IntechOpen

Rediscovery of Landraces as a Resource for the Future

Edited by Oscar Grillo

REDISCOVERY OF LANDRACES AS A RESOURCE FOR THE FUTURE

Edited by **Oscar Grillo**

Rediscovery of Landraces as a Resource for the Future
http://dx.doi.org/10.5772/intechopen.69576
Edited by Oscar Grillo

Contributors

Abe Gerrano, Nontuthuko Rosemary Ntuli, Froylan Rincon, Norma Ruiz, Oscar Grillo, Marisol Lo Bianco, Sebastiano Blangiforti, Laura Siracusa, Gianfranco Venora, Giuseppe Ruberto, Gianni Barcaccia, Fabio Palumbo, Andrés Cortés, Matthew W. Blair, Amjad Husaini, Najeeb Sofi, Enéas Ricardo Konzen, Siu Mui Tsai, Muhammad Amjad Nawaz, Gyuhwa Chung, Seung Hwan Yang, Musibau Azeez, Amos Adubi, Felicia Durodola

Notice

First published in London, United Kingdom, 2018 by IntechOpen
eBook (PDF) Published by IntechOpen, 2019
IntechOpen is the global imprint of INTECHOPEN LIMITED, registered in England and Wales, registration number: 11086078, The Shard, 25th floor, 32 London Bridge Street
London, SE19SG – United Kingdom
Printed in Croatia

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

Additional hard and PDF copies can be obtained from orders@intechopen.com

Rediscovery of Landraces as a Resource for the Future
Edited by Oscar Grillo

p. cm.

Print ISBN 978-1-78923-724-5

Online ISBN 978-1-78923-725-2

eBook (PDF) ISBN 978-1-83881-396-3

Meet the editor

Oscar Grillo was born in Catania (Sicily) in 1977. He is a food technologist with an international PhD degree in Applied and Environmental Botany. Since 2003, he has been working as a researcher at the *Stazione Sperimentale di Granicoltura per la Sicilia*, a government institute of agronomic research, mainly working with computer vision applied to food matrices and plant structures, primarily seeds, and in particular studying wheat and the related leguminous plants. For many years, he collaborated with the *Sardinian Germplasm Bank* of the *Biodiversity Conservation Centre* of the University of Cagliari on projects devoted to seed characterization and identification by image analysis. He is also working as a supervisor for many MSc and PhD degree students, making their own contributions to the agronomical and botanical research. Results of his work have been published in many peer-reviewed journals and international conference papers. A referee for a few peer-reviewed journals, many times, he has been invited as a visiting professor by national and international universities and research centers.

Contents

Preface

A cropped variety or cultivar is an intraspecific taxonomic entity characterized by a high level of homozygosis, especially for the genes that control the selected traits. Consequently, the individuals belonging to the same variety show homogeneous morphological and/or productive traits. Nevertheless, some differences in genetically controlled biochemical traits may exist within the same variety. These variations are defined as "biotypes." Different from varieties, landraces are natural populations put in cultivation, and as such, they are characterized by a wide adaptability to various environments. Considering all the abiotic factors, the high probability of interpopulation crosses and their heterozygosis condition, from the genetic point of view, these populations result to be more than a mixture of different pure lines.

On the other hand, a landrace is intrinsically different, from the genetic point of view, and recognizable as a distinct entity, allowing to distinguish one landrace from another or from modern cultivars for the same crop. This differentiation is also on the basis of the great assortment of landrace names, linked to their origin, or specific morphological, chromatic and productive traits, or particular biological cycles, or many other reasons.

Landraces are generally less productive than commercial cultivars, but many authors agree they have played a fundamental role in the history of crops worldwide, in crop improvement and in agricultural production, and they have been in existence since the origins of agriculture itself.

In recent years, all over the world, the attention paid to local and traditional productions is growing, especially in the agro-food sector. Maybe, it is not only due to the impact of globalization and the social and economic changes but also due to the increased consideration to health and nutritional aspects of food. Hence, for economic, social, historical and nutritional reasons, this trend has led to the rediscovery and reuse of landraces of many different crops, responding to requests for more and more demanding market.

This volume collects real examples of local crops and old landraces of different areas of the planet that testify the extreme importance of the relation existing among a land, the local productions, the historical traditions, the conservation of biodiversity, the health benefits, the environmental impact and the local economies, also including the significance to dedicate resources to scientific researches in local crops.

Oscar Grillo
Experimental Institute for Wheat, Italy

Section 1

Rediscovery of Landraces as a Resource for the Future

Chapter 1

Landraces and Crop Genetic Improvement

Musibau A. Azeez, Amos O. Adubi and
Felicia A. Durodola

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.75944

Abstract

Landraces are repository of gene pool that enrich biodiversity and maintain and stabilize ecosystem in a sustainable way to make it functional. Cultivation of traditional crops in different regions of the world, aside maintaining biodiversity in agriculture, also avails humanity of regulatory services such as nutrient cycling, carbon sequestration, control of soil erosion, reduction of greenhouse gas emission and control of hydrological processes. However, man through over-exploitation of some plant species with utter neglect to some other either deliberately or otherwise through modern agricultural systems that promote cultivation of a few high-input and high-yielding crop species caused disaffection to biodiversity with consequences of reduction in its regulatory services. In this chapter, different landraces of crops are examined, their usefulness in the maintenance of genetic diversity is explored, and implications of their depletion are discussed.

Keywords: genetic diversity, adaptation, conservation, heterogeneity, utilization

1. Introduction

Landraces are defined as dynamic populations of a cultivated plant with a 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 [1] . Sangam et al. [2] referred to plant landraces as heterogeneous local adaptations of domesticated species providing genetic resources that meet current and new challenges for farming in stressful environments. These local ecotypes can show variable phenology and low to moderate edible yield but are often highly nutritious. The main contributions of landraces to plant breeding have been traits for more efficient nutrient uptake and utilization, as well as useful genes

for adaptation to stressful environments such as water stress, salinity and high temperatures. A systematic landrace evaluation may define patterns of diversity, which will facilitate identifying alleles for enhancing yield and abiotic stress adaptation, thus raising the productivity and stability of staple crops in vulnerable environments. It can also 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 [3]. A landrace differs from a variety that has been selectively improved by breeders for particular characteristics.

Landraces are important genotypes for crop breeding owing to their high potential to adapt to specific environmental conditions and the large source of genetic variability that they provide [4]. Landraces are generally less productive than commercial cultivars, although in recent years, they have become important as sources of genetic variability in the search for genes for tolerance or resistance to biotic and abiotic factors of interest in agriculture [5]. The genetic diversity observed across landraces is the most important part of maize biodiversity, and local races represent an important fraction of the genetic variability exhibited by this genus. However, few agronomic and genetic data exist for such collections, and this scarcity has limited the use, management and conservation of this germplasm. In addition, a few improved genotypes with narrower genetic variability are quickly replacing maize landraces [6].

Zeven opined that landraces have played a fundamental role in the history of crops worldwide, in crop improvement and agricultural production, and they have been in existence since the origins of agriculture itself. During this time they have been subject to genetic modification through abiotic, biotic and human interactions. For centuries, crop landraces were the principal focus for agricultural production [7]. 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 [8]. This cycle remained current until the dawn of formal plant breeding and the generation of generally higher-yielding cultivars that subsequently replaced many traditional landraces [7, 9, 10].

2. Historical background of landrace origin

The origin of landraces encompasses both the temporal and spatial components of where landraces were first developed. They (landraces) have a relatively long history, significantly more than the ephemeral lifespan of modern cultivars. Many authors suggest that landraces have been growing ‘since time immemorial’ [11], ‘over long periods of time’ [9], ‘over hundreds even a thousand years’ [12], ‘for many years even centuries’ [13], ‘for generations’ [14], ‘for many centuries’ and ‘over a period of time’ [15]. Nevertheless, few are explicit about the amount of time a landrace must be grown to be considered a landrace. However, Louette [16] indicated for maize that the period of time must be ‘for at least one farmer generation (i.e. more than 30 years)’, while Astley referred to vegetable landraces being grown for ‘50–70 or even 100 years’.

Hawkes [17] opined that landraces are associated with one specific geographical location, in contrast to cultivars which are bred remotely, trialed in several locations and subsequently cultivated in diverse locations. Therefore, landraces are closely associated with 'specific locations' and often will take the name of the location [11]. Examples of this are Kent Wild White Clover from the UK county of Kent and Tuxpenõ maize from the Tuxpan region in Mexico. However, migrations (seed flow) of established landraces from their region of origin to new regions have also occurred as local informal variety introductions. Zeven [3] proposed two types of landraces: **autochthonous** (landraces cultivated for more than a century in a specific region) and **allochthonous** (a landrace that is autochthonous in one region introduced into another region and becoming locally adapted). In that case, the examples of Kent Wild White Clover and Tuxpenõ maize are cultivated in regions other than where they originated. Kent Wild White Clover is grown in some hilly areas of Scotland and Tuxpenõ maize in several regions of Southern Mexico. A third type known as a 'Creole' landrace may be derived from an originally bred variety [18, 19], which then becomes an effective landrace following numerous repeated cycles of planting and farmer seed selection in a specific location. For instance, Square Head Master Wheat, identified as a cultivar in the National List of the UK, has been grown continuously since 1930 by the family of Paul Watkin (a farmer from Suffolk, UK) saving seed each year.

Continuity and individual cultivation and discontinuity and collective cultivation are both significant. Individual farmers commonly lose and recover landraces as a result of their management of a dynamic portfolio of landraces [19] and seed replacement [20] and because of various stochastic events such as drought, floods, pests and diseases. Village or local community continuity may be maintained through farmer's seed exchange networks if cultivation is by more than one farmer. In fact, several papers have highlighted the relevance of seed exchange for the maintenance of landraces [20–22]. Such localized farmer exchange activities may help to define and ensure continuity of a landrace. However, the introduction of 'exotic' landraces to a locality is likely to adulterate the uniqueness and local adaptation of the local landraces. Therefore, many believe that the maintenance of an 'open' cultivation system, with routine local or more remote introductions of germplasm, is likely to be responsible for the maintenance of genetic diversity in landraces.

3. Lack of formal genetic improvement

Landrace production is associated with 'no human selection' [11] and 'it was naturally developed' [23]; thus, landraces have been developed as a result of time and natural selection alone. Other authors suggest that human selection has occurred but in the form of unconscious selection, and others suggest that a certain degree of consciousness is involved in the selection process, 'without or with only little mass selection' [23], 'subject to some deliberate selection' [24], 'artificial selection (probably largely of an unconscious nature)' [17] and 'breeding or selection … either deliberately or not' [14]. Where conscious human selection has been recognized as being significant in landrace development, it has nevertheless been distinguished that is applied to modern cultivars [7, 12] with qualifications such as 'more resistant to pests and diseases, have more yield stability' [25],

'grown in traditional farming systems' [7, 13], 'cultivated in low-input cultivation' [8], 'in a number of traits which together appear to form an adaptive complex' [3] and 'on a low selection pressure'.

It is generally accepted that farmers, gardeners and growers select and develop landraces [12–15, 17, 26], while formal plant breeders select and develop cultivars (**Figure 1**). However, even this division is not as clear as it first may appear if other considerations are included. Zeven [27] explained that 'continuous selection by some farmers for plants with desired characters is similar to the later proposed scientific selection within landraces to select by seekers for the best plants'. Examples of these are shown in vegetables that present special traits such as enormous size, developed by growers in the UK.

The situation concerning the involvement of landraces in participatory plant breeding is interesting, as Maxted [28] noted that care should be taken to ensure the security of the locally adapted genetic diversity or the former landrace could no longer be regarded as a landrace. Here, the decision over whether the former landrace may still be regarded as a landrace as described by Almekinders and Elings [29] depends on the degree of breeding and the quantity of external germplasm introgressed with the original landrace; the more of either the less the entity could be regarded as a landrace. Certainly, this would be the case for participatory varietal selection programs where external germplasm is introduced into an area and suitable material is selected by local farmers; even if the new germplasm is managed by the farmer in a manner usually associated with traditional farming and landrace maintenance, the use of the term landrace would be inappropriate. Yet, another consideration is understood by the term 'modern' crop improvement.

Simmonds [30] and Allard [31] further explained that modern professional crop improvement is based on the Darwinian theory of evolution through selection and the genetic mechanisms of evolution developed by Mendel, Johannsen, Nilsson-Ehle, East and others. Frankel and Bennett [9] used as a reference point the 19th century when conscious, individual plant selection commenced. Jarman and Leggett considered that 'modern' crop improvement started when formal breeding programs were initiated, in the UK, for example, in the 1920s. However, the fact that the history of crop improvement is different for each crop is also an important element to be considered [3]. Combining these considerations, formal crop improvement is understood as the application of genetic principles and practices to the development of cultivars by both classic breeding techniques (selection and hybridization) as well as more recent technologies (biotechnology, molecular biology, transgenics) within a crop improvement program. Virchow [32] when defining the characteristics of a landrace included the fact that landraces are not registered in official seed lists, but in the UK, several entities generally regarded as landraces, such as Kent Wild White Clover, are included on the National List and are regarded as landraces because they result from farmers' selection over millennia.

In fact, it is argued that inclusion of landraces on the UK National List is likely to promote their cultivation and thus conservation [33]. Landraces may therefore be more easily defined as being crop varieties which do not result in the first instance at least from formal crop

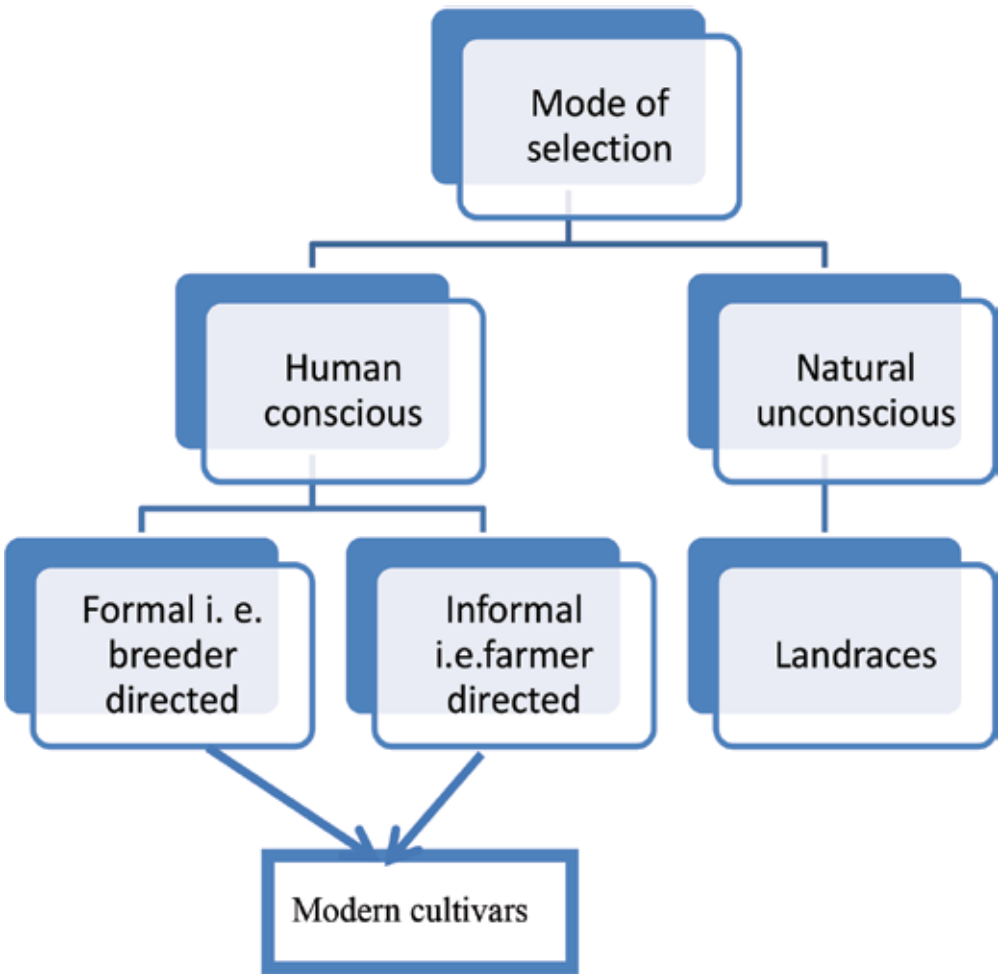

Figure 1. Different opinions about the types of landrace selection [36].

improvement programs, in contrast with modern cultivars which have resulted directly from these programs (**Figure 1**).

