganal M.W., Parkin I.A.P. (2016). A high-density SNP genotyping array for *Brassica napus* and its ancestral diploid species based on optimised selection of single-locus markers in the allotetraploid genome. *Theor Appl Genet.* doi: 10.1007/s00122-016-2746-7

Clegg M.T., Allard R.W. and Kahler A.L. (1972). Is the gene the unit of selection? Evidence from two experimental plant populations. *Proceedings of the National Academy of Sciences*, 69(9), 2474-2478.

Close T.J., Bhat P.R., Lonardi S., Wu Y., Rostoks N., Ramsay L., Druka A., Stein N., Svensson J.T., Wanamaker S., Bozdag S., Roose M.L., Moscou M.J., Chao S., Varshney R.K., Szűcs P., Sato K., Hayes P.M., Matthews D.E., Kleinhofs A., Muehlbauer G.J., DeYoung J., Marshall D.F., Madishetty K., Fenton R.D., Condamine P., Graner A. and Waugh R. (2009). Development and implementation of high-throughput SNP genotyping in barley. *BMC genomics*, 10(1), 582.

Comadran J., Kilian B., Russell J., Ramsay L., Stein N., Ganal M., Shaw P., Bayer M., Thomas W., Marshall D., Hedley P., Tondelli A., Pecchioni N., Francia E., Korzun V., Walther A. and Waugh R. (2012). Natural variation in a homolog of *Antirrhinum* CENTRORADIALIS contributed to spring growth habit and environmental adaptation in cultivated barley. *Nature genetics*, 44(12), 1388.

Cook J.P., McMullen M.D., Holland J.B., Tian F., Bradbury P., Ross-Ibarra J., Buckler E.S. and Flint-Garcia S.A. (2012). Genetic architecture of maize kernel composition in the nested association mapping and inbred association panels. *Plant Physiol.* 158, 824–834.

Darwin C. (1868). The variation of animals and plants under domestication (Vol. 2). O. Judd.

Dobzhansky T. and Dobzhansky T. G. (1970). Genetics of the evolutionary process (Vol. 139). Columbia University Press.

Dolezel J., Greilhuber J., Lucretti S., Meister A., Lysak M.A., Nardi L., Obermayer R. (1998). Plant genome size estimation by flow cytometry: Inter-laboratory comparison. *Ann Bot*-London 82:17-26

Dudley J.W. and Lambert R.J. (2004). 100 generations of selection for oil and protein in corn. *Plant breeding reviews*, 24(1), 79-110.

Dwivedi S.L., Ceccarelli S., Blair M.W., Upadhyaya H.D., Are A.K., Ortiz R. (2016). Landrace Germplasm for Improving Yield and Abiotic Stress Adaptation. *Trends Plant Sci* 21:31-42

Earl D.A. and vonHoldt B.M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* vol. 4 (2) pp. 359-361 doi: 10.1007/s12686-011-9548-7

Elena S.F. and Lenski R.E. (2003). Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nature Reviews Genetics*, 4(6), 457-469.

Elshire R.J., Glaubitz J.C., Sun Q., Poland J.A., Kawamoto K., Buckler E.S. and Mitchell S.E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE*, 6, e19379.

Endresen D.T.F., Street K., Mackay M., Bari A., De Pauw E. (2011). Predictive Association between Biotic Stress Traits and Eco-Geographic Data for Wheat and Barley Landraces. *Crop Sci* 51:2036-2055

Evanno G., Regnaut S. and Goudet J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular ecology*, 14(8), 2611-2620.

Excoffier L., Smouse P.E. and Quattro J.M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, 131(2), 479-491.

Excoffier L.G.L. (2001). Analysis of population subdivision.

Fan J.-B., Oliphant A., Shen R., Kermani B., Garcia F., Gunderson K., Hansen M., Steemers F., Butler S.L., Deloukas P., Galver L., Hunt S., Mcbride C., Bibikova M., Rubano T., Chen J., Wickham E., Doucet D., Chang W., Campbell D., Zhang B., Kruglyak S., Bentley D., Haas J., Rigault P., Zhou L., Stuelpnagel J. and Chee M.S. (2003). Highly parallel SNP genotyping. Cold Spring Harbor Symp. Quant. Biol. 68: 69–78

FAO (2008). *Country Report on the State of Plant Genetic Resources for Food and Agriculture, Croatia.* Available Online at: http://www.fao.org/docrep/013/i1500e/Croatia.pdf.

Fischbeck G. (2003). in Diversity in Barley (*Hordeum vulgare*), eds von Bothmer R, van Hintum T, Knüpffer H, Sato K (Elsevier, San Diego), pp 29–52.