Despite this improved clarification, there remains confusion as regards the effect of crop evolution on landraces. Crop evolution is not a linear process, and there are different points of view of the position occupied by landraces in relation to their wild relatives, on the one hand, and cultivars, on the other. Some authors such as Marchenay [34] suggested that some landraces exist on the borders of cultivation, not having been fully domesticated and might be better considered as ecotypes. Other authors raised the issue that some landraces have crossed freely with their wild relatives over millennia [24, 35] and as a result possess rudimentary characters or 'wild relative traits' not found in cultivars because of their more ephemeral existence. While

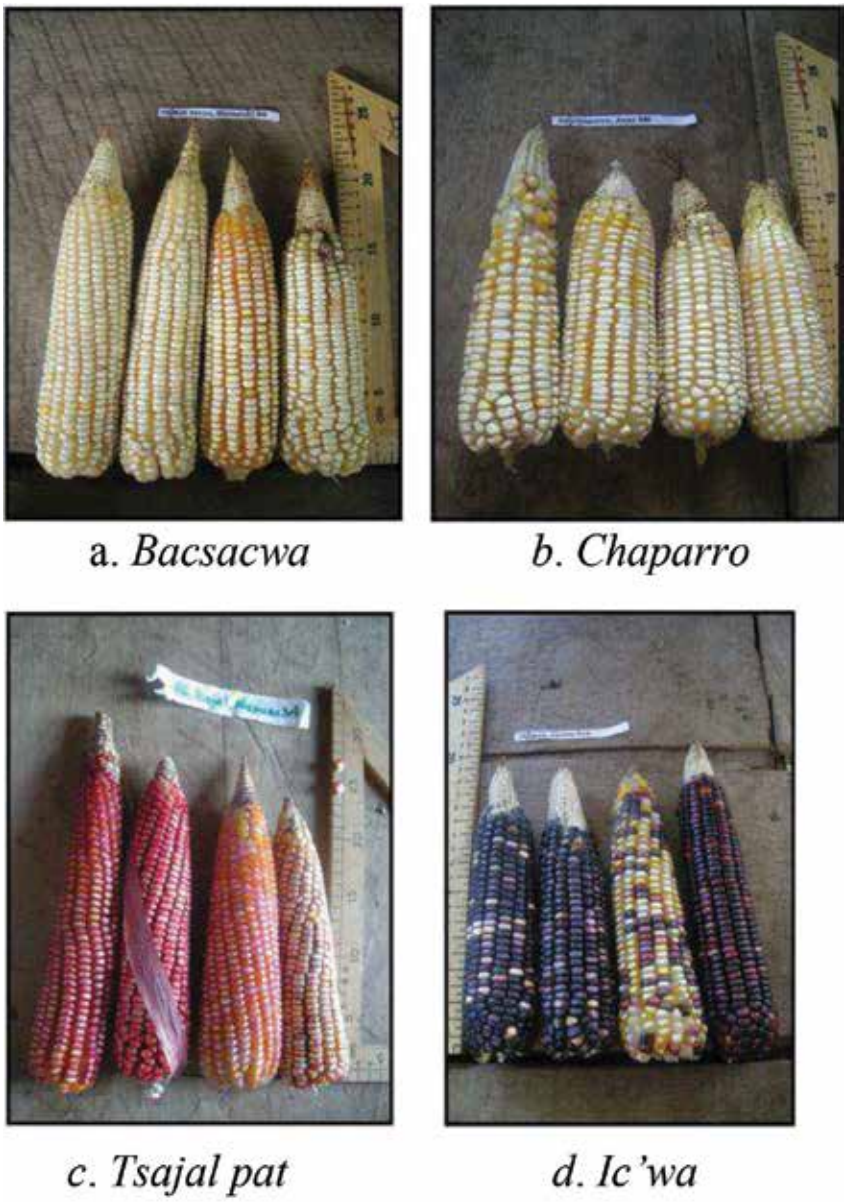

Figure 2. Types of landraces of maize [27].

others believe that landraces can even be selected from cultivars [18, 19], terms such as creolization or rustication are applied, and 'in the absence of traditional and formal maintenance breeding, any improved landrace (cultivar), including a hybrid variety, will regress with time into a landrace' [27]; 'a cultivar that has been growing under a low selection pressure for specific traits but not uniformity for a long time could be considered a landrace' (**Figure 2**).

4. Recognizable identity

Landrace must be intrinsically highly genetically diverse and recognized as a distinct entity via common-shared traits. These traits will allow the distinction of one landrace from another or from modern cultivars for the same crop. They will sometimes give rise to landrace names, but at other times, names may be determined by other factors such as use or origin. Therefore, landraces 'are each identifiable and usually have local names' [7], 'are recognized morphologically' [14], 'have a local name,' 'are a farmer selection based on local characteristics (specific use, local market, horticultural practices and locally adapted)' and 'are heterogeneous populations with a similar trait'.

However, this characteristic may be difficult to be applied universally as landraces identified on the basis of common names can be misleading because of non-associated synonyms and homonyms. Many disparate landraces may be named after their early flowering capability or seed color, for example. A landrace may be recognized by different names in different countries or communities [36], or conversely quite different landraces can be designated with the same name [14]. These factors contribute to one of the main problems associated with landraces, namely, their consistent identification and the determination of which traits can be consistently used to define the identity of a specific landrace.

5. High genetic diversity

The characteristics of landraces in relation to the magnitude of allelic and genetic diversity in contrast to cultivars are considered to be significantly more genetically diverse [37]. Thus, a landrace is a 'highly variable population in appearance' [7], 'highly diverse populations and mixtures of genotypes' [38], 'genetically heterogeneous' [13], 'not genetically uniform and containing high levels of diversity' [14], 'local diverse crop varieties' [26], 'heterogeneous crop populations' [39] and 'materials with variable levels of heterogeneity'. Frankel and Soule [40] indicated that the genetic diversity of landraces has two dimensions: between sites/populations and within sites/populations. The former is generated by heterogeneity in space and reproductive isolation, while the latter is generated by heterogeneity in time associated with both short-term variations between seasons and by longer-term climatic, biological and socio-economic changes.

Some authors have used the term 'meta-population' when referring to the diversity structure of a landrace. As such, a landrace constitutes a group of farmers' seed lots that are highly diverse both between and within themselves. In contrast however, Sanchez [41] when evaluating the genetic diversity of maize landraces of Mexico found that some landraces had very low levels of genetic diversity, and it was suggested that comparatively low diversity may be more associated with selfing crops. Bere barley, one of the oldest cereal varieties in Europe, is 'surprisingly homozygous', possibly because it has been maintained in isolation in marginal lands since the sixteenth century [42]. A similar picture is provided by Tibetan barley landraces

which proved to be much less diverse than modern barley cultivars due possibly to their relative geographic isolation, their relatively recent introduction to Tibet and the fact that they have been subject to very little natural or man-made selection [43]. Therefore, the dynamics of genetic diversity and changes over time of the genetic structure of landraces are likely to be crop specific. It is also likely to be associated with the mode of fertilization (self- versus cross) and propagation (sexual or asexual), which has over time resulted in genetic bottlenecks, varying outcrossing rates, recombination and gene flow. Thus, as Almekinders and Louwaars [24] conclude, 'a landrace is usually a complex heterogeneous population, but not necessarily so'.

6. Local genetic adaptation

Landraces are generally adapted to local environment. With the continued cycles of local planting, harvesting and farmers' selection, over time landraces will be selected for local environmental and agroecosystem conditions and practices, just as ecotypes of wild species are adapted to the local environmental conditions. Landraces 'are adapted to their growing conditions' [11]; 'possess adaptive complexes associated with the special conditions of cultivation, pure-stand associations, harvesting and others factors' [44]; 'are not only adapted to their environment, both natural and man-made, but they are also adapted to each other' [7]; 'are adapted to the areas in which they grow' [12]; 'are specifically adapted to local conditions' [13]; and 'are adapted to local conditions' [26].

Bennett [44] made the assumption that landraces are more suited to cultivation in particular locations than highly bred cultivars that are bred for cultivation in the most common environmental conditions. Inevitably, cultivars will be less suited to grow in suboptimal conditions and therefore have less of a competitive advantage in marginal environments where the local landraces are likely to have an adaptive advantage. These local conditions may be defined as abiotic (e.g. salinity, drought, etc.), biotic (e.g. pests, diseases, weeds) and human (e.g. cultivation, management and use). Landraces are perceived to have the ability 'to sensitively respond to even minor environmental influences' [44]; 'to have some built-in insurance against hazards' possibly due to their inherent population structure [7]; 'to accumulate resistance genes to limiting factors in the physical and biological environment—drought, cold, diseases, pests' [24]; and 'to be capable of producing in any but disaster seasons at a level which safeguards the survival of the cultivator' and so provide yield stability [24].

Several studies have demonstrated the relationship between landraces and local adaptation, for example; Frankel [8] and Brown [39] discuss landrace adaptation to marginal conditions associated with climatic, soil and disease stress. The evolution of local adaptation over millennia in these stressed environments ensures yield stability even in extremely adverse years. In this sense, Zeven [27] considers yield stability to be a principal characteristic of landraces.

However, even though there are numerous references to a specific relation between a landrace and local environmental conditions, there are exceptions. Zeven [3] indicated that 'some landraces are able to adapt themselves to a wide range of environments, whereas others are able to

adapt themselves only to a few environments'. Wood and Lenne [19] disagree with the assumption 'that all traditional varieties are locally adapted' and state that 'evidence against specific local adaptation in crop varieties is provided by the extensive interchange of traditional varieties of all crops'. Farmers employing an 'open' cultivation system where there is regular local or more exotic landrace introduction are less likely to have locally adapted landraces. Zeven [20] provided evidence of farmers' traditional practice of periodic seed replacement to combat so-called degradation, which indicates that in certain situations a 'closed' cultivation system that results in local adaptation of landraces may be deleterious. The farmer's criteria for seed selection also do not necessarily lead to selection for local adaptation; the varying environmental conditions under which traditional agriculture is carried out may in certain conditions not actually favor specific local adaptation. In this sense, some authors consider that local adaptation can comprise both wide adaptation in certain landrace characters and narrow adaptation in others.

7. Association with traditional farming systems

Traditional farming systems have often been considered beneficial reservoirs of landraces and intra-crop diversity [45]. Traditional farming systems involve traditional cultivation, storage and use practices, and integrated with these practical skills, traditional knowledge about landrace identification, cultivation, storage and uses is incorporated. In this sense, one important element of landraces conservation that has recently been the focus of researchers' attention is the way that landraces studies have focused on farmers' variety selection [46], farmers' seed exchange [22], farmers' seed networks [22], farmers' seed replacement [20], farmers' portfolios of varieties [19], farmers' landraces identification [47]and farmers' landrace uses [48]. Each has shown the role of farmers for the creation and maintenance of a landrace.

In fact, Zeven [27] suggested that landrace diversity can be explained by the combination of farmers' selection criteria on specific local landrace genotypes by means of farmers' seed saving and the introduction of variation by means of exchange with other farmers of other genotypes of the same crop. This indicates that landraces are more inherently dynamic than cultivars as they are maintained through repeated cycles of sowing, harvesting and replacing seed selection by farmers [49, 50] within complex informal systems. However, it is also important to consider that traditional farming systems are themselves also dynamic and that the frontier between them and other farming systems is not well defined. As such, traditional farming systems are subject to change, incorporating in some cases modern cultivars into their systems, growing them alongside landraces of the same species [51]. These have been managed and maintained by farmers.

8. Threat to landrace diversity

The current industrial agriculture system may be the single most important threat to biodiversity [2]. Also, Sarker and Erskine [52] opined that 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. Modern agriculture has contributed to decreasing agricultural biodiversity as most of humankind lives now on only crops, with wheat (*Triticum aestivum*) L.), rice (*Oryza sativa* L.), maize (*Zea mays* L.) and potato (*Solanum tuberosum* L.) accounting for 60% of diets [2]. For example, 74% of rice cultivars in Indonesia are derived from the same stock, while 50% of wheat, 75% of potato and 50% of soybeans in the USA. The genetic erosion was estimated at 72.4 and 72.8%, respectively [53]. Furthermore, the number of rice cultivars declined in India farms from about 400,000 before colonialism to 30,000 in the mid-nineteenth century with unknown thousands more being lost after the Green Revolution. Greece also lost 95% of its wheat landraces after being encouraged to replace them with modern cultivars.

9. The concept of local variety

The concepts of the local variety have already existed in the guidelines for the proper in situ, on-farm and ex situ conservation of plant varieties. A local variety is a variety or local crop that reproduces by seed or by vegetative process. It is a variable population, which is identifiable and usually has a local name. It lacks 'formal' genetic improvement and is characterized by specific adaptation to the environmental conditions of the area of cultivation (tolerant to the biotic and the abiotic stresses of that area) and is closely associated with the traditional use, knowledge, habits, dialects and celebrations of the people who developed and continue to grow it [54].

10. On-farm management of local seed diversity

In Nigeria, in spite of the event of the formal certified seed sector, many rural farmers continue to use traditional seeds or other planting materials to meet their seed need [55]. They have their own method of selection and conservation of seeds. This method varies slightly from one crop to another. Indeed, seeds are collected at maturity on apparent healthy plants and saved from season to season by individual farmers. As with selection, storage and conservation methods varied with crops. As such, seeds were stored either in packages and suspended at kitchen roofs (in the case of maize and cowpea, for example, in Yorubaland in Nigeria) or in grain and bottled (case of peppers, tomatoes, etc.). Yet, there can also be significant amounts of exchange between neighbors and relatives. They are also purchased when necessary. On-farm management of local seed diversity is predominant in the Nigeria seed sector since conversely to the cotton culture; no organized provision system exists for food crops.

11. Durum wheat

Durum wheat (*Triticum turgidum* var. L. *durum*) is grown on over 1 million ha. Forty-five percentage of which are sown in the arid and semiarid regions, 11% in high altitudes and 44% in more favorable areas [56]. The complexity of the population structure of wheat landraces

may arise from a number of different homozygotes and the occurrence and frequency of heterozygotes in populations. The assessment of genetic diversity between and within wheat landraces is essential to utilize landraces as donors of traits in wheat breeding and to identify priority areas for on-farm conservation.

Landraces could act as donors of important characteristics, such as drought and cold tolerance and mainly grain quality. In general, they represent significantly broader genetic diversity than modern varieties, and, therefore, they could contribute to extend the genetic base of modern cultivars. The identification of quality parameters such as protein content, gluten strength, yellow pigment and their integration in the improved varieties is a priority in research on durum wheat [57]. Mineral content in modern wheat cultivars has significantly decreased, including copper, iron, magnesium, manganese, phosphorus, selenium and zinc. High levels of these nutrients can be found in landraces and old low-yielding varieties [56].

Landraces displayed a wide range of genetic diversities. This local germplasm forms an interesting source of favorable quality traits such as protein content, gluten strength and yellow pigment content useful to durum wheat breeders. The persistent cultivation of durum wheat landraces in some regions attests to their continued value to farmers and to their competitive agronomic or nutritional advantage relative to modern varieties. Adding value of these landrace is the main motivating factor for their on-farm conservation. Fungi seed treatment against seed-borne diseases and chemical weeding at the right time could improve the landrace productivity in a simple way.

Furthermore, composite landraces made up of promising lines selected from landraces could be another way for durum wheat landrace valorization. But, on-farm conservation of durum wheat genetic resources in Morocco could be more efficient provided that legislation changes are made that make it possible to market landraces as diversified genetic materials and encourage their consumption [56]. Durum wheat landraces have over many generations become adapted to the local environment and cultural conditions under which they are grown. Development of new varieties from landraces could be a viable strategy to improve yield and yield stability, especially under stress and future climate change conditions.

12. Rice

Rice is among the most important crops worldwide. While much of the world's rice harvest is based on modern high-yield varieties, traditional varieties of rice grown by indigenous groups have a great importance as a resource for future crop improvement. These local landraces represent an intermediate stage of domestication between a wild ancestor and modern varieties, and they serve as reservoirs of genetic variation. Such genetic variation is influenced both by natural processes such as selection and drift and by the agricultural practices of local farmers. How these processes interact to shape and change the population genetics of landrace rice is unknown [58]. Compared to new rice cultivars, rice landraces have more complex genetic backgrounds and more abundant genetic diversity and heterogeneity, as well as strong adaptability to the environment, excellent resistance to diseases and pests, high yields

and good quality [59]. The Southwest China, Guizhou, Yunnan and Guangxi provinces, is one of the largest centers of rice genetic diversity and high-quality germplasm in the world [21, 60].

The genetic variability found within landraces affords the possibility of genetic flexibility; landraces have the potential to adapt to local field conditions, and they can adapt to changing environments, farming practices and specific uses such as animal vs. human consumption [61]. Moreover, the genetic diversity of traditional landrace varieties is the most immediately useful and economically valuable component of rice biodiversity [19]. To efficiently conserve, manage and use such germplasm resources, an understanding of structure, apportionment and dynamics of local landrace variation is required. Several studies have examined genetic variation and differentiation among rice landrace varieties [62, 63]. However, little to no information is available on how genetic diversity is structured within a given landrace.

Local adaptation plays an important role in maintaining yields in traditional agricultural systems. Selection for adaptation to each village environment by the farmer's seed selection enhances overall crop diversity and maintains evolutionary flexibility [64]. Almekinders [21] explained that farmers' selection in combination with natural selection results in landraces with high levels of adaptation to biotic and abiotic stresses and for agricultural traits. For example, the genetic diversity of *Phaseolus vulgaris* landraces in Italy has been shaped by local adaptation to microenvironments [65], and in wheat, selection by farmers has strongly influenced the evolution of neutral loci [66].

13. Advantages of landrace over modern cultivated agriculture

1. Landraces provide a medium to advertise information about the conservation and use of crop landraces.
2. Crop improvement often utilizes landrace diversity in the development of new cultivars [8, 24], particularly when developing cultivars for marginal environments. Although, breeders more routinely focus their efforts on a limited gene pool of advanced cultivars or breeders' lines which are more easily utilized without successive backcrossing to eradicate the undesirable traits introduced with the desirable [67, 68].
3. Landraces still present a unique source of specific traits for disease and pest resistance, nutritional quality and marginal environment tolerance [8].
4. Landrace tolerance to advert climatic condition.
5. They have become important as sources of genetic variability in the search for genes for tolerance or resistance to biotic and abiotic factors of interest in agriculture.
6. Knowledge of genetic distance among landraces will help the breeding of high-yielding, good quality cultivars that will increase crop production [14].
7. Landraces may provide new alleles for the improvement of commercially valuable traits.
8. Global climate change emphasizes the need to use better adapted cultivars of the main crops and landraces as potential donors of useful genes.

9. There is an increasing consumer concern worldwide about food safety and nutrition. Landraces or old crop cultivars may prove solutions as sources of healthy and nutritious food.

14. Disadvantages

1. 1. Landraces are generally less productive than commercial cultivars.

15. Implication of replacement of landraces by commercial cultivars

1. It increases and causes genetic erosion by the replacement of diverse landraces with comparatively few, homozygous modern cultivars, which caused considerable concern among conservationists and breeders alike.
2. Landrace replacement by modern cultivars demonstrated a marked reduction in overall genetic diversity.
3. People's concerns over this rapid extinction or erosion of landrace diversity resulted in widespread action to promote their conservation.