Flavell A.J., Dunbar E., Anderson R., Pearce S.R., Hartley R., and Kumar A. (1992). *Ty1*–copia group retrotransposons are ubiquitous and heterogeneous in higher plants. *Nucleic Acids Research*, 20(14), 3639-3644.

Flint-Garcia S.A., Thornsberry J.M. and Buckler IV E.S. (2003). Structure of linkage disequilibrium in plants. *Annual review of plant biology*, 54(1), 357-374.

Foll M. and Gaggiotti O.E. (2008). A genome scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics* 180: 977-993.

Frankel O.H., Brown A.H.D. and Burdon J.J. (1998). The Conservation of Plant Biodiversity, 2nd edn. Cambridge: Cambridge University Press, pp. 56 –78.

Franklin I.R. (1980). Evolutionary change in small populations.

Friedt W. and Ordon F. (2013). Barley Production and Breeding in Europe: Modern Cultivars Combine Disease Resistance, Malting Quality and High Yield. In: Zhang

G, Li C, Liu X (eds) Advance in Barley Sciences: Proceedings of 11th International Barley Genetics Symposium. Springer Netherlands, Dordrecht, pp 389-400

Fuller D.Q., Willcox G., Allaby R.G. (2011). Cultivation and domestication had multiple origins: arguments against the core area hypothesis for the origins of agriculture in the Near East. *World Archaeology* 43:628-652

Futuyma D.J. and Bennett A.F. (2009). The importance of experimental studies in evolutionary biology. Experimental evolution: Concepts, methods, and applications of selection experiments, 15-30.

Ganal M.W., Durstewitz G., Polley A., Bérard A., Buckler E.S., Charcosset A., Clarke J.D., Graner E.-M., Hansen M., Joets J., Le Paslier M.-C., McMullen M.D., Montalent P., Rose M., Schön C.-C., Sun Q., Walter H., Martin O.C. and Falque M. (2011). A large maize (*Zea mays* L.) SNP genotyping array: development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome. *PLoS ONE*, 6, e28334.

Ganal M.W., Polley A., Graner E.M., Plieske J., Wieseke R., Luerssen H. and Durstewitz G. (2012). Large SNP arrays for genotyping in crop plants. *Journal of biosciences*, 37(5), 821-828.

Gao Y., Liu H., An C., Shi Y., Liu X., Yuan W., Zhang B., Yang J., Yu C. and Gao H. (2013). *Arabidopsis FRS 4/CPD 25* and *FHY 3/CPD 45* work cooperatively to promote the expression of the chloroplast division gene ARC 5 and chloroplast division. *The Plant Journal*, 75(5), 795-807.

Garland Jr T. (2003). Selection experiments: an under-utilized tool in biomechanics and organismal biology. *Vertebrate biomechanics and evolution*, 23-56.

Garland T. and Kelly S.A. (2006). Phenotypic plasticity and experimental evolution. *Journal of Experimental Biology*, 209(12), 2344-2361.

Goudet J. and Jombart T. (2015). Hierfstat: estimation and tests of hierarchical F-statistics. R package version 0.04-22.

Grandbastien M.A. (1998). Activation of plant retrotransposons under stress conditions. *Trends in plant science*, 3(5), 181-187.

Graner A. and Bauer E. (1993) Rflp Mapping of the *Ym4* Virus-Resistance Gene in Barley. *Theor Appl Genet* 86:689-693

Hadado T., Rau D., Bitocchi E. and Papa R. (2009). Adaptation and diversity along an altitudinal gradient in Ethiopian barley (*Hordeum vulgare* L.) landraces revealed by molecular analysis. *BMC Plant Biology*. https://doi.org/10.1186/1471-2229-10-121

Harlan H.V. and Martini M.L. (1929). A Composite Hybrid Mixture 1. *Agronomy Journal*, 21(4), 487-490.

Harlan J.R. (1975). Our vanishing genetic resources. *Science*. 188:618-621

Harlan J.R. (1978). On the origin of barley. Pages 10 - 36 in : Barley: Origin , Botany

Culture, Winter Hardiness, Genetics, Utilization, pests. Agriculture Handbook 338. US. Department of Agriculture, Washinton, DC

Harlan J.R. (1995). Agricultural origins and crop domestication in the Mediterranean region. *Diversity* 11, 14–16. Ibrahim

Harlan J.R. and Wet J.M.J. (1971). Toward a rational classification of cultivated plants. Taxon 20:509- 517

Hasan M., Odat N., Qrunfleh I., Shakhatreh Y. and Saifan, S. (2018). Microsatellite analysis of genetic diversity and population structure of Jordanian barley (*Hordeum vulgare* l.) reveals genetic polymorphism and divergence associated with inflorescence type. *Research on Crops*, *19*(1), 86–96. https://doi.org/10.5958/2348-7542.2018.00015.3

Hawkes J.G. (1983). The Diversity of Crop Plants. Cambridge, MA: Harvard University Press. p. 102

Hirochika H. and Hirochika R. (1993). *Ty1*-copia group retrotransposons as ubiquitous components of plant genomes. *The Japanese Journal of Genetics*, 68(1), 35-46.

Hopf M. and Zohary D. (1988). Domestication of plants in the Old World.

Huang X., Ouyang X., Yang P., Lau O.S., Li G., Li J., Chen H. and Deng X.W. (2012). *Arabidopsis FHY3* and *HY5* positively mediate induction of COP1 transcription in response to photomorphogenic UV-B light. *Plant Cell* 24: 4590–4606.

Istat - Istituto Nazionale di Statistica (http://agri.istat.it)

Jia G., Huang X., Zhi H., Zhao Y., Zhao Q., Li W., Chai Y., Yang L., Liu K., Lu H., Zhu C., Lu Y., Zhou C., Fan D., Weng Q., Guo Y., Huang T., Zhang L., Lu T., Feng Q., Hao H., Liu H., Lu P., Zhang N., Li Y., Guo E., Wang S., Wang S., Liu J., Zhang W., Chen G., Zhang B., Li W., Wang Y., Li H., Zhao B., Li J., Diao X. and Han B. (2013). A haplotype map of genomic variations and genome-wide association studies of agronomic traits in foxtail millet (*Setaria italica*). *Nat. Genet.* 45, 957–961.

Jarvis D.I. and Hodgkin T. (1999). Wild relatives and crop cultivars: detecting natural introgression and farmer selection of new genetic combinations in agroecosystems. *Molecular ecology*, 8, S159-S173.

Jarvis D.I., Hodgkin T., Sthapit B.R., Fadda C. and Lopez-Noriega I. (2011). An Heuristic framework for identifying multiple ways of supporting the conservation and use of traditional crop varieties within the agricultural production system. *Critical Reviews in Plant Sciences*, *30*(1–2), 125–176. https://doi.org/10.1080/07352689.2011.554358

Jensen H.R., Dreiseitl A., Sadiki M. and Schoen D.J. (2012). The Red Queen and the seed bank: pathogen resistance of ex situ and in situ conserved barley. *Evolutionary Applications*, 5(4), 353-367.

Jombart T., Kamvar Z.N., Collins C., Lustrik R., Beugin M.P., Knaus B.J. and Jombart M.T. (2018). Package 'adegenet'.

Jones G., Charles M.P., Jones M.K., Colledge S., Leigh F.J., Lister D.A., Smith L.M.J., Powell W., Brown T.A., Jones H. (2013). DNA evidence for multiple introductions of barley into Europe following dispersed domestications in Western Asia. *Antiquity* 87:701-713

Kami J., Becerra Velasquez V., Debouck D.G. and Gepts P. (1995). Identification of presumed ancestral DNA sequences of phaseolin in *Phaseolus vulgaris*. *Proc. Natl. Acad. Sci*. U.S.A. 92, 1101–1104. doi: 10.1073/pnas.92.4.1101.

Kao K.C. and Sherlock G. (2008). Molecular characterization of clonal interference during adaptive evolution in asexual populations of *Saccharomyces cerevisiae*. *Nature genetics*, 40(12), 1499-1504.

Kasha K.J. and Kao K.N. (1970). High Frequency Haploid Production in Barley (*Hordeum vulgare* L.). *Nature* 225:874-876

Kihara H. (1983). in Proceedings of the 6th International Wheat Genetic Symposium, ed Sakamoto S (Plant Germplasm Institute, University of Kyoto, Kyoto, Japan)

Kilian B., Ozkan H., Walther A., Kohl J., Dagan T., Salamini F. and Martin W. (2007). Molecular diversity at 18 loci in 321 wild and 92 domesticate lines reveal no reduction of nucleotide diversity during *Triticum monococcum* (einkorn) domestication: Implications for the origin of agriculture. *Mol Biol Evol* 24:2657-2668

Lazaridi E., Ntatsi G., Savvas D. and Bebeli P.J. (2017). Diversity in cowpea (*Vigna unguiculata* (L.) Walp.) local populations from Greece. *Genetic Resources and Crop Evolution*, *64*(7), 1529–1551. https://doi.org/10.1007/s10722-016-0452-6.

Leinonen T., McCairns R.S., O'hara R.B. and Merilä J. (2013). Q_{ST}–F_{ST} comparisons: evolutionary and ecological insights from genomic heterogeneity. *Nature Reviews Genetics*, 14(3), 179.

Lenski R.E., Rose M.R., Simpson S.C. and Tadler S.C. (1991). Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. *The American Naturalist*, 138(6), 1315-1341.