16. Landrace conservation

Conservation of all gene pools is a high priority for sustaining food security and coping with current and future climate change effects. Not only must landraces be conserved, but so should local varieties that have been replaced by new and more productive ones. Older varieties, due to the emphasis on landraces and more exotic materials, must not be forgotten, and older varieties, as well as other breeding materials, need to be conserved as a source of genetic diversity. Despite the enormous efforts made by national and international programs to conserve landrace diversities, eventually the conservation of germplasm and characterization of key traits will provide specific information to breeders that will promote the use of genetic resources by the scientific community.

The discovery of abiotic stress at tolerant alleles in landraces of maize, rice and wheat clearly shows the importance of conserving and exploring landrace germplasm as a means to identify gnomonically beneficial alleles for enhancing adaptation and productivity in stress-prone environments [2].

16.1. Specifically, several challenges need attention:

i. Dealing with duplication where tracking is lost when moving germplasm from one place to other, particularly if a unique notation is not used
ii. Genetic diversity of collections widely determined by DNA markers available in gene bank facilities

iii. Diversity being well retained during collection through the use of molecular markers and visual observation and by using internationally accepted conservation and characterization standards in seed gene banks

iv. Increasing in situ conservation

v. Functional multiplication programs

vi. Organizing regular national or regional collection programs with functional surveys that gather high-quality information related to germplasm being collected

vii. Reliable 'passport' information being available with GPS coordinates

viii. Using internationally accepted database management programs in gene banks

ix. Providing a worldwide data system among gene banks

x. Spreading research results in a database system linked with gene banks

These activities, once established, will greatly improve the targeted use of genetic resources and will help scientists and breeders strategically extract and use allelic variation for important traits.

17. Conclusion and future perspective

Loss of genetic diversity has been recognized as a genetic bottleneck imposed on crop plants during domestication and through modern plant breeding practices. Allelic variation of genes originally found in the wild but gradually lost through domestication and breeding has been recovered only by going back to landraces. Landraces with increased biomass and total photosynthesis have potentially new allelic variation that should be exploited in plant breeding.

Landraces are heterogeneous with variable phenology, are low to moderate but stable edible yield and are often nutritionally superior. Traditional agricultural production systems in the past have played a vital role in the evolution and conservation of on-farm diversity, allowing farmers to circumvent crop failure by reducing vulnerability environmental stresses. A systematic evaluation of landraces for assessing the pattern of diversity is urgently needed to identify alleles for enhancing yield and adaptation to abiotic stress for raising the productivity of the staple food crops in stressful environments.

Acknowledgements

Authors thanked the authority of LAUTECH, Ogbomoso, Nigeria, for the provision of some of the facilities used in this study.

Conflict of interest

Authors declare no conflict of interest.

Author details

Musibau A. Azeez*, Amos O. Adubi and Felicia A. Durodola

*Address all correspondence to: maazeez@lautech.edu.ng

Laboratory of Ecotoxicology, Genetics and Nanobiotechnology, Environmental Biology Unit, Department of Pure and Applied Biology, Ladoke Akintola University of Technology Ogbomoso, Nigeria

References

[1] Camacho-Villa TC, Maxted N, Scholten M, Ford-Lloyd B. Defining and identifying crop landraces. Plant Genetic Resources, Characterization and Utilization. 2005;**3**:373-384

[2] Dwivedi SL, Ceccarelli S, Blair MW, Upadhyaya HD, Are AK, Ortiz R. Landrace germplasm for improving yield and abiotic stress adaptation. Trends in Plant Science. 2016;**21**(1):31-42

[3] Zeven AC. Landraces: A review of definitions and classifications. Euphytica. 1998; **104**:127-139

[4] Paterniani E, Nass LL, Santos MX. O Valor dos RecursosGenéticos de Milhopara o Brasil: Uma AbordagemHistórica da UtilizaçãodoGermoplasma. In: Udry CW, Duarte W, editors. Uma HistóriaBrasileiradoMilho: O Valor dos RecursosGenéticos. Brasília: Paralelo; 2000. pp. 11-14

[5] Araújo PM, Nass LL. Caracterização e avaliação de populações de milho crioulo. Science in Agriculture. 2002;**59**:589-593

[6] Pollack LM. The history and success of the public-private project on germplasm enhancement of maize (GEM). Advances in Agronomy. 2003;**78**:45-87 Valeria Negri Renzo Torricelli Nigel

[7] Harlan JR. Our vanishing genetic resources. Science. 1975;**188**:618-621

[8] Frankel OH, Brown AHD, Burdon JJ. The Conservation of Plant Biodiversity. 2nd ed. Cambridge: Cambridge University Press; 1998. pp. 56-78

[9] Frankel OH, Bennett E. Genetic resources—Introduction. In: Frankel OH, Bennett E, editors. Genetic Resources in Plants—Their Exploration and Conservation. International Biological Programme Handbook No. 11. Oxford: Blackwell; 1970. pp. 1-32

[10] Frankel OH, Hawkes JG, editors. Crop Genetic Resources for Today and Tomorrow. Cambridge: Cambridge University Press; 1975

[11] von Rünker K. Die systematischeEinteilung und Benen-ung der Getreidesortenfu¨rpr aktischeZwecke. Jahrbuch der Deutschenlandwirtschafts-Gesellschaft. 1908;**23**:137-167

[12] Tudge C. Food Crops for the Future. Oxford: Basil Blackwell; 1988. p. 83

[13] Villa TCC, Maxted N, Scholten M and Ford-Lloyd B. Defining and identifying crop landraces. Plant Genetic Resources. 2006;**3**(3):373-384

[14] FAO (Food and Agricultural Organisation) (1998). The State of the World's Genetic Resources for Food and Agriculture. Rome, Italy: FAO; 1998

[15] Almekinders CJM, Louwaars NP. Farmer's Seed Production: New Approaches and Practices. London: Intermediate Technology Publications; 1999

[16] Louette D, Charrier A, Berthaud J. In situ conservation of maize in Mexico: Genetic diversity and maize seed management in a traditional community. Economic Botany. 1997;**51**:20-38

[17] Hawkes JG. The Diversity of Crop Plants. Cambridge, MA: Harvard University Press; 1983. p. 102

[18] Bellon MR, Brush S. Keepers of maize in Chiapas, Mexico. Economic Botany. 1994;**48**:196-209

[19] Wood D, Lenné JM. The conservation of agrobiodiversity on-farm: Questioning the emerging paradigm. Biodiversity and Conservation. 1997;**6**:109-129

[20] Zeven AC. The traditional inexplicable replacement of seed and seed ware of landraces and cultivars: A review. Euphytica. 1999;**110**:181-191

[21] Almekinders CJM, Louwaars NP, de Bruijn GH. Local seed systems and their importance for an improved seed supply in developing countries. Euphytica. 1994;**78**:207-216

[22] Louette D, Smale M. Genetic diversity and maize seed management in a traditional Mexican community: Implications for in situ conservation of maize. Natural Resources Group, Paper 96-03: International Centre for Maize and Wheat Improvement (CIMMYT); 1996. 22 pp

[23] Banga O. Veredeling van tuinbouwgewassen. Zwolle: Bergboek; 1944

[24] Frankel OH. Natural variation and its conservation. In: Muhammed A, Aksel R, Von Borstel RC, editors. Genetic Diversity in Plants. New York: Plenum Press; 1977. pp. 29-34

[25] Schindler J. EinigeBemerkungenüber die züchterische und wirtschaftlicheBedeutung der LandrassenunsererKul-urpflanzen. Deutsche Landwirt. Presse. 1918;**45**(25):155-156

[26] Brush SB. The issues of in situ conservation of crop genetic resources. In: Brush S, editor. Genes in the Field. Rome: International Plant Genetic Resources Institute; 1999. pp. 3-26

[27] Zeven AC. Traditional maintenance breeding of landraces: 1 data by crop. Euphytica. 2000;**116**:65-85

[28] Maxted. Novel characterization of crop wild relative and landrace resources as a basis for improved crop breeding; 2013

[29] Almekinders CJM, Elings A. Collaboration of farmers and breeders: Participatory crop improvement in perspective. Euphytica. 2001;**122**:425-438

[30] Simmonds NW. Principles of Crop Improvement. London: Longman Group; 1979

[31] Allard RW. Principles of Plant Breeding. 2nd ed. New York: John Wiley and Sons; 1999

[32] Virchow D. Conservation of Genetic Resources: Costs and Implications for a Sustainable Utilization of Plant Genetic Resources for Food and Agriculture. Berlin: Springer; 1999. p. 4

[33] Scholten MA, Maxted N and Ford-Lloyd BV. UK National Inventory of Plant Genetic Resources for Food and Agriculture. School of Biosciences, University of Birmingham. 2004. p. 95. www.brochwell-bake.org.uk/pics/UK_inventory.pdf

[34] Marchenay P. A la recherche des varie'te's locales de plantescultive'es. Paris: Burean des resources génétiques; 1987

[35] Asfaw Z. The barley in Ethiopia. In: Brush S, editor. Genes in the Field. International Plant Genetic Resources Institute: Rome; 1999. pp. 77-107

[36] Cleveland DA, Soleri D, Smith SE. Folk crop varieties: Do they have a role in sustainable agriculture? BioScience. 1994;**44**:740-751

[37] Fowler C, Mooney P. Shattering: Food, Politics and the Loss of Genetic Diversity. Tucson: University Arizona Press; 1990

[38] Hoyt E. Conserving the Wild Relatives of Crops. Rome: IBPGR, IUCN, WWF; 1992

[39] Brown AHD. The genetic structure of crop landraces and the challenge to conserve them in situ on farms. In: Brush S, editor. Genes in the Field. International Plant Genetic Resources Institute: Rome; 1999. pp. 29-48

[40] Frankel OH, Soule´ ME. Conservation and Evolution. Cambridge: Cambridge University Press; 1981. pp. 177-223

[41] Sanchez JJ, Goodman MM, Stuber CW. Isozymatic and morphological diversity in the races of maize in Mexico. Economic Botany. 2000;**54**:43-59

[42] Jarman RJ. Bere barley: A living link with 8th century. Plant Varieties and Seeds. 1996;**9**:191-196

[43] Choo T-M. Genetic Resources of Tibetan Barley in China. Beijing: Ma Dequan, China Agriculture Press; 2002

[44] Bennett E. Adaptation in wild and cultivated plant populations. In: Frankel OH, Bennett E, editors. Genetic Resources in Plants—Their Exploration and Conservation. International Biological Programme Handbook No. 11. Oxford: Blackwell; 1970. pp. 115-129

[45] Altieri M, Merrick LC. In situ conservation of crop genetic resources through maintenance of traditional farming systems. Economic Botany. 1987;**41**:86-96

[46] Bellon MR. The dynamics of crop infraspecific diversity: A conceptual framework at the farmer level. Economic Botany. 1996;**50**:26-39

[47] Boster J. Human cognition as a product and agent of evolution. In: Ellen R, Fukui K, editors. Redefining Nature: Ecology, Culture and Domestication. Oxford: Berg; 1996. pp. 269-289

[48] Zimmermer KS. Managing diversity in potato and maize fields of the Peruvian Andes. Journal of Ethnobiology. 1991;**11**(1):23-49

[49] Maxted N, Ford-Lloyd B, Hawkes JG. Complementary conservation strategies. In: Maxted N, Ford-Lloyd B, Hawkes JG, editors. Plant Genetic Conservation: The In Situ Approach. London: Chapman and Hall; 1997. pp. 15-39

[50] Qualset CO, Damania AB, ACA Z, Brush SB. Locallybased cropplant conservation. In: Maxted N, Ford-Lloyd BV, Hawkes JG, editors. Plant Genetic Conservation: The In Situ Approach. London: Chapman and Hall; 1997. pp. 160-175

[51] Brush SB. In situ conservation of landraces in centers of crop diversity. Crop Science. 1995;**35**:346-354

[52] Sarker A, Erskine W. Resent progress in the ancient lentil. The Journal of Agricultural Science. 2003;**144**:19-29

[53] Hammer K et al. Estimating genetic erosion in landraces- two cases studies. Genetic Resources and Crop Evolution. 1996;**43**:329-336

[54] Lorenzetti F, Negri V. The European seed legislation on conservation varieties. In: Vetelainen M, Negri V, Maxted N. editors. "European landraces: On-farm Conservation, Management and Use" Bioversity Technical Bullettin No 15. Rome, Italy: Bioversity International Publ.; 2009. pp 287-295

[55] Lafia M. Contribution a l'etude des systemessemencierstraditionnels au Benin :cas du mais, du sorgho, du niebe, du soja, et du goussi. These de DEA, FSA/UAC; 2006. p. 103

[56] Nasserlehaq N, Amamou A, Taghouti M, Annicchiarico P. Adaptation of Moroccan durum wheat varieties from different breeding eras. Journal of Plant Breeding and Crop Science. 2011;**3**(2):34-40

[57] Nachit MM. Durum wheat breeding for Mediterranean dry land of North Africa and West Asia. In: Nachit MM, Baum M, Impiglia A, Ketata H., 1995. Studies on Some Quality Traits in Durum Wheat Grown in Mediterranean Environments.*Options Mediterraneennes*. Serie A. ICARDA/CIHEAM/CIMMYT. 1992;**22**:181-188

[58] Pusadee T, Jamjod S, Chiang Y, Rerkasem B, Schaal BA. Genetic structure and isolation by distance in a landrace of Thai rice. PNAS. 2009;**106**(33):13880-13885

[59] Qiyw ZDL, Zhang HL, Wang MX, Sun JL, Wei XH, Qiu ZE, Tang SX, Cao YS, Wang XK. Genetic diversity of rice cultivars (Oryza sativa L.) in China and the temporal trends in recent fifty years. Chinese Science Bulletin. 2006;**51**(6):681-688. DOI: 10.1007/s11434-006-0681-8

[60] Glaszman JC. Isozymes and classification of Asian rice varieties. Theoretical and Applied Genetics. 1998;**74**(1):21-30. DOI: 10.1007/BF00290078. 34

[61] McCouch S. Diversifying selection in plant breeding. PLoS Biology. 2004;**2**:1507-1512

[62] Fukuoka S et al. Diversity in phenotypic profiles in landrace populations of Vietnamese rice: A case study of agronomic characters for conserving crop genetic diversity on farm. Genetic Resources and Crop Evolution. 2006;**53**:753-761 Available from: 13884_www.pnas.org_cgi_doi_10.1073_pnas.0906720106 Pusadee et al

[63] Bajracharya J et al. Rice landrace diversity in Nepal: Variability of agromorphological traits and SSR markers in landraces from a high-altitude site. Field Crops Research. 2006;**95**:327-335

[64] Alvarez N et al. Farmers' practices, metapopulation dynamics, and conservation of agricultural biodiversity on-farm: A case study of sorghum among the Duupa in sub-sahelian Cameroon. Biological Conservation. 2005;**121**:533-543

[65] Tiranti B, Negri V. 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. 2007;**16**:4942-4955

[66] Goldringer I, Enjalbert J, Raquin A-L, Brabant P. Strong selection in wheat populations during ten generations of dynamic management. Genetics, Selection, Evolution. 2001;**33**:441-463

[67] Duvick DN. Genetic diversity in major farm crops on the farm and in reserve. Economic Botany. 1984;**38**:162-178

[68] Peeters JP, Galwey NW. Germplasm collections and breeding needs in Europe. Economic Botany. 1988;**42**:503-521

Chapter 2

Phenolic Fingerprinting and Glumes Image Analysis as an Effective Approach for Durum Wheat Landraces Identification

Oscar Grillo, Marisol Lo Bianco,
Sebastiano Blangiforti, Laura Siracusa,
Gianfranco Venora and Giuseppe Ruberto

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.79595

Abstract

In the last decade, encouraged by economic, social and nutritional reasons, the trend towards the rediscovery and reuse of durum wheat landraces moved on in Sicily. This growing attention in local wheat landraces made necessary to design new effective and objective identification methods that are able to distinguish landraces. Considering the difficulties coming from the genetic and morphological heterogeneity of a landrace, in this chapter a multidisciplinary approach for durum wheat landraces identification is proposed. Nine Sicilian wheat landraces were investigated from the genotypic and phenotypic point of view, studying their polyphenolic profile, and analyzing the glumes morpho-colorimetric traits, in search of similarities and/or differences. In particular, hydro-alcoholic extracts from whole wheat grains were analyzed by means of HPLC/DAD and HPLC/ESI-MS, revealing 13 metabolites mainly belonging to the classes of hydroxycinnamic acids and flavones *C*-glycosides. The quantitative pattern of the 13 phenolic markers allowed to perfectly identify all the wheat samples, confirming a specific and genotype-dependent pattern of phenolics concentration. Moreover, computerized image analysis techniques were applied to compare the wheat samples on the basis of 138 quantitative morpho-colorimetric variables descriptive of glumes size, shape, color and texture, confirming the possibility to undoubtedly identify wheat samples belonging to local landraces.

Keywords: traceability, biodiversity, local populations, morpho-colorimetric analysis, old varieties, phytochemicals, polyphenols, *Triticum* L

1. Introduction

After the massive use of processed acorns, as food source for prehistoric nomadic populations, the most important food discovery was undoubtedly that of cereals. Even now, wheat (*Triticum*) is one of the main food sources in the world. According to the last FAO report, wheat world production for 2017 was approximately expected in 740 million tons, exceeding the previous last one crop year by 1.2%, and covering about 15% of the world's arable surface [1]. In this scenario, durum wheat production reaches around 30 million tons, accounting approximately for 5–6% of the total world wheat production [2]. Canada, USA, Mexico and Russia are some of the countries around the world where durum wheat is cropped, although the Mediterranean region covers about 60% of world durum wheat production [3], being the EU (Italy, Spain, France and Greece) the leading global producer [2]. South Italy is one of the regions historically most voted to the cereal crops, where the durum wheat varietal biodiversity is particularly high [4].

For geographical position and ecological condition, Sicily represents the perfect environment for the cultivation of cereals, especially for durum wheat. In addition to the pedo-climatic conditions [5], some historical and socio-cultural aspects had also contributed to enrich the varietal heritage, such as the many invasions that characterized the island during the centuries. All these conditions, together with the historically conducted mass selection and the more recent genetic improvement programs based on the artificial crosses, had contributed to build the extremely wide currently existing varietal panorama [4].

In Sicily, old and new durum wheat commercial varieties are currently cropped, but also many ancient landraces or populations characterized by specific bio-morphological traits and qualitative features [6, 7].

In recent years, all over the world, the attention paid to local and traditional productions is growing, especially in the agro-food sector. Maybe, it is due to the impact of globalization and the social and economic changes, but also to the increased consideration to health and nutritional aspects of food. Also in Sicily, this trend has led to the rediscovery and reuse of landraces both of wheat and other crops, responding to requests for more and more demanding market. The rising price of these local productions are contributing to the farmers' satisfaction, changing an unprofitable job in a renewed professional opportunity also for young businessmen. Furthermore, many recent research studies testify the high healthy and nutraceutical value of landraces, both for high amount of antioxidant compounds and for their natural aptitude to organic production [8–11].

This growing interest in local landraces has inspired to find effective and objective identification methods, able to distinguish landraces [12, 13].

In this chapter a multidisciplinary practical approach based on genotype and phenotype characterization of durum wheat Sicilian landraces is proposed. In particular, the polyphenolic profile of whole wheat grains was analyzed by means of HPLC/DAD and HPLC/ESI-MS.

Moreover, computerized image analysis techniques were applied to compare glume wheat samples, implementing a statistical classificator able to discriminate the landraces.

2. Materials and methods

2.1. Polyphenolic profile analysis

2.1.1. Samples details

Nine durum wheat (*Triticum durum* Desf.) landraces ("Margherito," "Manto di Maria," "Ruscia," "Russello SG8," "Scavuzza," "Tumminia SG3," "Trentino," "Tripolino," "Urria") were selected for phenolics profile evaluation.

Grains were cropped, in three plots of 10 m^2 each, using 350 viable seeds/m^2, during three consecutive years (2012, 2013, 2014), in the fields of the Stazione Sperimentale di Granicoltura per la Sicilia, sited in Santo Pietro - Caltagirone [37°07′12″N; 14°31′17″E; 313 m a.s.l.] (CT, Sicily, Italy). 40 kg N/ha and 90 kg P_2O_5/ha were supplied at sowing carried out at the beginning of December; nitrogen fertilization with 50 kg N/ha were applied before the beginning of stem elongation stage (20–30 code in the BBCH-scale for cereals). Mechanical weed control methods were carried out in spring time and harvest was performed when physiological maturity of each genotype was reached.

Whole grain samples were milled to a fine powder by a laboratory mill (1093 Cyclotec Sample Mill, Tecator Foss, Hillerød, Denmark) equipped with a 1 mm sieve, immediately cooled to −20°C and kept at this temperature until analysis to protect bioactive components from degradation [14].

2.1.2. Chemicals

All solvents and reagents used in this study were high purity laboratory solvents by Carlo Erba (Milano, Italy); HPLC grade water and acetonitrile were obtained from VWR (Milano, Italy). Pure vitexin (apigenin 8-*C*-glucoside) and orientin (luteolin 8-*C*-glucoside) were provided by Extrasynthese (Lyon, France) whilst vanillic acid, ferulic acid, p-coumaric acid and caffeic acid were purchased from Sigma (Sigma-Aldrich s.r.l., Milano, Italy).

2.1.3. Extraction of free and bound phenolic compounds

Phenolic acids and flavonoids represent the most common form of phenolic compounds found in whole grains, existing as soluble free compounds, soluble conjugates esterified to sugars and other low molecular mass components, and insoluble bound forms either encapsulated in the cell-wall structures or chemically bound at molecular level [15].

According to Lo Bianco et al. [11], free phenolics were recovered by applying the method proposed by Dinelli et al. [14] with few changes. In brief, 1 g of whole wheat flour was mixed

under vigorous stirring for 10 min with 20 mL of an acidic aqueous methanol solution (80% methanol, 19% water, 1% formic acid). The resulting heterogeneous mixture was transferred into standard glass sample tubes and centrifuged at 2500 g/min for 10 min. After that, the supernatant was removed and the extraction was repeated. Collected supernatants were pooled, evaporated to dryness, and then stored at −20°C until use.

The solid residue from the free phenolic extraction was subjected to alkaline hydrolysis to recover the bound phenolic compounds, according to Mattila et al. [16]. Distilled water (12 mL) and 5 mL of 10 M NaOH were added to the residue and stirred overnight at room temperature. The mixture was acidified to pH = 2 and then extracted three times with 15 mL of a 1:1 (v/v) mixture of cold diethyl ether and ethyl acetate by manually shaking and centrifuging. Organic layers were combined, evaporated to dryness, and dissolved into 2 mL of the aqueous methanol solution to analytical determinations.

2.1.4. HPLC/DAD quantitative analyses

For HPLC/DAD analyses dry extracts were reconstituted in 3 mL of the extracting solvent and immediately analyzed. Quantitative analyses were carried out on a UltiMate3000 "UHPLC focused" instrument equipped with a binary high pressure pump, a Photodiode Array detector, a Thermostatted Column Compartment and an Automated Sample Injector (Thermo Scientific, Italy). Collected data were processed through a Chromeleon Chromatography Information Management System v. 6.80. Chromatographic runs were all performed using a reverse-phase column (Gemini C_{18}, 250 × 4.6 mm, 5 μm particle size, Phenomenex, Italy) equipped with a guard column (Gemini C_{18} 4 × 3.0 mm, 5 μm particle size, Phenomenex, Italy). Wheat polyphenols were eluted with the following gradient of B (formic acid, 2.5% solution in acetonitrile) in A (2.5% solution of formic acid in water): 0 min: 5% B; 10 min: 15% B; 30 min: 25% B; 35 min: 30% B; 50 min: 90% B; then kept for 7 min at 100% B. The solvent flow rate was 1 mL/min and. Quantifications were carried out at 350 nm using orientin (R^2 = 0.9999) as external standard; the detector was set at 280 nm to build the calibration curve for vanillic acid (R^2 = 0.9997), whilst vitexin (R^2 = 0.9999), caffeic acid and ferulic acid were quantified at 330 nm using the corresponding reference substances (R^2 = 0.9999 and R^2 = 0.9998, respectively). The same reference wavelength was used for the quantification of coumarins against p-coumaric acid (R^2 = 0.9998). All analyses were carried out in triplicate.

2.1.5. Identification of main components via HPLC/ESI-MS

In order to unambiguously identify the chromatographic signals and/or to confirm peak assignments, a series of HPLC/ESI/MS analyses were performed on wheat samples. In this case, variable aliquots (1.0–1.5 mL) of the above mentioned hydro-alcoholic solutions coming from quantitative analyses (see previous paragraph) were transferred into standard laboratory vials and brought to dryness *in vacuo* with a rotary evaporator (Heidolph Laborota 400). The resulting yellowish residues were then re-dissolved in 500 μL of the original hydroalcoholic solution and submitted to qualitative analyses. The HPLC apparatus used was the same described above, whilst ESI mass spectra were acquired by a Thermo Scientific Exactive Plus Orbitrap MS (Thermo Fisher Scientific, Inc., Milan, Italy), using a heated electrospray ionization (HESI II) interface. Mass spectra were recorded operating in negative ion mode in the *m/z*

range 120–1800 at a resolving power of 25,000 (full-width-at-half-maximum, at *m*/*z* 200, RFWHM), resulting in a scan rate of >1.5 scans/s when using automatic gain control target of 1.0×10^6 and a C-trap inject time of 250 ms, under the following conditions: capillary temperature 300°C, nebulizer gas (nitrogen) with a flow rate of 60 arbitrary units; auxiliary gas flow rate of 10 arbitrary units; source voltage 3 kV; capillary voltage 82.5 V; tube lens voltage 85 V. The Orbitrap MS system was tuned and calibrated in positive modes, by infusion of solutions of a standard mixture of sodium dodecyl sulfate (Mr 265.17 Da), sodium taurocholate (Mr 514.42 Da) and Ultramark (Mr 1621 Da). Data acquisition and analyses were performed using the Xcalibur software.

2.2. Glume image analysis

2.2.1. Samples details

For the glumes image analysis, ears of the same ten wheat landraces were reaped, at the time of maximum ripening, in order to include a widest morphological and environmental variability, the wheat ears were collected during three consecutive years (2012, 2013, 2014).

From three to six ears were sampled and from two to four glumes were removed from the spikelets of the ear middle section and from the both sides of each ear. The glumes were stored at room temperature under controlled conditions (20°C and 50% RH).

2.2.2. Images acquisition

Digital images of glumes samples were acquired using a flatbed scanner (ScanMaker 9800 XL, Microtek Denver, CO), applying the same resolution and scanning area conditions reported in Grillo et al. [4]. As suggested by Venora et al. [17], before digital image capture, the scanner was standardized according to the calibration protocol proposed by Shahin and Symons [18]. Morpho-colorimetric features were only measured for sound intact glumes, rejecting that ones with broken beak or shoulder, distinguishing in right and left side of the ear. A total of 902 wheat glumes were analyzed (**Table 1**).

Code	Variety/landrace	Sample amount
mar6	Margherito 06	97
mm1	Manto di Maria 01	95
rsc9	Ruscia 09	97
russg8	Russello 13 SG8	97
sca1	Scavuzza 01	95
tre2	Trentino 02	119
tri2	Tripolino 02	80
tumsg3	Tumminia SG3	94
urr1	Urrìa 01	88

Table 1. List of the ten different wheat local varieties studied.

2.2.3. Image processing and analysis

All the images were processed and analyzed using the software package KS-400 V. 3.0 (Carl Zeiss, Vision, Oberkochen, Germany). The same macro used by Grillo et al. [4], specifically developed for the characterization of wheat glumes was applied to perform automatically all the morpho-colorimetric measurements on the glume samples of the present study.

The macro allowed to compute 138 quantitative variables measured for each analyzed left and right glume (**Tables 2** and **3**). In particular, it was possible to measure 18 parameters descriptive of the glume size and shape and 20 features descriptive of the glume surface color. Afterwards, applying the same procedure reported by Orrù et al. [19], 78 quantitative Elliptic Fourier Descriptors (EFDs) were used to describe the shape of the glume. Finally, the macro was kitted to compute 11 Haralick's descriptors including the relative standard deviations, as reported in Lo Bianco et al. [20].

$$G = \begin{bmatrix} p(1,1) & p(1,2) & \cdots & p(1,N_g) \\ p(2,1) & p(2,2) & \cdots & p(2,N_g) \\ \vdots & \vdots & \ddots & \vdots \\ p(N_g,1) & p(N_g,2) & \cdots & p(N_g,N_g) \end{bmatrix} \quad (1)$$

2.3. Statistics

The data, obtained from chemical and image analysis, were used to build a global database. Statistical elaborations were executed using SPSS software package release 16.0 (SPSS Inc. for Windows, Chicago, Illinois, USA), and the stepwise Linear Discriminant Analysis (LDA) method was applied to identify and discriminate among the investigated wheat samples [23]. This approach is commonly used to classify/identify unknown groups characterized by quantitative and qualitative variables [24–27], finding the combination of predictor variables with the aim of minimizing the within-class distance and maximizing the between-class distance simultaneously, thus achieving maximum class discrimination [28–31]. Then, the stepwise procedure, carried out as explained in [4], identifies and selects the most statistically significant features among the chemical metabolites and the 138 traits measured on each glume. Finally, a cross-validation procedure was applied to verify the performance of the identification system, testing individual unknown cases and classifying them on the basis of all others [32].

All the raw data were standardized before starting any statistical elaboration. Moreover, in order to evaluate the quality of the discriminant functions achieved for each statistical comparison, the Wilks' Lambda, the percentage of explained variance and the canonical correlation between the discriminant functions and the group membership, were computed. The Box's M test was executed to assess the homogeneity of covariance matrices of the features chosen by the stepwise LDA while the analysis of the standardized residuals was performed to verify the homoscedasticity of the variance of the dependent variables used to discriminate among the groups' membership [33]. Kolmogorov-Smirnov's test was performed to compare the empirical distribution of the discriminant functions with the relative cumulative distribution function of the reference probability distribution, while the Levene's test was executed to assess the equality of variances for the used discriminant functions calculated for groups membership [34].

	Feature	Equation
Har 1	Angular second moment	$\sum_i \sum_j p(i,j)^2$
Har 2	Contrast	$\sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j) \right\}, \lvert i,j \rvert = n$
Har 3	Correlation	$\frac{\sum_i \sum_j (ij) p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}$ where μ_x, μ_y, σ_x and σ_y are the means and the standard deviations of p_x and p_y.
Har 4	Sum of square: variance	$\sum_i \sum_j (i-\mu)^2 p(i,j)$
Har 5	Inverse difference moment	$\sum_i \sum_j \frac{1}{1+(i-j)^2} p(i,j)$
Har 6	Sum average	$\sum_{n=2}^{2N_g} i p_{x+y}(i)$ where x and y are the coordinates (row and column) of an entry in the co-occurrence matrix, and $p_{x+y}(i)$ is the probability of co-occurrence matrix coordinates summing to $x + y$.
Har 7	Sum variance	$\sum_{i=2}^{2N_g} (i - f_8)^2 p_{x+y}(i)$
Har 8	Sum entropy	$-\sum_{i=2}^{2N_g} p_{x+y}(i) \log\{p_{x+y}(i)\} = f_8$
Har 9	Entropy	$-\sum_i \sum_j p(i,j) \log[p(i,j)]$
Har 10	Difference variance	$\sum_{n=0}^{N_{g-1}} i^2 p_{x-y}(i)$
Har 11	Difference entropy	$-\sum_{n=0}^{N_{g-1}} p_{x-y}(i) \log\{p_{x-y}(i)\}$

The basis for these features is the gray-level co-occurrence matrix (G in Eq. (1)). This matrix is square with dimension Ng, where Ng is the number of gray levels in the image. Element $[i,j]$ of the matrix is generated by counting the number of times a pixel (p) with value i is adjacent to a pixel with value j and then dividing the entire matrix by the total number of such comparisons made. Each entry is therefore considered to be the probability that a pixel with value i will be found adjacent to a pixel of value j.

Table 2. Haralick's descriptors measured as reported in Haralick et al. [21].

	Feature	Description
A	Area	Seed area (mm^2)
P	Perimeter	Seed perimeter (mm)
P_{conv}	Convex Perimeter	Convex perimeter of the seed (mm)
P_{Crof}	Crofton Perimeter	Crofton perimeter of the seed (mm)
P_{conv}/P_{Crof}	Perimeter ratio	Ratio between convex and Crofton's perimeters
D_{max}	Max diameter	Maximum diameter of the seed (mm)
D_{min}	Min diameter	Minimum diameter of the seed (mm)
D_{min}/D_{max}	Feret ratio	Ratio between minimum and maximum diameters
Sf	Shape factor	Seed shape descriptor = $(4 \times \pi \times \text{area})/\text{perimeter}^2$ (normalized value)
Rf	Roundness factor	Seed roundness descriptor = $(4 \times \text{area})/(\pi \times \text{max diameter}^2)$ (normalized value)
Ecd	Eq. circular diameter	Diameter of a circle with equivalent area (mm)
F	Fiber length	Seed length along the fiber axis
C	Curl degree	Ratio between $\boldsymbol{D_{max}}$ and $\boldsymbol{F}$
$Conv$	Convessity degree	Ratio between $\boldsymbol{P_{Crof}}$ and $\boldsymbol{P}$
Sol	Solidity degree	Ratio between $\boldsymbol{A}$ and convex area
Com	Compactness degree	Seed compactness descriptor = $[\sqrt{(4/\pi)}\,\boldsymbol{A}]/\boldsymbol{D_{max}}$
EA_{max}	Maximum ellipse axis	Maximum axis of an ellipse with equivalent area (mm)
EA_{min}	Minimum ellipse axis	Minimum axis of an ellipse with equivalent area (mm)
R_{mean}	Mean red channel	Red channel mean value of seed pixels (gray levels)
R_{sd}	Red std. deviation	Red channel standard deviation of seed pixels
G_{mean}	Mean green channel	Green channel mean value of seed pixels (gray levels)
G_{sd}	Green std. deviation	Green channel standard deviation of seed pixels
B_{mean}	Mean blue channel	Blue channel mean value of seed pixels (gray levels)
B_{sd}	Blue std. deviation	Blue channel standard deviation of seed pixels
H_{mean}	Mean hue channel	Hue channel mean value of seed pixels (gray levels)
H_{sd}	Hue std. deviation	Hue channel standard deviation of seed pixels
L_{mean}	Mean lightness ch.	Lightness channel mean value of seed pixels (gray levels)
L_{sd}	Lightness std. dev.	Lightness channel standard deviation of seed pixels
S_{mean}	Mean saturation ch.	Saturation channel mean value of seed pixels (gray levels)
S_{sd}	Saturation std. dev.	Saturation channel standard deviation of seed pixels
D_{mean}	Mean density	Density channel mean value of seed pixels (gray levels)
D_{sd}	Density std. deviation	Density channel standard deviation of seed pixels
S	Skewness	Asymmetry degree of intensity values distribution (gray levels)
K	Kurtosis	Peakness degree of intensity values distribution (densit. units)
H	Energy	Measure of the increasing intensity power (densitometric units)

Table 3. List of morphometric features measured on seeds, excluding the elliptic Fourier descriptors (EFDs) calculated according to Hâruta [22] and the Haralick's descriptors reported in **Table 2**.

To graphically highlight the differences among groups, multidimensional plots were drawn using the first three discriminant functions.