Li G., Siddiqui H., Teng Y., Lin R., Wan X.Y., Li J., Lau O.-S., Ouyang X., Dai M., Wan J., Devlin P.F., Deng X.W. and Wang H. (2011). Coordinated transcriptional regulation underlying the circadian clock in *Arabidopsis*. *Nature cell biology*, 13(5), 616.

Lin R. and Wang H. (2004). *Arabidopsis FHY3/FAR1* gene family and distinct roles of its members in light control of *Arabidopsis* development. *Plant physiology*, *136*(4), 4010-4022.

Lioi L. (1989). Geographical variation of phaseolin patterns in an old world collection of*Phaseolus vulgaris*. *Seed Sci. Technol.* 17, 317–324.

Lipka A.E., Tian F., Wang Q., Peiffer J., Li M., Bradbury P.J., Gore M.A., Buckler E.S. and Zhang, Z. (2012). GAPIT: genome association and prediction integrated tool.

Bioinformatics, 28(18), 2397-2399.

Malysheva-Otto L.V., Ganal M.W. and Röder M.S. (2006). Analysis of molecular diversity, population structure and linkage disequilibrium in a worldwide survey of cultivated barley germplasm (*Hordeum vulgare* L.). *BMC genetics*, 7(1), 6.

Manova V., Georgieva R., Borisov B. and Stoilov L. (2016). Efficient removal of cyclobutane pyrimidine dimers in barley: differential contribution of light-dependent and dark DNA repair pathways. *Physiologia plantarum*, 158(2), 236-253.

Mazzucato A., Papa R., Bitocchi E., Mosconi P., Nanni L., Negri V., Picarella M.E., Siligato F., Soressi G.P., Tiranti B., Veronesi F. (2008). Genetic diversity, structure and marker-trait associations in a collection of Italian tomato (*Solanum lycopersicum* L.) landraces. *Theoretical and Applied Genetics*, *116*(5), 657–669. https://doi.org/10.1007/s00122-007-0699-6

Morrell P.L. and Clegg M.T. (2007). Genetic evidence for a second domestication of barley (*Hordeum vulgare*) east of the Fertile Crescent. Proceedings of the National Academy of Sciences 104:3289-3294

Mayer K.F.X., Ariyadasa R., Berges H., Bhat P., Brown J.W.S., Close T.J., Felder M., Fincher G.B., Frenkel Z., Graner A., Groth M., Hedley P., Korol A., Langridge P., Liu H., Madishetty K., Mascher M., Matsumoto T., Morgante M., Morris J., Moscou M., Muehlbauer G.J., Platzer M., Poursarebani N., Resnik J., Sampath D., Sato K., Scalabrin S., Scholz U., Schulman A., Schulte D., Shi B.J., Stein N., Steuernagel B., Svensson J.T., Swarbreck D., Taudien S., Vendramin V., Waugh R., Wise R.P., Zhou R.N., Zuccolo A. Consoritium IBGS (2012). A physical, genetic and functional sequence assembly of the barley genome. *Nature* 491:711

Merilä J. and Crnokrak P. (2001). Comparison of genetic differentiation at marker loci and quantitative traits. *Journal of Evolutionary Biology*, 14(6), 892-903.

Mason A.S., Higgins E.E., Snowdon R.J., Batley J., Stein A., Werner C. and Parkin I. (2017). A user guide to the *Brassica* 60K Illumina InfiniumTM SNP genotyping array. *Theoret. Appl. Genet.* 130, 621–633.

McClintock B. (1984). The significance of responses of the genome to challenge. *Science*, 226, 792–801.

Negri V. (2005). Agro-Biodiversity Conservation in Europe: Ethical Issues. *Journal of Agricultural and Environmental Ethics*, *18*(1), 3–25. https://doi.org/10.1007/s10806-004-3084-3.

Nei M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences*, 70(12), 3321-3323.

Nei M. (1977). F-statistics and analysis of gene diversity in subdivided populations. *Ann. Hum. Genet.* 41: 225–233.

Nei M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, 23, 341–369.

Nei M. (1987). Molecular Evolutionary Genetics. Columbia University Press.

Nevo E. and Shewry P. (1992). Origin, evolution, population genetics and resources for breeding of wild barley, *Hordeum spontaneum*, in the Fertile Crescent. Barley: genetics, biochemistry, molecular biology and biotechnology:19-43

Oliver R.E., Tinker N.A., Lazo G.R., Chao S., Jellen E.N., Carson M.L., Rines H.W., Obert D.E., Lutz J.D., Shackelford I., Korol A.B., Wight C.P., Gardner K.M., Hattori J., Beattie A.D., Bjørnstad Å., Bonman J.M., Jannink J.-L., Sorrells M.E., Brown-Guedira G.L., Mitchell Fetch J.W., Harrison S.A., Howarth C.J., Ibrahim A., Kolb F.L., McMullen M.S., Murphy J.P., Ohm H.W., Rossnagel B.G., Yan W., Miclaus K.J., Hiller J., Maughan P.J., Redman Hulse R.R., Anderson J.M., Islamovic E. and Jackson E.W. (2013). SNP discovery and chromosome anchoring provide the first physically-anchored hexaploid oat map and reveal synteny with model species. *PLoS ONE*, 8, e58068.