3. Results and discussion

3.1. Phenolic profile in wheat landraces

Phenolics are mainly concentrated in the outer layers of kernel and contribute to the wheat flour nutraceutical value owing to their antioxidant, anti-inflammatory and anticancer properties [35]. In literature, ca. 70 different phenolic compounds, including coumarins, phenolic acids, anthocyanins, flavones, isoflavones, proanthocyanidins, stilbenes and lignans, were identified in durum wheat genotypes [14].

Referring to flavones, whose interest has grown enormously due to their putative beneficial effects against atherosclerosis, osteoporosis, diabetes mellitus and certain cancers [36] 5,7,4′-trihydroxyflavone (apigenin) and 5,7,3′,4′-tetrahydroxyflavone (luteolin) are the main representatives in wheat, where they accumulate as 6-*C* and/or 8-*C*-glycosidic conjugates. The 8-*C*-glucosides of apigenin and luteolin are also known as vitexin and orientin, respectively.

Hydro-alcoholic extracts from wheat grains were exhaustively analyzed by means of HPLC/DAD and HPLC/ESI-MS. Although the major portion of phenolics in grains exist in the bound form [37], there is a general trend for studying polyphenols in the free form when dealing with chemotaxonomic studies [38, 39]. The chromatograms relating to free phenolics profile of durum wheat grains showed ca. 20 different signals, eluting in the range from 7 to 30 min. Among these, 13 signals were tentatively identified: a preliminary analysis of the UV–VIS (in terms of spectrum shape and absorption maximum, see **Table 4**) spectra of the peaks revealed the presence of compounds belonging to the chemical subclasses of hydroxycinnamic acids and organic acids; several peaks showing the typical spectrum of apigenin derivatives were also detected.

The use of mass spectrometry as detector was helpful in tentatively identifying wheat metabolites (**Table 4**); peak assignments were further confirmed by comparison with literature data [14, 40, 41] and co-injection with pure reference standards when available (see material and methods).

According to Lo Bianco et al. [11], three hydroxycinnamic acids were identified in wheat grains: caffeic acid (peak 4), ferulic acid (peak 10) and another member of this class (peak 1) for which unfortunately the MS spectrum was not determined. Vanillic acid (peak 3) was identified for its diagnostic UV–VIS and mass spectrum; the assignment was confirmed with co-injection with the corresponding standard. Peaks 2 and 6, showing almost identical UV–VIS spectra (a symmetrical absorption with λ max = 317 nm) were tentatively identified as coumarins; furthermore, peak 2 showed a clear mass spectrum with a pseudomolecular ion at 145.14 m/z $(M\text{-}H)^-$. Presence of coumarins in durum wheat has been reported by other authors [14, 40]. The UV–VIS spectrum of peak 5 (λ max = 268, 35 nm) was typical of that of luteolin derivatives; the corresponding mass spectrum exhibited a base peak of 609.52 m/z (pseudomolecular ion) with no signals ascribable to fragments generated by the loss of sugars. The peak corresponding to luteolin aglycone was

Peak #	Retention time (min)	λ ass. (nm)	Selected ion	*m/z* calculated	Tentative identification	Phenolic subclass
1	13.66	295(sh), 316		n.d	Hydroxycinnamic acid	Hydroxycinnamic acid
2	14.24	317	$[M-H]^-$	145.14	Coumarin	Coumarin
3	14.57	258, 291	$[M-H]^-$	167.04	Vanillic acid	Hydroybenzoic acid
4	14.98	290 (sh), 323	$[M-H]^-$	179.16	Caffeic acid	Hydroxycinnamic acid
5	16.57	268, 348	$[M-H]^-$	609.52	Luteolin di-*C*-hexoside (lucenin-2 isomer)	Flavone-*C*-glycoside
6	17.10	316		n.d	Coumarin	Coumarin
7	17.99	270, 334	$[M-H]^-$	563.14	Apigenin *C*-hexoside-*C*-pentoside	Flavone-*C*-glycoside
8	18.95	270, 335	$[M-H]^-$	563.14	Apigenin *C*-hexoside-*C*-pentoside	Flavone-*C*-glycoside
9	19.52	271, 335	$[M-H]^-$	563.14	Apigenin *C*-hexoside-*C*-pentoside	Flavone-*C*-glycoside
10	22.54	295 (sh), 323	$[M-H]^-$	193.05	Ferulic acid	Hydroxycinnamic acid
11	25.18	272, 332	$[M-H]^-$	769.18	Apigenin *C*-hexoside-*C*-hexoside *O*-glucuronide	Flavone-*C*-glycoside
12	26.00	272, 332	$[M-H]^-$	769.18	Apigenin *C*-hexoside-*C*-hexoside *O*-glucuronide	Flavone-*C*-glycoside
13	28.99	270, 334	$[M-H]^-$	431.10	Apigenin *C*-hexoside	Flavone-*C*-glycoside

(sh) is for shoulder.

Table 4. Phenolic compounds detected in the free form extracts from durum wheat grains.

absent as well. These data are usually diagnostic of the presence of *C*-bound glycosides; the peak was then tentatively identified as luteolin di-*C*-hexoside (lucenin-2 isomer). This is in discordance to what was reported by Dinelli et al. [14, 40] who found in durum wheat grains several isomers of lucenin 1/3, the *C*-hexoside-*C*-pentoside derivative of luteolin. Peaks 7, 8, 9, 11, 12 and 13 showed UV–VIS spectra whose shapes and absorption maxima clearly recalled apigenin (**Table 4**); in this case mass analysis was determinant in the assignments. Peaks 7, 8 and 9 all exhibited a mass spectrum with a base peak of 563.14 m/z units, corresponding to the pseudomolecular ion of an hexoside- pentoside derivative; absence of intermediate fragments lead us to assign the peaks as *C*-hexoside *C*-pentoside derivatives of apigenin (**Table 4**). Similarly, peak 13 was tentatively identified as apigenin *C*-hexoside, whilst peaks 11 and 12, both showing a base peak of 769.18 m/z units, were tentatively identified as apigenin *C*-hexoside *C*-hexoside *O*-glucuronide.

3.2. Phenolic content in wheat landraces

The determination of free phenolics in whole grains extracted by a hydroalcoholic solution (see experimental) was carried out through calibration curves obtained via HPLC/DAD triplicate injection of standard solutions. In **Table 5** the concentration of 13 phenolic markers and total free phenolics for the all investigated wheat genotypes is given.

Some of the phenolic markers identified in the free form, were quantitatively quite different among the genotypes studied. For example, coumarin (peak 6), ranging from 2.57 μg/g in Tumminia SG3 to 0.09 μg/g in Tripolino, and vanillic acid (peak 3) from 1.34 μg/g in Manto di Maria to 0.25 μg/g in Tumminia SG3. The apigenin *C*-hexoside-*C*-pentoside (peak 7) content was significantly different among wheat grains, recording values above 21 μg/g for Tumminia SG3 and about 6 μg/g for Trentino.

Luteolin di-*C*-hexoside (lucenin-2 isomer), present in all genotypes in low concentration (mean value 0.33 μg/g, excluding the extremes of the interval, Tumminia SG3 and Manto di Maria) was about 50-times more abundant in Tumminia SG3 (18.15 μg/g) than in other genotypes.

Free ferulic acid content resulted almost 3-times higher in Tumminia SG3 (5.81 μg/g) with respect to the mean value (1.84 μg/g).

Total phenolics concentration ranged from 65.65 μg/g of grain in Russello SG8 to 104.84 μg/g of grain in Scavuzza, and a mean value of 82.78 μg/g was recorded. Three landraces (Tumminia SG3, Tripolino, Scavuzza) showed a content higher than the average.

In general, genotype has been demonstrated to affect the phenolic content of wheat grains. Previous investigations reported on highly significant differences of polyphenol content among different wheat cultivars, suggesting the genotype-specificity of this characteristic [9, 14]. Moreover, the comparison of wheat cultivars grown at different locations showed that environmental and growing conditions may have a certain effect on the biosynthesis and accumulation of phenolic compounds [42].

With regard to the bound phenolic fraction subjected to alkaline hydrolysis, the main component is undoubtedly ferulic acid, as already observed by other authors [43], and confirmed by co-injection with the corresponding analytical standard; this metabolite is present ubiquitously in all the genotypes considered with a mean value of 543.20 μg/g. The landrace Ruscia showed the highest level of ferulic acid content (673.58 μg/g), while Scavuzza the lowest (375.13 μg/g) (**Table 5**).

3.3. Landraces statistical comparison

In order to discriminate among the studied wheat landraces, a statistical classification system was implemented using the data from the 15 analyzed chemical variables and the 138 measured morpho-colorimetric parameters. An overall percentage of correct identification of 100.0% was achieved, proving the peculiarity of the nine studied Sicilian wheat landraces and, on the other hand, the absolute effectiveness of the proposed method (**Table 6**).

Finally, in the evaluation of the parameters that more than other influenced the discrimination process of the studied landraces, none of the assessed variables chosen by the stepwise LDA highlighted particular statistical weight, proving that a high amount of quantitative information is necessary to distinguish and characterize botanical entities so heterogeneous, under chemical, phenotypical and genetic profile, such as landraces.

This work represent the first attempt of wheat landraces identification based on glume phenotypic characters, applying image analysis techniques, coupled with phenolic fingerprinting.

Peak #	Peak ID	Tumminia SG3	Russello SG8	Manto di Maria	Margherito	Ruscìa	Tripolino	Scavuzza	Trentino	Urrìa
1	Hydroxycinnamic acid	n.d.	0.06 ± 0.00	0.48 ± 0.02	n.d.	n.d.	0.20 ± 0.01	n.d.	0.05 ± 0.00	0.16 ± 0.01
2	Coumarin	1.51 ± 0.06	0.52 ± 0.02	1.06 ± 0.04	0.47 ± 0.02	1.28 ± 0.05	0.84 ± 0.03	1.76 ± 0.07	1.29 ± 0.05	1.39 ± 0.06
3	Vanillic acid	0.25 ± 0.01	0.86 ± 0.03	1.34 ± 0.05	0.51 ± 0.02	0.42 ± 0.02	0.58 ± 0.02	1.05 ± 0.04	0.94 ± 0.04	0.91 ± 0.04
4	Caffeic acid	0.11 ± 0.01	0.17 ± 0.01	0.41 ± 0.02	0.07 ± 0.00	0.63 ± 0.03	0.25 ± 0.01	1.29 ± 0.05	0.31 ± 0.01	1.28 ± 0.05
5	Luteolin di-*C*-hexoside	18.15 ± 0.72	0.41 ± 0.02	n.d.	0.20 ± 0.01	0.25 ± 0.01	1.01 ± 0.04	0.22 ± 0.01	0.07 ± 0.00	0.15 ± 0.01
6	Coumarin	2.57 ± 0.10	0.84 ± 0.03	1.28 ± 0.05	0.34 ± 0.01	0.87 ± 0.03	0.09 ± 0.00	0.17 ± 0.01	0.54 ± 0.02	0.32 ± 0.01
7	Apigenin *C*-hexoside-*C*-pentoside	21.46 ± 0.85	7.91 ± 0.31	12.15 ± 0.48	9.25 ± 0.36	11.67 ± 0.46	15.89 ± 0.63	18.81 ± 0.74	5.86 ± 0.23	11.80 ± 0.47
8	Apigenin *C*-hexoside-*C*-pentoside	3.76 ± 0.15	4.83 ± 0.19	3.09 ± 0.12	5.12 ± 0.20	3.97 ± 0.16	6.82 ± 0.27	5.90 ± 0.23	5.19 ± 0.20	4.84 ± 0.19
9	Apigenin *C*-hexoside-*C*-pentoside	10.38 ± 0.41	26.13 ± 1.03	22.78 ± 0.90	28.06 ± 1.11	24.30 ± 0.96	27.71 ± 1.09	29.42 ± 1.16	23.92 ± 0.94	35.32 ± 1.39
10	Ferulic acid	5.81 ± 0.23	0.88 ± 0.04	1.28 ± 0.05	1.04 ± 0.04	1.47 ± 0.06	0.90 ± 0.04	1.88 ± 0.07	1.53 ± 0.06	1.81 ± 0.07
11	Apigenin *C*-hexoside-*C*-hexoside *O*-glucuronide	7.22 ± 0.29	4.26 ± 0.17	10.81 ± 0.43	7.41 ± 0.29	9.40 ± 0.37	13.79 ± 0.54	16.92 ± 0.67	6.32 ± 0.25	5.48 ± 0.22
12	Apigenin *C*-hexoside-*C*-hexoside *O*-glucuronide	17.51 ± 0.69	14.98 ± 0.59	18.26 ± 0.72	22.66 ± 0.89	17.88 ± 0.70	18.67 ± 0.74	21.51 ± 0.85	18.21 ± 0.72	15.07 ± 0.60
13	Apigenin *C*-hexoside	4.87 ± 0.19	3.79 ± 0.15	5.36 ± 0.21	7.01 ± 0.28	3.11 ± 0.12	2.90 ± 0.11	5.91 ± 0.23	5.03 ± 0.20	7.76 ± 0.31
—	Total phenolics	93.61 ± 3.69	65.65 ± 2.59	78.30 ± 3.08	82.14 ± 3.23	75.26 ± 2.96	89.64 ± 3.53	104.84 ± 4.13	69.26 ± 2.73	86.30 ± 3.4
—	Ferulic acid in the bound phenolic fraction	522.30 ± 20.56	516.73 ± 20.34	603.90 ± 23.78	577.40 ± 22.73	673.58 ± 26.52	558.92 ± 22.01	328.67 ± 12.94	560.45 ± 22.07	546.83 ± 21.53

Table 5. Phenolics detected in the durum wheat landraces extracts.

	Margherito	Manto di Maria	Ruscia	Russello SG8	Trentino	Tripolino	Tumminia SG3	Urrìa	Total
Margherito	100.0% (192)	—	—	—	—	—	—	—	100.0% (192)
Manto di Maria	—	100.0% (192)	—	—	—	—	—	—	100.0% (192)
Ruscia	—	—	100.0% (192)	—	—	—	—	—	100.0% (192)
Russello SG8	—	—	—	100.0% (192)	—	—	—	—	100.0% (192)
Trentino	—	—	—	—	100.0% (282)	—	—	—	100.0% (282)
Tripolino	—	—	—	—	—	100.0% (192)	—	—	100.0% (192)
Tumminia SG3	—	—	—	—	—	—	100.0% (192)	—	100.0% (192)
Urrìa	—	—	—	—	—	—	—	100.0% (192)	100.0% (192)
Overall									100.0% (1626)

Percentages refer to the classification performance; in parentheses, the number of analyzed glumes.

Table 6. Percentages identification among the studied landraces.

The achieved results here discussed allowed to demonstrate the usefulness of this discrimination system for the identification and classification wheat landraces, notoriously very difficult to do. The technique here proposed, conveniently sustained by a conspicuous database, can be undoubtedly considered a helpful identification tool both for commercial varieties and for no genetically defined samples, such as populations or landraces.

Considering the heterogeneous nature of the wheat landrace samples used in this study, in order to validate these preliminary achievements, further trials will have to be conducted focusing on the collection of new data, enriching the database with new and accurate information, allowing to the system to give results more and more reliable.

Author details

Oscar Grillo[1]*, Marisol Lo Bianco[1], Sebastiano Blangiforti[1], Laura Siracusa[2], Gianfranco Venora[1] and Giuseppe Ruberto[2]

*Address all correspondence to: oscar.grillo.mail@gmail.com

1 Stazione Sperimentale di Granicoltura per la Sicilia, Caltagirone (CT), Italy

2 Italian National Research Council - Institute of Biomolecular Chemistry ICB, Catania, Italy

References

[1] FAO. Available from: http://www.fao.org/faostat/en/#home [Accessed: September 29, 2017]

[2] Cebola Lidon F, Almeida AS, Leitão AL, Silva MM, Pinheiro N, Maçãs B, Costa R. A synoptic overview of durum wheat production in the Mediterranean region and processing following the European Union requirements. Emirates Journal of Food and Agriculture. 2014;**26**(8):693-705

[3] Morancho J. Production et commercialization du blé dur dans le monde. In: Di Fonzo N, Kaan F, Nachit M, editors. Durum Wheat Improvement in Mediterranean Region. New Challenges. Vol. 40. Zaragoza: CIHEAM; 2000. pp. 19-34. Options Méditerranéennes A

[4] Grillo O, Blangiforti S, Venora G. Wheat landraces identification through glumes image analysis. Computers and Electronics in Agriculture. 2017;**141**:223-231

[5] Lombardo GM. Evoluzione varietale del frumento duro in Sicilia nel corso del XX secolo. In: Maimone G, editor. I Frumenti Siciliani–Un Patrimonio da Mantenere e Valorizzare. Catania: A cura della Stazione Sperimentale di Granicoltura per la Sicilia; 2004

[6] Spina A, Blangiforti S, Cambrea M, Gallo G, Licciardello S. Valorizzazione di germoplasma locale di frumento duro per la produzione di pani tipici. In: Proceedings of the VIII Convegno Nazionale "La biodiversità – una risorsa per sistemi multifunzionali; 21/23 April 2008; Lecce, Italy; 2008

[7] Sciacca F, Cambrea M, Licciardello S, Pesce A, Romano E, Spina A, Virzì N, Palumbo M. Evolution of durum wheat from Sicilian landraces to improved varieties. In: Porceddu E, Damania AB, Qualset CO, editors. Proceedings of the International Symposium on Genetics and Breeding of Durum Wheat. Bari: CIHEAM; 2014. pp. 139-145. (Options Méditerranéennes: Série A. Séminaires Méditerranéens; n. 110). In: International Symposium: Genetics and Breeding of Durum Wheat; Rome, Italy; May 27–30, 2013

[8] Gallo G, Porfiri O, Venora G, Lombardo GM, Blangiforti S. I Frumenti Siciliani: Patrimonio da Mantenere e Valorizzare. Catania: Maimone; 2004

[9] Pasqualone A, Delvecchio LN, Mangini G, Taranto F, Blanco A. Variability of total soluble phenolic compounds and antioxidant activity in a collection of tetraploid wheat. Agricultural and Food Science. 2014;**23**(4):307-316

[10] Migliorini P, Spagnolo S, Torri L, Arnoulet M, Lazzerini G, Ceccarelli S. Agronomic and quality characteristics of old, modern and mixture wheat varieties and landraces for organic bread chain in diverse environments of northern Italy. European Journal of Agronomy. 2016;**79**:131-141

[11] Lo Bianco M, Siracusa L, Dattilo S, Venora G, Ruberto G. Phenolic fingerprint of Sicilian modern varieties and old durum wheat landraces: A tool to assess with diversity. Cereal Chemistry; 2017;**94**:1045-1051

[12] Grillo O, Blangiforti S, Venora G. Traditional food with authentic raw material: The Sicilian wheat landraces. In: Proceedings of the "Bonding Tradition with Innovation: Successful Strategies in Food Chain Value; 12/13 September 2016; Olsztyn, Poland; 2016

[13] Venora G, Blangiforti S. In: Fate L, editor. I grani antichi siciliani. Manuale per il riconoscimento delle varietà locali dei frumenti siciliani. Ragusa, Italy. 2017. ISBN: 978-88-940976-6-5

[14] Dinelli G, Segura Carretero A, Di Silvestro R, Marotti I, Fu S, Benedettelli S, Ghiselli L, Fernández Gutiérrez A. Determination of phenolic compounds in modern and old varieties of durum wheat using liquid chromatography coupled with time-of-flight mass spectrometry. Journal of Chromatography. A. 2009;**1216**:7229-7240

[15] Acosta-Estrada BA, Gutiérrez-Uribe JA, Serna-Saldívar SO. Bound phenolics in foods, a review. Food Chemistry. 2014;**152**:46-55

[16] Mattila P, Pihlava JM, Hellstrom J. Contents of phenolic acids, alkyl- and alkenylresorcinols, and avenanthramides in commercial grain products. Journal of Agricultural and Food Chemistry. 2005;**53**:8290-8295

[17] Venora G, Grillo O, Saccone R. Quality assessment of durum wheat storage centers in Sicily: Evaluation of vitreous, starchy and shrunken kernels by image analysis system. Journal of Cereal Science. 2009;**49**:429-440

[18] Shahin MA, Symons SJ. Colour calibration of scanners for scanner independent grain grading. Cereal Chemistry. 2003;**80**:285-289

[19] Orrù M, Grillo O, Lovicu G, Venora G, Bacchetta G. Morphological characterization of *Vitis vinifera* L. seeds by image analysis and comparison with archaeological remains. Vegetation History and Archaeobotany. 2013;**22**:231-242

[20] Lo Bianco M, Grillo O, Sarigu M, Cremonini R, Venora G. Characterisation of Italian bean landraces (*Phaseolus vulgaris* L.) using seed image analysis and texture descriptors. Australian Journal of Crop Science. 2015;**9**(11):1022-1034

[21] Haralick RM, Shanmugam K, Its'hak D. Textural features for image classification. IEEE Transactions on Systems, Man, and Cybernetics. 1973;**3**:610-621

[22] Hâruta O. Elliptic Fourier analysis of crown shapes in Quercus petraea trees. Annals of Forest Research. 2011;**54**:99-117

[23] Grillo O, Draper D, Venora G, Martínez-Laborde JB. Seed image analysis and taxonomy of Diplotaxis DC. (*Brassicaceae Brassiceae*). Systematics and Biodiversity. 2012;**10**(1):57-70

[24] Duda R, Hart P, Stork D. Pattern Classification. 2nd ed. New York, USA: John Wiley & Sons; 2012

[25] Fisher RA. The use of multiple measurements in taxonomic problems. Annals of Eugenics. 1936;**7**(2):179-188

[26] Fisher RA. The precision of discriminant functions. Annals of Eugenics. 1940;**10**(4):422-429

[27] Fukunaga K. Introduction to Statistical Pattern Recognition. 2nd ed. San Diego, USA: Academic Press Professional, Inc; 1990

[28] Hastie T, Tibshirani R, Friedman J. In: Hastie T, Tibshirani R, Friedman J, editors. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2nd ed. New York: Springer; 2009. pp. 1-764

[29] Holden JE, Finch WH, Kelley K. A comparison of two-group classification methods. Educational and Psychological Measurement. 2011;**71**:870-901

[30] Kuhn M, Johnson K. In: Kuhn M, Johnson K, editors. Applied Predictive Modeling. New York: Springer; 2013. pp. 1-600

[31] Rencher AC, Christensen WF. Methods of Multivariate Analysis. 3rd ed. New York: John Wiley & Sons; 2012. pp. 1-800

[32] SPSS. Application Guide, SPSS Version 16.0. Chicago: SPSS Inc.; 2007

[33] Box GEP. A general distribution theory for a class of likelihood criteria. Biometrika. 1949; **36**:317-346

[34] Levene H. Robust tests for equality of variances. In: Olkin I, Ghurye SG, Hoeffding W, Madow WG, Mann HB, editors. Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling. Palo Alto, CA, USA: Stanford University Press; 1960. pp. 278-292

[35] Adom KK, Sorrells ME, Liu RH. Phytochemicals and antioxidant activity of milled fractions of different wheat varieties. Journal of Agricultural and Food Chemistry. 2005;**53**:2297-2306

[36] Cermak R. Effect of dietary flavonoids on pathways involved in drug metabolism. Expert Opinion on Drug Metabolism & Toxicology. 2008;**4**:17-35

[37] Adom KK, Liu RH. Antioxidant activity of grains. Journal of Agricultural and Food Chemistry. 2002;**50**:6182-6187

[38] Oomah BD, Mazza G. Flavonoids and antioxidative activities in buckwheat. Journal of Agricultural and Food Chemistry. 1996;**44**:1746-1750

[39] Kiprovski B, Mikulic-Petkovsek M, Slatnar A, Veberic R, Stampar F, Malencic D, Latkovic D. Comparison of phenolic profiles and antioxidant properties of European Fagopyrum esculentumcultivars. Food Chemistry. 2015;**185**:41-47

[40] Dinelli G, Segura-Carretero A, Di Silvestro R, Marotti I, Arráez-Román D, Benedettelli S, Ghiselli L, Fernadez-Gutierrez A. Profiles of phenolic compounds in modern and old common wheat varieties determined by liquid chromatography coupled with time-of-flight mass spectrometry. Journal of Chromatography. A. 2011;**1218**:7670-7681

[41] Leoncini E, Prata C, Malaguti M, Marotti I, Segura-Carretero A, Catizone P, Dinelli G, Hrelia S. Phytochemical profile and nutraceutical value of old and modern common wheat cultivars. PLoS One. 2012;**7**:e45997

[42] Martini D, Taddei F, Ciccoritti R, Pasquini M, Nicoletti I, Corradini D, D'Egidio MG. Variation of total antioxidant activity and of phenolic acid, total phenolics and yellow coloured pigments in durum wheat (*Triticum turgidum* L. var. durum) as a function of genotype, crop year and growing area. Journal of Cereal Science. 2015;**65**:175-185

[43] Adom KK, Sorrells ME, Liu RH. Phytochemical profiles and antioxidant activity of wheat varieties. Journal of Agricultural and Food Chemistry. 2003;**51**:7825-7834

Chapter 3

Rice Biodiversity in Cold Hill Zones of Kashmir Himalayas and Conservation of Its Landraces

Amjad M. Husaini and Sofi NR

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.74591

Abstract

Jammu & Kashmir is an agri-horticulture state (in India) where a large population is economically dependent to agriculture and horticulture, directly or indirectly for livelihood, food and nutritional security. Rice, the staple food of majority population, is cultivated in diverse agro-ecological situations extending from subtropical area (<1000 m amsl) of Jammu, through temperate valley to cold high altitudes regions (1650–2400 m amsl) of Kashmir, and therefore rice biodiversity is rich. Some of the landraces fall into *indica* type and thrive well in temperate regions of valley, while others fall into *japonica* type and are prevalent in hilly areas. Cold tolerance is a special feature in most of these peculiar landraces, which are different from rest of the country. However, with the advent of High Yielding Varieties the local biodiversity got neglected and remained confined to seed banks. With new emerging challenges like climate change, population explosion, limited land and water resources, demand for organic products, local landraces have assumed tremendous importance, either for direct exploitation or indirect use. In this chapter, we attempt to bring out information about these landraces in a comprehensive manner and discuss the issues pertaining to their conservation, utilization, cultivation and revival through approaches like participatory plant breeding, participatory varietal selection or plant biotechnology.

Keywords: *Oryza sativa*, climate change, high yielding varieties, cold tolerance, characterization, India

1. Introduction

The state of Jammu and Kashmir comprises the extreme western part of the Himalayas (32.44°N and 74.54°E), with altitude ranging from 200 to 7000 m amsl. The valley of Kashmir

(temperate zone) is approximately 120 km long and 32 km wide with an altitude range of 1524–2286 m amsl. Annual precipitation varies from 700 to 1500 mm. The temperature remains generally low, varying from −10°C during winter to 30°C during summer, with a yearly average of around 13°C. Generally, soil in rice growing areas is clay loam with a neutral pH. The economy of Jammu and Kashmir is predominantly agrarian, rice being the staple food and the most important food crop occupying an area of almost 140 thousand hectares. There are many well-known landraces of rice in Kashmir, noted for their unique qualities, peculiar taste and texture after cooking, early maturity and cold tolerance. When Green Revolution made its impact in India, the state of Jammu and Kashmir did not lag behind. Some introductions from China like China-1039, China 1007 and locally bred varieties like K-39, Jehlum and Chenab spread to every nook and corner of the state through vigorous political and infrastructural support by way of different Govt. sponsored schemes. As a result, faster replacement of indigenous low yielding landraces by modern High Yielding Varieties (HYVs) took place in the sixties and the entire rice area of Kashmir valley got covered under a few varieties.

China 1039 and China 1007 were the most popular varieties during 60s and 70s; K39 was the dominant rice variety from 80s to mid-1990s; while Jhelum took over as the most popular variety for the decade beginning from mid 90s. This variety has an excellent cooking quality and is most preferred even today, but due to its susceptibility to blast there has been a decline in its area. In order to save farmer's interests a blast tolerant variety Shalimar Rice-1 was released in 2005 by the State Varietal Release Committee for the valley basin irrigated areas of Kashmir. This was followed by release of Shalimar Rice-2, Shalimar Rice-3, Shalimar Rice-4, all of which are blast tolerant and have good cooking quality (**Table 1**).

Variety and year of release	Cross combination	Salient feature
Plains (upto 1650 m amsl)		
China 1039 (1955)	Introduction	Yield potential 50–55 q/ha, cold tolerant, lodging and shattering susceptible, maturity 136–140 days (1650 m amsl). Moderately susceptible to blast, recommended for cultivation upto an altitude of 1650 m amsl
China 1007 (1956)	Introduction	Yield potential 50–60 q/ha, lodging and shattering resistant, cold tolerant at early stages of growth, maturity 145–147 days, special attribute: resistant to blast and suitable for low lying areas.
China 988 (1956)	Introduction	Lodging and shattering resistant, cold tolerant at early stages of growth, maturity 147–150 days, moderately tolerant to blast, suitable for low lying areas.
China 972 (1956)	Introduction	Lodging and shattering resistant, cold tolerant at early stages of growth, maturity 147–152 days, moderately tolerant to blast, suitable for low lying areas
K-60 (1962)	China 47/RIKUU-132	Resistant to shattering and lodging, resistant to cold, maturity 140–145 days, blast resistant
K-65 (1966)	NORIN-8/China-47	Easy threshability, lodging susceptible, low head rice recovery, maturity 140–145 days, suitable for low lying areas
K-39 (1978)	China 1039/IR-580	Yield potential 58–62 q/ha, high yielding, recommended upto 1650 m amsl, mod: susceptible to blast, resistant to lodging, maturity 140–145 days. High head rice recovery (%age)

Variety and year of release	Cross combination	Salient feature
Chenab (1996)	K-21/IR-2053	Yield potential 60–65 q/ha, coarse grained, moderately tolerant to blast, cold tolerant, possess complete synchronous flowering recommended upto 1650 m amsl, better cooking quality, maturity 138–140 days.
Jehlum (1996)	JAKKOKU/IET-1444	Yield potential 60–65 q/ha, high yielding, better cooking quality, greater tolerance to cold, moderately susceptible to blast. Recommended upto 1650 m amsl, maturity 138–140 days.
Shalimar Rice-1 (2005)	China 1007/IET 1444	Yield potential 65–70 q/ha, high yielding, better cooking quality, greater tolerance to cold, highly resistant to blast. Recommended upto 1650 m amsl, maturity 142–145 days, medium bold grain size.
Shalimar Rice-2 (2012)	VL Dhan 221/K 39	High yielding (80–85 q/ha) moderately blast resistant *indica* rice variety having long panicle, more number of grains/panicles strong stem, resistant to lodging, high phenotypic acceptability takes 139–143 days to mature. Recommended for plains of the valley (upto 1650 m).
Shalimar Rice-3 (2012)	IR32429-47-3-2-2/K 438	High yielding (80–85 q/ha), early maturing, cold tolerant *indica* variety having moderate resistance to blast, erect plant type, easy thresh ability, erect flag leaf, recommended for cultivation in plains of the valley (upto 1650 m)
Shalimar Rice-4 (2016)	Jehlum/84017-IR745-12-1	High yielding (75–80 q/ha), early maturing, cold tolerant *indica* variety having resistance to blast, easy thresh ability, recommended for cultivation in plains of the valley (upto 1650 m)
Higher belts (upto 1800–2100 m amsl)		
Shenei (1967)	Introduction	Yield potential (30–35 q/ha), moderately tolerant to blast, cold tolerant, maturity 130–135 days, recommended ecology 1850–2200 m amsl
China 971 (1967)	Introduction	Yield potential (30–35 q/ha), moderately tolerant to blast, cold tolerant, maturity 130–135 days, recommended ecology 1850–2200 m amsl
Barkat (1974)	Shenei/China 971	Yield potential (38–40 q/ha), cold tolerant, high head rice recovery, susceptible to blast, maturity 140–145 days. Suitable for cultivation under mid altitude condition 1650–1850; universal donor for cold tolerance
K-332 (1982)	Shenei/Norin 11	Yield potential 40–45 q/ha, *japonica* type, cold tolerant, high head rice recovery, moderately resistant to blast, maturity 130–140 days.
Kohsar (2002)	Shenei/GINMASARI	Yield potential 45–50 q/ha, *japonica* type, cold tolerance particularly at seedling stage, high degree of resistance to blast, suitable for high hills of the valley (1800–2150 m). The variety has high head rice recovery, easy threshability, sticky type of grain, short bold grains, and matures in 138–140 days
Subtropical area (<1000 m amsl)		
Ranbir Basmati (1996)	Selection from Basmati 370	Suited to non-basmati growing areas, yield potential 20–30 q/ha; altitudinal tolerance upto 1000 m.

Table 1. The most popular rice varieties and their salient features.

Narrow genetic base increases the vulnerability of rice production system to biotic/abiotic stresses and has resulted in yield stagnation. Among the biotic stresses rice blast is the most damaging disease in Kashmir, while low temperature ranks first among the abiotic factors limiting rice production [1]. Further, climate change is also a potential threat to rice production owing to erratic rise/fall in temperature, and change in the dynamics of pests and diseases. These nutritional and food security related issues demand the immediate attention of rice breeders and biotechnologists. Locally available rice germplasm is a valuable repository of traits which could help address most of these concerns effectively. Novel gene pools could be generated from the characterized germplasm and donors for yield, quality and resistance to biotic and abiotic stresses identified. Besides, there are some quality rice varieties which when cultivated properly can boost the farm income substantially.

2. Collection, conservation and documentation of rice biodiversity - How and why?

The advent of Green Revolution had an overwhelming impact on rice production in Kashmir, as was elsewhere in the country. In 1966, the International Rice Research Institute (IRRI) released the first high yielding rice variety in the Philippines. In the subsequent decade Rice Breeders at Rice Research Station, Khudwani introduced Chinese high yielding varieties viz., China 1039, China 1007, China 988 and China 972 in selected pockets of Kashmir valley. However, in a matter of few years these varieties and the rice varieties developed by the station through crossing programme viz., K-60, K-65 and K-39 (most popular) almost completely replaced hundreds of the traditional rice landraces previously cultivated by the farmers [2]. The new rice varieties had a higher harvest index (grain/straw ratio), and the benefit of delivering significantly higher yields when combined with accompanying management practices, including irrigation, weedicide and fertilizer application. These high yielding varieties spread in favorable environments, where the natural and infrastructural setting allowed for such practices. In unfavorable environments, in which irrigation and mechanization were not possible or agrochemicals were not available, the cultivation of the traditional landraces persisted. These marginal areas (upland environments, high altitude belts, very cold areas, etc.) could serve as a repository of indigenous rice germplasm/landraces. One such niche area is the 'Sagam' belt of Anantnag district, which continues to grow aromatic landraces *Kamad* and *Mushk budji*.