Orabi J., Backes G., Wolday A., Yahyaoui A., Jahoor A. (2007). The Horn of Africa as a centre of barley diversification and a potential domestication site. *Theor Appl Genet* 114:1117-1127.

Ouyang X., Li J., Li G., Li B., Chen B., Shen H., Huang X., Mo X., Wan X., Lin R., Li S., Wang H. and Deng X.W. (2011). Genome-wide binding site analysis of *FAR-RED ELONGATED HYPOCOTYL3* reveals its novel function in *Arabidopsis* development. *The Plant Cell*, tpc-111.

Papa R., Attene G. and Veronesi F. (1994). Diversity and Adaptation in Sardinian Barley (*Hordeum vulgare* L.) Landraces. In Plant Production on the Threshold of a New Century (pp. 477-479). *Springer*, Dordrecht.

Papa R., Attene G., Barcaccia G., Ohgata A. and Konishi T. (1998). Genetic diversity in landrace populations of *Hordeum vulgare* L. from Sardinia, Italy, as revealed by RAPDs, isozymes and morphophenological traits. *Plant Breeding* 117:523-530

Paradis E. (2010). *Pegas*: an R package for population genetics with an integrated–modular approach. *Bioinformatics*, 26(3), 419-420.

Parzies H.K., Spoor W. and Ennos R.A. (2000). Genetic diversity of barley landrace accessions (*Hordeum vulgare* ssp . *vulgare*) conserved for different lengths of time in ex situ gene banks, *84* (August 1999).

Pasam R.K., Sharma R., Walther A., Ozkan H., Graner A. and Kilian B. (2014). Genetic Diversity and Population Structure in a Legacy Collection of Spring Barley Landraces Adapted to a Wide Range of Climates. *Plos One* 9

Percival J. (1921). The Wheat Plant. Duckworth, London

Petersen L., Ostergard H., Giese H. (1994). Genetic Diversity among Wild and Cultivated Barley as Revealed by Rflp. *Theor Appl Genet* 89:676-681

Pickering R.A. (1984). The Influence of Genotype and Environment on Chromosome Elimination in Crosses between *Hordeum Vulgare* L. X *Hordeum Bulbosum* L. *Plant*

Sci Lett 34:153-164

Pickering R. (2000). Do the wild relatives of cultivated barley have a place in barley improvement? In: Barley Genetics VIII. Proc. 8"Int. Barley Genet., Symp., Adelaide, Australia, pp. 223- 230

Piffanelli P., Ramsay L., Waugh R., Benabdelmouna A., D'Hont A., Hollricher K., Jørgensen J.H., Schulze-Lefert P., Panstruga R. (2004). A barley cultivation-associated polymorphism conveys resistance to powdery mildew. *Nature* 430:887-891

Poets A.M., Fang Z., Clegg M.T. and Morrell P.L. (2015). Barley landraces are characterized by geographically heterogeneous genomic origins. *Genome Biol* 16:1-11

Poland J.A., Brown P.J., Sorrells M.E. and Jannink J.-L. (2012). Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS ONE*, 7, e32253.

Polegri L. and Negri V. (2010). Molecular markers for promoting agro-biodiversity conservation: a case study from Italy. How cowpea landraces were saved from extinction. *Genetic Resources and Crop Evolution*, *57*(6), 867–880. https://doi.org/10.1007/s10722-009-9526-z

Pourkheirandish M., Hensel G., Kilian B., Senthil N., Chen G., Sameri M., Azhaguvel P., Sakuma S., Dhanagond S., Sharma R., Mascher M., Himmelbach A., Gottwald S., Nair S.K., Tagiri A., Yukuhiro F., Nagamura Y., Kanamori H., Matsumoto T., Willcox G., Middleton C.P., Wicker T., Walther A., Waugh R., Fincher G.B., Stein N., Kumlehn J., Sato K., Komatsuda T. (2015). Evolution of the Grain Dispersal System in Barley. *Cell* 162:527-539

Pritchard J.K., Stephens M., Rosenberg N.A. and Donnelly P. (2000). Association mapping in structured populations. *The American Journal of Human Genetics*, 67(1), 170-181.