Waking up to the loss of original indigenous varieties of agri-horticulture crops a 'National Agricultural Technology Project on Sustainable Management of Plant Biodiversity (1995–2005)' was undertaken by scientists of SKUAST-K. Under this project a total 1911 germplasm accessions, which were under cultivation before the introduction of improved or imported varieties, were collected. The collected biodiversity included 742 accessions in cereals, 38 in pseudo cereals, 28 in millets, 71 in oilseeds in pulses, 377 in vegetable crops, 21 in spices and condiments, 13 in fodder crops, 204 medicinal and aromatic plants, 55 in fruit crops and 4 in

others. The university deposited 1447 germplasm specimens with the National Gene Bank for storage and handed over 557 germplasm accessions to the Germplasm Handling Unit of the National Bureau of Plant Genetic Resources (NBPGR), New Delhi for long term storage. For collection of rice germplasm, expeditions were undertaken by a team of rice experts and scientists to collect the rare germplasm from different areas, among which Tral (Wagad, Shikargah), Pahalgam (Batekoot, Puhri-pajal, Khayar Hapath-nard), Shopian (Balpora, Ganapora, Shadimarg, Kalampora), Badgam (Khan sahab, Chadoora), Kupwara (Nagri malpora), Rajouri and Uri were the prominent ones. Nearly 100 existing landraces of rice were collected and assigned accession numbers, for conservation and maintenance in the seed bank at Rice Research & Regional Station, Khudwani.

Currently most of the rice fields in Jammu & Kashmir are occupied by merely a small number of high yielding rice varieties, of which K-39, Jehlum, Chenab, Shalimar Rice 1, Rambir Basmati are the prominent ones. This trend is no different than rest of India and other Asian regions. In India the most widely grown rice varieties are Swarna, Samba Mahsuri, Sona Masuri, Jaya, Ratna, etc. In Philippines almost half of the rice area is devoted to four of the most widespread HYVs, Cambodia one single IRRI variety (IR66) accounts for around 90% of the rice area, and in Pakistan only four HYVs are planted on 99% of the country's rice fields. This illustrates the immense 'genetic erosion' that has occurred in farmers' fields since the onset of the Green Revolution.

At present, 580 germplasm accessions, indigenous and exotic, are being maintained as three row material at Mountain Research Centre for Field Crops (erstwhile Rice Research & Regional Station), Khudwani. Almost all these accessions have been characterized for different agro-morphological traits, diseases resistance score (leaf blast, panicle blast) and aroma in a period of 5 years (2008–2013). A systematic and exhaustive morphological and molecular characterization of the 'speciality' rice types has revealed some interesting results [3]. DNA fingerprints of 16 pigmented and aromatic genotypes, mostly of Western Himalayan region (32° 44′–35° 2′ N and 74° 28′–75° 48′ E at altitude range of 1540–2200 mamsl in Kashmir; and 31° 17′ N and 76° 51′ E at an altitude of 1190 mamsl in Himachal Pradesh (India) have been developed and these were evaluated for genetic diversity using SSR markers. Various population parameters viz. range, mean, skewness and kurtosis from the data generated have shown wide range of variability (unpublished). This has enormous implications for future studies on gene/allele mining, as germplasm in Western Himalayas could serve as a vital resource of genetic repository for marginal areas.

The germplasm bank at MRCFC Khudwani maintains some well adapted exotic introductions (including *Goshigon, Chengshi, Koshihikari, Cheolwean 32, Kunusa rex*) and some indigenous introductions (like *Heera, Dullar, Bahrigu dhan*) from other rice growing regions of India. It also maintains some international blast differentials (like C101-LAC, C101-A51, C101-PKT, C105-TTP-4-1-23, RIL 10, RIL 29, NP 125, USEN, Tadukan, Shia-tai-tsau, HR 12, CO 39 etc.) which are being used in the development of blast tolerant rice genotypes [4] (**Figure 1**).

Figure 1. Biodiversity of rice, as depicted through a range in morphological variability.

3. Why are landraces important?

3.1. Genetic and agronomic value

The local varieties are a repository of diverse genes, some of which are of practical importance in the changing socio-economic as well as edaphic/climatic conditions. While all high yielding varieties in Kashmir are white, mostly with medium bold grains, local rice varieties often exhibit tremendous morphological diversity. The color of the outer layer (pericarp) can range from black/purple to red and brownish to white. The grain weight of landraces in Kashmir, as characterized by the thousand-kernel weight, varies between 15 and 32 g, while the HYVs varies only between 22 and 34 g. In fact, more than 90% out of 92 identified landraces cultivated in the area are no longer under cultivation (**Figure 2**).

The findings of a recent study, conducted somewhere else, are presented here to illustrate the importance of conserving and exploring the 'genetic value' of landraces. A major quantitative trait locus for phosphorus-deficiency tolerance, *Pup1*, was identified almost a decade back in the traditional *aus*-type Indian rice variety 'Kasalath' [5, 6]. The locus was sequenced and *Pup1*-specific protein kinase gene located, which was recently named as phosphorus-starvation tolerance 1 (*PSTOL1*) gene. It is important to emphasize that this gene is absent from the rice reference genome and other phosphorus-starvation-intolerant modern varieties. The overexpression of *PSTOL1* in such varieties significantly enhances early root growth and grain yield in phosphorus-deficient soil by enabling plants to acquire more phosphorus and other nutrients [7]. The absence of genes like *PSTOL1* and submergence-tolerance gene *SUB1A* from modern rice varieties underlines the value of traditional rice germplasm.

3.2. Socio-economic and cultural value

Rice is not only the dominant staple food, but also an integral part of culture in rural Kashmir. Some landraces have a special cultural value, for example scented landraces *Mushk budji* and *Kamad* have traditionally been served on particular occasions like marriage ceremonies and festivals. These varieties used to be sold as 'food for the royal families' in the local markets of

http://dx.doi.org/10.5772/intechopen.74591

Figure 2. Local biodiversity of rice as depicted by the difference of panicle type, grain type, grain color, awn type etc.

Srinagar during the times of Maharajas, which could only be afforded by the affluent and rich families owing to their 4–5 times higher market price. These landraces are highly valued even in present times and are served to distinguished visitors and dignitaries. Some red rice types called '*Zag*' are used for the preparation of snacks like '*Vazul bate*' for pregnant ladies owing to their higher nutritive value, while some are preferred for preparation of munchies like '*Bate laaye*', '*Mur-murei*' and '*Chewrei*'. The varieties' names in the local language often reflect the rice's appearance (*Kaw kreer, Laer beoul, Nika katwor, Shala kew*), smell (*Mushk budji, Mushkandi*), color (*Zag, Safed Khuch, Bari safed, Kaw kreer, Khuch, Sig safed, Safeed braz, Safed cheena*), cultivator's name (*Aziz beoul, Begum, Qadir baig, Rehman bhatti, Noormiree*), etc. Many varieties are characterized by a very specific taste, and their seeds are exchanged among neighbors/relatives, or are gifted for eating purposes in the form of roasted rice, locally called '*Bayel tamul*'. The long-grained basmati type varieties are cooked as '*Kashmiri pulao*' and served along with dry fruit and raisins in the famous Kashmiri cuisine locally called as '*Wazwan*'. The loss of biodiversity therefore would also imply a fading rural culture (**Figure 3**).

3.3. Health value

The taste, texture and organoleptic value of rice depend on many factors. Rice is a starchy staple food which provides almost 90% of dietary energy to an average Kashmiri. It is a part of their culture, and even in parties, conventions, get-togethers rice is an indispensible component of their dishes. No party can ever be even imagined without '*bate*' (cooked rice)! Starch properties of the traditional varieties are therefore an important factor determining the choice of varieties for different occasions. Varieties differ widely in the proportion of starch composition/fractions of amylose (consisting of linearly linked glucose molecules) and amylopectin (composed of glucose molecules with branched links). The waxy rice varieties consist of amylopectin only, and have a sticky texture because these absorb less water upon cooking. On the other hand, rice landraces with more than 25% amylose content absorb more water and have a fluffy texture after cooking. Higher proportion of amylose is further associated with low glycemic index (slow starch digestion) and longer feeling of satiation after ingestion. On the other hand fast

Figure 3. Agricultural biodiversity shapes food culture (a) Kashmiri salted rice *Taeher*, generally served at *Sufi dargahs* (b) Sweet *Kashmiri pulao* served with *Wazwan* (c) *Bate laaye* sold by local vendors as snacks for children (d) Cooked pigmented rice by a *Gujjar* family.

digestion can cause a sensation of hunger shortly after the ingestion of waxy rice and is considered unfavorable because its long term consumption can induce type II diabetes (i.e. non-insulin dependent diabetes) in adults. There is a prompt and pronounced increase of the blood glucose level (= high glycemic index) after the ingestion of rice (especially waxy rice), similar to that caused by white bread or pure glucose. Therefore in some parts of Jammu & Kashmir such waxy varieties are used for the preparation of sweet/salty snacks (*Meethay chaval*, *Taeher*) by cooking them and then cooling before consumption. Some varieties are preferred for the preparation of bread (*Tomul tschot*). Cooling after cooking has been shown to substantially slow down starch digestion due to physiochemical changes in the starch structure (retrogradation) [8]. Besides starch composition many other factors specific to these varieties viz., physiochemical starch structure or the size of the starch granules, also contribute to delayed starch digestion.

3.4. Grain nutritional value

Another aspect that makes rice landraces attractive is the wide range in palatability, texture, and nutritional value depending upon their genetic makeup. With the advent of high yielding varieties and large scale adoption of the rice milling technology the spread of vitamin

B-deficiency (*beri-beri*) has become more common, due to the loss of vitamins through the disposal of rice bran. Earlier farmers used to process their rice manually and remove only the fibrous hull. Rice was then consumed as 'brown rice' that included the bran layer. The 'nutritional diversity' of rice bran added value to the rice bowl, because bran is more diverse in its composition and contains protein, lipids, fiber, vitamins, and minerals. The major vitamins present in the rice bran are vitamin E (α-tocopherol) and the B-vitamins (thiamin, riboflavin, and niacin) while the major minerals are phosphorus, potassium, and magnesium. It is a general practice nowadays to process rice grain by additionally removing the bran layer from the endosperm to obtain milled rice. This is done due to the consumers' preferences but it takes away the nutritive 'value' from nutritionally rich landraces, making them at par with other landraces or varieties. A study on Philippine rice landraces has suggested that their average lipid content is significantly higher than that of the HYVs collected from the same area. The lipid content of HYVs (brown rice) ranged between 2.0 and 2.1%, while its average value for the landraces was 2.3%, with some individual varieties having 3.2% lipid content. Similar study was undertaken on landraces of Kashmir [9]. A highly aromatic landrace of Kashmir Himalayas 'Mushk budji', which is grown in mid altitude cold regions, showed maximum total protein (8.86%) content as well as highest fiber content (3.31%). However, it recorded the lowest starch content of 70.45% while a popular high yielding variety grown in the plains of Kashmir valley 'Shalimar rice-1' recorded the highest (79.36%) starch content as well as maximum amylose percentage (24.34%). Total Phenol content showed a wide range from 4.87 to 1.02 mg/g, with maximum in a pigmented rice genotype 'Purple rice' while lowest in 'Jhelum', a popular high yielding rice variety. Besides, purple rice also had maximum total anthocyanin (9995.34 μg/g) content, while lowest (5943.14 μg/g) was recorded in Jhelum. Similarly, total carotenoids too varied in a wide range, with 'Khuch' recording almost 10 times (0.022 μg/g) than the lowest (0.002 μg/g) in Shalimar Rice-1 and Jhelum. These results indicate that scented and pigmented rice genotypes of Kashmir Himalayan region are of better nutritional quality than the conventional high yielding varieties and could be promoted as 'Specialty' rice for better economic returns to the farmers [9].

Similarly, higher levels of β-carotene are generally found in pigmented (colored) rice varieties [10]. Such landraces of Kashmir with colored pericarp viz., brown (*Zag, Khuch, Khuch niver, Mir zag, Niver, Mir sagi, GS 10, GS 44, GS 51, GS 52, GS 80, GS 83, GS 224, GS 268, GS 289, GS 484, GS 501*), purple (*Purple rice*), black (*Zager, Kaw kreer*) are only cultivated in remote areas maintaining a higher diversity of rice genotypes. Carotenoids being fat soluble, such varieties generally possess higher lipid content too. From a nutritional point of view this is favorable because it ensures the supply of unsaturated fatty acids necessary for the transformation of β-carotene into vitamin A. As with vitamins, minerals like iron and zinc are chiefly located in the bran of the rice grain. Generally, iron and zinc contents tend to be higher in aromatic and colored (red and black) rice varieties than in colorless varieties and ordinary HYVs [11, 12] (**Figure 4**).

3.5. Disease management value

Magnaporthe grisea, the causal agent of blast disease, is a serious pathogen of graminaceous species and is best known as the causal agent of the rice blast disease. The disease is a serious

Figure 4. Pigmented rice types (a) *Purple rice* (b) *Niver* (c) *Zag* (d) *Zager*

production constraint for rice in the northwestern Himalayan region of India, comprising the states of Jammu and Kashmir, Uttaranchal and Himachal Pradesh [13]. The disease is endemic to most rice growing areas of Jammu and Kashmir due to prevailing blast-conducive environmental conditions during the crop season. Although, chemical control of the disease is available, it is economically expensive for resource poor farmers and is environmentally undesirable. Since host resistance offers cost effective and eco-friendly method for disease management, a study in China has demonstrated how diversification of rice varieties is able to significantly reduce rice blast infestation [14]. Genetic diversity is a defense against diseases and pests owing to the presence of diverse genes and genetic components, which give selective advantage to these varieties under heavy selection pressure imposed by diseases. The rice blast disease is one of the major diseases in Kashmir, which exists as a combination of pathogenic races. Therefore, rice resistance genes often remain effective only for a few years of agricultural production, before succumbing to new pathogenic races. In the Chinese study diversification as a pest management strategy was so successful that farmers were able to abandon the use of fungicides in just 2 years. Similar results were obtained in Philippines, where more than 50 rice landraces were cultivated in two upland municipalities, and there were no reports of any rice pest infestation, except rats and birds [10]. A preliminary study with this purpose was conducted at MRCFC, Khudwani wherein rice germplasm was screened for blast tolerance under temperate conditions of Kashmir valley. Evaluation was

http://dx.doi.org/10.5772/intechopen.74591

Figure 5. Popular landraces of rice in Kashmir (central two are scented, and outer two are red rice types).

made by the IRRI's standard evaluation system of rice on 0–9 scale. The promising test entries with tolerance to blast can increase prospects of producing rice organically.

4. Initiatives in biodiversity management of rice

4.1. Purification of scented local landraces

Kashmir is well known for the cultivation of some local scented landraces grown in different agroecological niches and maintained by farmers since time immemorial. Of the aromatic cultigens, *Mushk budji* and *Kamad* are in great demand due to their excellent cooking and eating qualities [15, 16]. Earlier, in absence of blast infestation these varieties would yield grain of 3.0–3.5 t/ha in farmers field. However, the yields declined drastically due to their

vulnerability to virulent races of rice blast pathogen (*Magnaporthe grisea*). Therefore, these varieties were thrown out of cultivation, except in some small isolated pockets where these varieties are being grown as complex mixtures. These varieties fetch 5–6 times higher price in local markets than the commonly grown high yielding varieties. In view of this an attempt was made at MRCFC, Khudwani to purify the farmers bulks and identify superior pure lines of *Mushk budji* and *Kamad* with good quality and agronomic attributes to boost the income of farmers. Fifty-five single plant selections of *Mushk budji* and 64 single plant selections of *Kamad* were chosen from the farmers' fields in *kharif* 2008 and threshed separately to raise the head to row progeny during *kharif* 2009. Based on yield, disease and quality traits, 18 promising progenies of *Mushk budji* and *Kamad* were selected for further evaluation during *kharif* 2010. Out of 16 *Mushk budji* pure line selections made from 55 single plant selections collected from different areas of Kashmir valley, three selections were finally found to be at par with each other in terms of grain yield and aroma and were chosen for seed multiplication, while one *Kamad* line was selected for its seed multiplication. During *kharif* 2011 these lines were multiplied to produce nucleus seed, and evaluated for agro-physiological characters and distinguishing morphological descriptors. During *kharif* 2012 these were being multiplied for distribution among farmers. The *Mushk budji* seed was distributed at MRCFC Khudwani among the farmers of Sagam, Kokernag, and Batengu; while in districts of Budgam, Kulgam, and Pulwama it was distributed to farmers through the respective Krishi Vigyan Kendras (KVKs). However, not much success could be achieved in identifying blast tolerant genotypes among these accessions and blast resistance remained a challenge. A programme on Marker Assisted introgression of blast resistance genes (pita, pi54) in *Mushk budji* was introduced recently (2014) with some progress already achieved (personal communication).

4.2. Identification and evaluation of local red rice type land races

As discussed earlier, a fruitful outcome of germplasm characterization was that some of the promising nutritive red rice types having colored pericarp were identified for evaluation, molecular intervention and biofortification. Preliminary evaluation of the 13 major red rice types viz., *Zag*, *Kupwara Zag*, *Uri Zag*, *Mir Zag*, *Khuch*, *Mir Sagi* and some other accessions (*GS 51*, *GS 80*, *GS 83*, *GS 224*, *GS 268*, *GS 289*, *GS 484*) was also completed for yield and agronomic traits (2011–2012). The biochemical and mineral contents (Iron & Zinc) of these landraces are being determined using modern techniques (**Figures 5** and **6**).