Pusadee T., Jamjod S., Chiang Y.-C., Rerkasem B. and Schaal B.A. (2009). Genetic structure and isolation by distance in a landrace of Thai rice. *Proceedings of the National Academy of Sciences*, *106*(33), 13880–13885. https://doi.org/10.1073/pnas.0906720106

Qi-lun Y., Ping F., Ke-cheng K. and Guang-tang P. (2008). Genetic diversity based on SSR markers in maize *(Zea mays L.)* landraces from Wuling mountain region in China. Retrieved from http://www.ias.ac.in/jgenet/.

R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Rafalski J.A. (2002). Novel genetic mapping tools in plants: SNPs and LD-based approaches. *Plant science*, 162(3), 329-333.

Rau D. (2002). Effetto dei fattori agro-climatici sulla struttura della diversità genetica del sistema pianta-patogeno *Hordeum vulgare-Pyrenophora teres*. Dottorato di Ricerca (PhD) Thesis, Università degli Studi di Sassari, Sassari, Italy.

Rellstab C., Gugerli F., Eckert A.J., Hancock A.M. and Holderegger R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, 24(17), 4348-4370.

Remy E., Cabrito T.R., Baster P., Batista R.A., Teixeira M.C., Friml J., Sá-Correia I. and Duque P. (2013). A major facilitator superfamily transporter plays a dual role in polar auxin transport and drought stress tolerance in *Arabidopsis*. *The Plant Cell*, tpc-113.

Remington D.L., Thornsberry J.M., Matsuoka Y., Wilson L.M., Whitt S.R., Doebley J., Kresovich S., Major Goodman M. and Buckler VI E.S. (2001). Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proceedings of the National Academy of Sciences*, 98(20), 11479-11484.

Reynolds M., Dreccer F. and Trethowan R. (2006). Drought-adaptive traits derived from wheat wild relatives and landraces. *Journal of Experimental Botany*, *58*(2), 177–186. https://doi.org/10.1093/jxb/erl250

Rodriguez M., Rau D., Papa R. and Attene G. (2008). Genotype by environment interactions in barley (*Hordeum vulgare* L.): different responses of landraces, recombinant inbred lines and varieties to Mediterranean environment. *Euphytica*, 163(2), 231-247.

Rodriguez M., Rau D., O'Sullivan D., Brown A.H., Papa R. and Attene G. (2012). Genetic structure and linkage disequilibrium in landrace populations of barley in Sardinia. *Theoretical and Applied Genetics*, 125(1), 171-184.

Romdhane M.B., Riahi L., Selmi A., Jardak R., Bouajila A., Ghorbel A. and Zoghlami N. (2017). Low genetic differentiation and evidence of gene flow among barley landrace populations in Tunisia. *Crop Science*, *57*(3), 1585–1593. https://doi.org/10.2135/cropsci2016.05.0298

Saintenac C., Jiang D., Wang S. and Akhunov E. (2013). Sequence-based mapping of the polyploid wheat genome. G3 (Bethesda), 3, 1105–1114.

Saisho D. and Purugganan M.D. (2007). Molecular phylogeography of domesticated barley traces expansion of agriculture in the Old World. *Genetics* 177:1765-1776.

Sarker A. and Erskine W. (2006). Recent progress in the ancient lentil. *The Journal of Agricultural Science*, 144(1), 19-29.

Sharma-Poudyal D., Chen X.M., Wan A.M., Zhan G.M., Kang Z.S., Cao S.Q., Jin S.L., Morgounov A., Akin B., Mert Z., Shah S.J.A., Bux H., Ashraf M., Sharma R.C., Madariaga R., Puri K.D., Wellings C., Xi K.Q., Wanyera R., Manninger K., Ganzález M.I., Koyda M., Sanin S., Patzek L.J. (2013). Virulence characterization of international collections of the wheat stripe rust pathogen, *Puccinia striiformis* f. sp. *tritici*. *Plant Disease* 97: 379–386.

Sicard D., Nanni L., Porfiri O., Bulfon D. and Papa R. (2005). Genetic diversity of *Phaseolus vulgaris* L. and *P. coccineus* landraces in central Italy. Retrieved from https://www.researchgate.net/publication/224946461.

Sim S.C., Durstewitz G., Plieske, J., Wieseke R., Ganal M.W., Van Deynze A., Hamilton

J.P., Buell C.R., Causse M., Wijeratne S. and Francis D.M. (2012). Development of a large SNP genotyping array and generation of high-density genetic maps in tomato. *PLoS ONE*, 7, e40563.

Slatkin M. (1987). Gene flow and the geographic structure of natural populations. *Science*, 236(4803), 787-792.

Stirnberg P., Zhao S., Williamson L., Ward S. and Leyser O. (2012). *FHY3* promotes shoot branching and stress tolerance in *Arabidopsis* in an AXR1-dependent manner. *The Plant Journal*, 71(6), 907-920.

Storz J.F. (2005). INVITED REVIEW: Using genome scans of DNA polymorphism to infer adaptive population divergence. *Molecular ecology*, 14(3), 671-688.

Subrahmanyam N. and von Bothmer R. (1987), Interspecific hybridization with *Hordeum bulbosum* and development of hybrids and haploids. *Hereditas*, 106: 119-127. doi:10.1111/j.1601-5223.1987.tb00244.x

Suneson C. A. (1956). An Evolutionary Plant Breeding Method 1. *Agronomy Journal*, 48(4), 188-191.

Sutton T., Baumann U., Hayes J., Collins N.C., Shi B.-J., Schnurbusch T., Hay A., Mayo G., Pallotta M., Tester M. (2007). Boron-toxicity tolerance in barley arising from efflux transporter amplification. *Science* 318:1446-1449

Swallow J.G. and Garland T. (2005). Selection experiments as a tool in evolutionary and omparative physiology: insights into complex traits—an introduction to the symposium. Integrative and comparative Biology, 45(3), 387-390.

Taketa S., Amano S., Tsujino Y., Sato T., Saisho D., Kakeda K., Nomura M., Suzuki T., Matsumoto T., Sato K., Kanamori H., Kawasaki S. and Takeda K. (2008). Barley grain with adhering hulls is controlled by an *ERF* family transcription factor gene regulating a lipid biosynthesis pathway. *P Natl Acad Sci* USA 105:4062-4067.

Tang W., Wang W., Chen D., Ji Q., Jing Y., Wang H. and Lin R. (2012). Transposase-derived proteins *FHY3/FAR1* interact with PHYTOCHROME-INTERACTING FACTOR1 to regulate chlorophyll biosynthesis by modulating *HEMB1* during deetiolation in *Arabidopsis*. *The Plant Cell*, tpc-112.

Tannahill R. (1988). Food in History, rev. ed. Penguin, London

Tanno K., Willcox G. (2006) How fast was wild wheat domesticated? *Science* 311:1886-1886

Tenaillon O., Rodríguez-Verdugo A., Gaut R.L., McDonald P., Bennett A.F., Long A.D. and Gaut B.S. (2012). The molecular diversity of adaptive convergence. *Science*, 335(6067), 457-461.

Tian F., Bradbury P.J., Brown P.J., Hung H., Sun Q., Flint-Garcia S., Rocheford T.R., McMullen M.D., Holland J.B. and Buckler VI E.S. (2011). Genome-wide association study of leaf architecture in the maize nested association mapping population. *Nat. Genet.* 43, 159–162.

Tiranti B. and Negri V. (2007). Selective microenvironmental effects play a role in shaping genetic diversity and structure in a *Phaseolus vulgaris* L. landrace: implications for on-farm conservation. *Molecular Ecology*, *16*(23), 4942–4955. https://doi.org/10.1111/j.1365-294X.2007.03566.x

Travisano M. (2009). Long-term experimental evolution and adaptive radiation. Experimental evolution: concepts, methods, and applications of selection experiments (eds Garland T. Jr., Rose MR), 111-133.

Ullstrup A.J. (1972). The impact of the southern corn leaf blight epidemics of 1970–1971. *Annu Rev Phytopathol* 10:37–50

Van Hung P. (2016). Phenolic compounds of cereals and their antioxidant capacity. *Crit Rev Food Sci* 56:25-35

Van Poecke R.M.P., Maccaferri M., Tang J., Truong H.T., Janssen A., van Orsouw N.J., Salvi S., Sanguineti M.C., Tuberosa R. and van der Vossen E.A.G. (2013). Sequence-based SNP genotyping in durum wheat. *Plant Biotechnol. J.* 11, 809–817.

Vasemägi A., Nilsson J. and Primmer C.R. (2005). Expressed sequence tag-linked microsatellites as a source of gene-associated polymorphisms for detecting signatures of divergent selection in Atlantic salmon (*Salmo salar* L.). *Molecular Biology and Evolution*, 22(4), 1067-1076.

Vigouroux Y., McMullen M., Hittinger C.T., Houchins K., Schulz L., Kresovich S., Matsuoka Y. and Doebley J. (2002). Identifying genes of agronomic importance in maize by screening microsatellites for evidence of selection during domestication. *Proceedings of the National Academy of Sciences*, 99(15), 9650-9655.

Villa T.C.C., Maxted N., Scholten M. and Ford-Lloyd B. (2005). Defining and identifying crop landraces. *Plant Genetic Resources*, 3(3), 373-384.

Voytas D.F., Cummings M.P., Koniczny A., Ausubel F.M. and Rodermel S.R. (1992). Copia-like retrotransposons are ubiquitous among plants. *Proceedings of the National Academy of Sciences*, 89(15), 7124-7128.

von Bothmer R. (1992). The wild species of Hordeum: relationships and potential use for improvement of cultivated barley.

von Bothmer R, Jacobsen N., Baden C., Jørgensen C., Linde-Laursen I. (1995). An ecogeographical study of the genus *Hordeum*. Systematic and ecogeographic studies on crop genepools, 7. IPGRI, Rome, 2nd ed 129p

von Bothmer R., Sato K., Komatsuda T., Yasuda S. and Fischbeck G. (2003). The domestication of cultivated barley. Diversity in barley (*Hordeum vulgare*), 9-27.

von Rünker K. (1908). Die systematische Einteilung und Benen-ung der Getreidesorten für praktische Zwecke. Jahrbuch der Deutschen landwirtschafts-Gesellschaft 23: 137–167.

Vos P.G., Uitdewilligen J.G., Voorrips R.E., Visser R.G. and van Eck H.J. (2015). Development and analysis of a 20K SNP array for potato (*Solanum tuberosum*): an

insight into the breeding history. *Theoretical and Applied Genetics*, 128(12), 2387-2401.

Weir B.S., Allard R.W. and Kahler A.L. (1972). Analysis of complex allozyme polymorphisms in a barley population. *Genetics*, 72(3), 505-523.

Weir B.S. and Cockerham C.C. (1984). Estimating F-statistics for the analysis of population structure. evolution, 38(6), 1358-1370.

Weiss E., Kislev M.E. and Hartmann A. (2006). Autonomous cultivation before domestication. Science 312:1608-1610

Wicker T., Narechania A., Sabot F., Stein J., Vu G.T., Graner A., Ware D., Stein N. (2008). Low-pass shotgun sequencing of the barley genome facilitates rapid identification of genes, conserved non-coding sequences and novel repeats. Bmc Genomics 9

Willcox G., Fornite S. and Herveux L. (2008). Early Holocene cultivation before domestication in northern Syria. *Veg Hist Archaeobot* 17:313-325.

Willing E.M., Dreyer C. and Van Oosterhout C. (2012). Estimates of genetic differentiation measured by FST do not necessarily require large sample sizes when using many SNP markers. *PloS one,* 7(8), e42649.

Winfield M.O., Allen A.M., Burridge A.J., Barker G.L., Benbow H.R., Wilkinson P.A., Coghill J., Waterfall C., Davassi A., Scopes G., Pirani A., Webster T., Brew F., Bloor C., King J., West C., Griffiths S., King I., Bentley A.R. and Edwards K.J. (2016). High-density SNP genotyping array for hexaploid wheat and its secondary and tertiary gene pool. *Plant Biotechnol. J.* 14, 1195–1206.

Wright S. (1946). Isolation by distance under diverse systems of mating. *Genetics*, 31(1), 39.

Wright S. (1949). The genetical structure of populations. *Ann Hum Genet*. 15:323–354.

Wright S. (1978). Genetic variability in natural populations: methods. Evolution and the Genetics of Populations, 4.

Xu X., Liu X., Ge S., Jensen J.D., Hu F., Li X., Dong Y., Gutenkunst R.N., Fang L., Huang L., Li J., He W., Zhang G., Zheng X., Zhang F., Li Y., Yu C., Kristiansen K., Zhang X., Wang J., Wright M., McCouch S., Nielsen R., Wang J. and Wang W. (2012). Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. *Nat. Biotechnol.* 30, 105–111.

Yu S.Y., Long H., Deng G.B., Pan Z.F., Liang J.J., Zeng X.Q., Tang Y.W., Tashi N., Yu M.Q. (2016). A Single Nucleotide Polymorphism of Nud Converts the Caryopsis Type of Barley (*Hordeum vulgare* L.). *Plant Molecular Biology Reporter* 34:242-248

Zhao K., Tung C.W., Eizenga G.C., Wright M.H., Ali M.L., Price A.H., Norton G.J., Islam M.R., Reynolds A., Mezey J., McClung A.M., Bustamante C.D. and McCouch S.R. (2011). Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nat. Commun.* 2, 467.

Zhang Z., Ersoz E., Lai C.Q., Todhunter R.J., Tiwari H.K., Gore M.A., Bradbury P.J., Yu J., Arnett D.K, Ordovas J.M. and Buckler, E. S. (2010). Mixed linear model approach adapted for genome-wide association studies. *Nature genetics*, 42(4), 355.

Zeven A.C. (1998). Landraces: A review of definitions and classifications. *Euphytica*. 104:127-139

Zohary D. (1988). Domestication of Plants in the Old World: The Origin and Spread of Cultivated Plants in West Asia, Europe, and the Nile Valley 3rd Edition