Figure 6. Evaluation, purification and selection of best genotype from *Zag*, *Mushk budji* and *Kamad* at MRCFC, Khudwani.

http://dx.doi.org/10.5772/intechopen.74591

5. The way forward: What do we need to do?

Rice is cultivated in different agro-ecological regions of J&K, comprising sub-tropical area >1000 m amsl of Jammu region; mid altitude areas (1000 to <1650 m amsl) of Poonch, Rajouri and Doda districts; temperate or valley basin area (1650 to 1900 m amsl) and cold high altitude areas (>1950 to 2400 m amsl) of mountainous terrain of Kashmir. Nearly 10–12% of total rice cultivated area of the valley falls in the higher altitude region. The population of this region lives in harsher climate and difficult hilly/mountainous terrain. The farmers in this region still grow old non-descript varieties/cultivars which have poor yield potential and are susceptible to Paddy blast. Low temperature and very short summer months reduce yield and affect nutrient availability/mobilization rate from the soil. These are a big impediment to the introduction of varieties from mainland India, most of which thrive well under subtropical conditions [17]. Attempts were made in the past to develop high yielding cold tolerant rice varieties like Barkat, K332, Kohsar etc. [1, 15, 16]. Similarly, an innovative programme on development of a hybrid rice was started at SKUAST (K), Khudwani, after procuring cytoplasmic male sterile (CMS) lines and their maintainers from various institutions such as the International Rice Research Institute (Philippines), Directorate of Rice Research (Hyderabad), Central Rice Research Institute (Cuttack), and Punjab Agricultural University (Ludhiana). Studies on the performance of these CMS and maintainer lines (of tropical and subtropical regions) for various agro-morphological traits under temperate conditions of Kashmir revealed that these lines, because of poor phenotypic acceptability, cannot be used to develop experimental hybrids [18]. In addition, a good number of hybrids released in India were also evaluated along with their parental lines under temperate conditions and were found to be not suitable for cultivation in such an environment [19]. Thus, efforts were made to develop new CMS lines in the background of agronomically adapted and popular varieties of the region in order to fully exploit this technology. This led to the development of two cold-tolerant CMS lines suitable for Kashmir Himalayas [19]. These CMS lines were then successfully employed for development of medium-bold rice hybrids with good grain quality for Kashmir Himalayas [20].

The challenges to nutritional and food security need to be addressed. At the same time weightage of an equal measure needs to be given towards the conservation and utilization of rice genetic diversity. This can be done using a multi-pronged strategy involving the following:

5.1. Conservation and rejuvenation

Most of locally adapted aromatic and non-aromatic rice genotypes have evolved as a consequence of natural and human selection, and are highly adapted to specific ecological niches carrying the genes for adaptability, early maturity and cold tolerance. These genotypes, having evolved under specific ecological niches of Kashmir carry combined adaptive traits for such difficult ecological regime, and are not much amenable to high input agriculture. Therefore, these need to conserved/maintained, and periodically cultivated for evaluation under resource poor and marginal conditions of far flung areas.

Conservation should aim to preserve all of the genetic variation that is available in a population, and is best insured by seed rejuvenation in an environment as similar as the native habitat of the population. On the contrary, evaluation should aim at the identification and isolation of obviously useful genotypes, involving selection and purification in the process. Splitting of original stocks into pure lines, and repeated seed increase cycles, are adjuncts of evaluation and utilization. Conservation maintains the germplasm inputs, while evaluation and utilization make these conservation efforts worthwhile. The dynamics of any crop needs to be understood prior to initiating any on-farm conservation and utilization programme. The conventional approach so far has been to transfer technologies generated elsewhere to the farmers. But such an approach has not only been less efficient in the adoption of the technologies by the farmers, but has also led to replacement or erosion of local genetic resources. This raises the question of how to generate relevant and farmer preferred technologies, while attempting to conserve, manage and utilize the rice diversity at community level.

Can participatory plant breeding be a guiding principle for redesigning/revitalizing the landraces to suit modern times? And can participatory varietal selection help in conserving the landraces in pure form wherever they fit farmers' criteria and consumers' preferences.

5.2. Participatory plant breeding and participatory varietal selection

There is a serious concern among the farmers, scientists, policymakers and environmentalists regarding continuous erosion of genetic biodiversity. When uniformity becomes the cause of genetic vulnerability, genetic diversity is the only insurance against it. In the era of climate change complex biotic and abiotic stresses shall cause the high yielding varieties succumb and lose their comparative economic advantage. The crisis is further expected to deepen in scope as well as intensity, when pressure due to population increase and urbanization causes shrinkage in rice area, as well as shift in rice cultivation towards newer areas with untested soil types, different climatic patterns, and new pathogenic interactions. Under these circumstances, can local landraces be a part of the solution?

Although many landraces are preserved by breeders in seed banks, farmers do not have access to these for cultivation. Moreover, preservation in seed banks does not allow these landraces to adapt to changing environmental settings and changing agricultural practices. In order to address both these issues and meet the challenges (discussed in above paragraph), systematic screening of the desired rice types by local farmers of an area, through participatory varietal selection (PVS), can lead to useful site-specific introductions [21]. In this perspective it may be renamed as participatory varietal dispersal (PVD), and can lead to high seed replacement rate. PVS will also generate wealth of information about the (a) farmer preferred traits, and (b) trade-off between traits, which can then be followed by participatory plant breeding (PPB) for incorporating such useful traits/genes into the existing varieties and landraces. This would ensure genetic diversity on ground, and guarantee sustained levels of high productivity (**Figure 7**).

Figure 7. Participatory varietal selection (PVS) could be a good strategy for evaluation, dispersion and distribution of local landraces.

5.3. Value addition and promotion of domestic/international trade

Many landraces, mostly of *japonica* background, are known for their superior agro-morphological and quality traits like taste, aroma, texture, colored aleurone, early maturity and cold tolerance. Notable among them are *Mushk budji*, *Kamad*, *Nun beoul*, *Laer beoul*, *Zag* and *Khuch*. While their collection, characterization and conservation in hot spots of the valley is important from genetic point of view, their purification and commercialization is of special importance from the farmers' perspective of livelihood security. Therefore, there is an urgent need to add value to these landraces by enriching them with minerals and micronutrients like iron, zinc etc., and promote their commercialization as 'speciality rices'.

Further, scientific studies on aromatic rice of Kashmir have not given ample attention to their domestic trade in the past, despite the fact that scented rice varieties have competitive international price and the state can earn foreign exchange from them. However, in the present decade Ministry of Agriculture, Government of J&K has keenly supported the promotion and revival of these varieties. The Ministry of Commerce, Government of India, has permitted export of *Sarveshwar* Basmati of Jammu & Kashmir to USA, Jordan and Saudi Arabia so that the farmers of the state could get benefited [2]. In order to further the cause it is important for the farmers, scientists, policymakers and environmentalists to understand the pattern of domestic trade of aromatic rices at the micro level to address the basic issues of promoting their cultivation, production and export. The marketing and price-spread patterns of aromatic rices of Jammu and Kashmir like *Mushk budji*, *Kamad*, Ranbir basmati have to be examined and policy interventions suggested with regard to their production and trade in the state. A beginning in this direction was made in the year 2012 by setting up demonstrative plots of *Mushk budji* and *Kamad* on an area of 0.8 ha in Sagam, and cultivated by local farmers under the guidance of SKUAST-K. The revival programme was a huge success. The landrace *Mushk budji* was released by Hon'ble Governor J&K at 3rd Agricultural Science Congress of J&K (2014) under Public Private partnership mode by SKUAST-K and Sarveshwar Overseas Ltd. The farmers that were involved in the programme bagged the Plant Genome Savior Community Award (2016) from Protection of Plant Varieties and Farmers Rights Authority, Ministry of Agriculture & Farmers Welfare, Govt. of India (**Figure 8**).

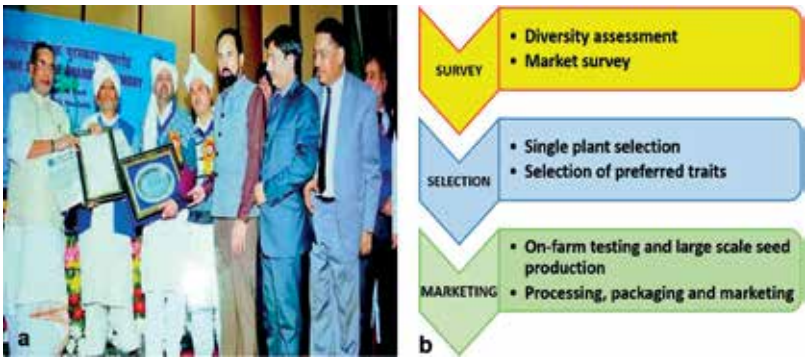

Figure 8. a) Award ceremony; b) Flow chart showing strategy for enhancing market 'value' of local landraces.

5.4. Production of organic rice

The continuous cultivation of only a few select rice varieties has led to the loss of genetic diversity of landraces. Although many landraces are preserved in seed banks these are not accessible to the farmers. Ensuring genetic diversity requires that rice landraces are cultivated continuously, and not simply stored in seed banks. In view of a preferred shift towards organic cultivation, and consumers' readiness to pay higher price for organically cultivated rice, the scope of traditional rice landraces has increased tremendously. Earlier HYVs enjoyed a distinct advantage, owing to their being fertilizer responsive, but under organic cultivation these are bound to lose the yield advantage. Therefore on-farm cultivation of landraces organically shall not only be a profitable economic enterprise for the farmer but would also lead to their conservation and adaptation. This process would be dynamic, i.e. the landraces would get subjected to continuous selection by the farmers, and would thus be allowed to develop and evolve.

The challenges in production of organic rice would however be in the form of 'paddy blast' and 'weed growth', both of which are currently controlled by chemical intervention.

5.5. Promotion of brown rice

As rice bran contains up to 20% lipids, this makes brown rice susceptible to rancidity. In earlier times, rancidity due to these rice lipids was prevented by removing the hull shortly before its consumption, and thus protecting it from oxidation. Rice (with bran) could then be stored for about 1 year, without leading to rancidity. Rice bran is characterized by high nutritional value. It contains high proportion (nearly 80%) of unsaturated fatty acids, which are known to have blood cholesterol lowering effects. The major unsaturated fatty acids in rice are oleic acid (a monounsaturated acid) and linoleic acid (an essential polyunsaturated fatty acid), and these are not synthesized in humans and therefore need to be taken from outside. These lipids play important roles in cell membrane function and functioning of the nervous system. The consumers' preference for milled rice has further reduced the availability of iron and zinc substantially in their diet. The variability in mineral content among different rice landraces of

Kashmir is quite pronounced, with zinc content ranging between 30 and 80 mg/kg and iron content between 16 and 55 mg/kg.

The current prevalence of milled rice on the market reduces the rice's nutritional value and essentially turns it into a simple carbohydrate food. Therefore, in addition to developing more nutritious varieties, awareness of the benefits of eating brown rice should be raised among rice consumers. Such a combined approach would ultimately result in a sustainable enhancement of the essential nutrient supply in rice-based diets.

5.6. Gene mining using genomics and biotechnological approaches

Rice is a diverse crop that grows in different ecosystems. Genomics-based strategies for gene discovery, coupled with the validation of transgenes by genetic transformation, have accelerated the identification of candidate genes from this broad genetic diversity. It is therefore important to explore landraces as well as wild rice species, and characterize their genes for further use rather than storing them in gene banks. Current gene revolution has broadened the scope for the application of biotechnology in rice, across ecosystems and genetic barriers. In order to prevent biopiracy policymakers should look into the potential use of biotechnology in safeguarding intellectual property rights of rice farmers and scientists by promotion of finger-printing technologies for molecular characterization of rice germplasm.

With accumulation of genes from a few elite parental lines in the current generation rice varieties, the genetic base has plateaued leading to comparatively lower genetic gains over the existing high yielding varieties. The primary attention of converging genes for yield and yield component traits needs to be diversified. Grain quality and physiological traits like cold tolerance, thermosensitivity, source-sink relationship and harvest index, lodging resistance, better nutrient absorption, efficient number of productive tillers per hill, etc. have not been explored biotechnologically to generate meaningful results for higher and sustainable crop yields in rice under Kashmir conditions.

5.7. Combining high yields with high nutritional value through molecular breeding and biotechnology

The immense genetic diversity in rice landraces is reflected by their multiplicity of nutritional characteristics. Suitable rice varieties exist for enhancing the supply of various nutrients, including protein, essential lipids, certain minerals, and to some extent β- carotene also. The diversity of such favorable nutritional characteristics is not represented in most of the widespread HYVs currently prevailing in Kashmir. These HYVs have been developed mainly to optimize the quantitative yields, and not the nutritional value. The high nutritional quality of rice landraces can form a solid base for changing priorities in rice breeding, putting more emphasis on the grain nutritional value. In order to meet the targets of nutritional security and food security, biotechnology and molecular breeding techniques, like marker-assisted selection, marker-assisted backcrossing, and genetic transformation need to be employed for accelerating the development of more nutritious rice varieties. Combining high yields and high grain nutritional value thus appears to be possible through these molecular interventions.

Acknowledgements

The lead author expresses his gratitude to the Fellow Scientists (Prof. G.A. Parray, Dr. M. Anwar Bhat, Dr. Gulzar S. Sanghera, Dr. Subhash Kashyap, Dr. Ashaq Hussain, Dr. Asif Shikari, Dr. M.A. Mantoo and Dr. Manzoor A. Ganai) for their support, and wishes to express appreciation for the supporting staff of PBG Section (Haji Gh M. Bhat (TA), Mohd Shaban (FCLA), Gh Nabi Dar (FCLA) etc.) for always carrying out instructions with sincerity and dedication.

Author contributions

AMH has contributed ideas, conceived, structured and written the whole chapter including preparation of figures; while SNR has contributed in preparation of **Table 1**, **Figures 4** and **8**).

Author details

Amjad M. Husaini[1]* and Sofi NR[2]

*Address all correspondence to: amjadhusaini@skuastkashmir.ac.in

1 Division of Plant Biotechnology, Sher-e-Kashmir University of Agricultural Sciences and Technology of Kashmir, Srinagar, Jammu and Kashmir, India

2 Mountain Research Centre for Field Crops, Sher-e-Kashmir University of Agricultural Sciences and Technology of Kashmir, Srinagar, Jammu and Kashmir, India

References

[1] Sanghera GS, Husaini AM, Anwar A, Kashyap S. Evaluation of some IRCTN rice genotypes for cold tolerance and leaf blast disease under temperate Kashmir conditions. Journal of Hill Agriculture. 2011;**2**:28-32

[2] Husaini AM, Parray GA, Wani SA. Biodiversity of rice germplasm and its conservation in Kashmir: Targets, accomplishments and way forward. In: University Research in Kashmir Rice - A Historical Perspective. India: SKUAST-K; 2013. pp. 27-47

[3] Ashraf H, Husaini AM, Bhat MA, Parray GA, Khan S, Ganai NA. SSR based genetic diversity of pigmented and aromatic rice (Oryza sativa L.) genotypes of the western Himalayan region of India. Physiology and Molecular Biology of Plants. 2016;**22**:547-555

[4] Husaini AM, Bhat ZA, Sanghera GS, Bhat GN, Rather AG. Screening and development of blast resistant rice varieties for Kashmir valley. Journal of Himalayan Ecology and Sustainable Development. 2011;**6**:1-6

[5] Wissuwa M, Yano M, Ae N. Mapping of QTLs for phosphorus-deficiency tolerance in rice (*Oryza sativa* L.). Theoretical and Applied Genetics. 1998;**97**:777-783

[6] Wissuwa M, Ae N. Genotypic variation for tolerance to phosphorus deficiency in rice and the potential for its exploitation in rice improvement. Plant Breeding. 2001;**120**:43-48

[7] Gamuyao R, Chin JH, Pariasca-Tanaka J, Pesaresi P, Catausan S, Dalid C, Slamet-Loedin I, Tecson-Mendoza EM, Wissuwa M, Heuer S. The protein kinase Pstol1 from traditional rice confers tolerance of phosphorus deficiency. Nature. 2012;**488**:535-539

[8] Frei M, Becker K. Studies on the in vitro starch digestibility and the glycemic index of six different indigenous rice cultivars from the Philippines. Food Chemistry. 2003;**83**: 395-402

[9] Ashraf H, Murtaza I, Nazir N, Wani AB, Naqash S, Husaini AM. Nutritional profiling of pigmented and scented rice genotypes of Kashmir Himalayas. Journal of Pharmacognosy and Phytochemistry. 2017;**6**(6):910-916

[10] Frei M, Becker K. Agro-biodiversity in subsistence-oriented farming systems in a Philippine upland region: Nutritional considerations. Biodiversity and Conservation. 2004;**13**:1591-1610

[11] Gregorio GB. Progress in breeding for trace minerals in staple crops. Journal of Nutrition. 2002;**132**:500S-502S

[12] Kennedy G, Burlingame B. Analysis of food composition data on rice from a plant genetic resources perspective. Food Chemistry. 2003;**80**:589-596

[13] Sharma TR, Chauhan RS, Singh BM, Sagar V, Paul R, Rathour R. RAPD and virulence analysis of *Magnoporthe grisae* rice populations from north-western Himalayan region of India. Journal of Phytopathology. 2002;**150**:649-656

[14] Zhu Y, Chen H, Fan J, Wang Y, Li Y, Chen J, Fan J, Yang S, Hu L, Leung H, Mew TM, Teng PS, Wang Z, Mundt CC. Genetic diversity and disease control in rice. Nature. 2000;**406**:718-722

[15] Parray GA, Shikari AB, Ganai MA, Husaini AM. Local aromatic rice cultivars of North Himalayas - Present status and future strategies for their sustainable improvement. Crop Research. 2008;**36**(3):380-381

[16] Parray GA, Shikari AB, Husaini AM, Wani SA, Rather AG, Teli MA. Kohsar: A new high yielding *japonica* rice variety released for high altitude irrigated areas of Jammu & Kashmir. Research on Crops. 2008;**9**(2):508-510

[17] Husaini AM, Parray GA, Rather AG, Sanghera GS. Performance of elite basmati rice varieties of subtropical India under valley temperate conditions of Kashmir. International Rice Research Notes. 2009:1-5

[18] Sanghera GS, Ahmed N, Zargar MA, Rather MA. Studies on the performance of some rice CMS lines under temperate conditions of Kashmir. SKUAST Journal of Research. 2003;**5**:121-124

[19] Sanghera GS, Husaini AM, Rather AG, Parray GA, Shikari AB. SKAU 7A and SKAU 11A: New cold tolerant CMS lines from Kashmir, India. International Rice Research Notes. 2010:1-3

[20] Sanghera GS, Husaini AM, Parray GA, Rather AG, Shikari AB, Wani SA. Generation of cold tolerant CMS lines of rice and identification of maintainers/restorers for hybrid rice development in Kashmir. Indian Journal of Crop Sciences. 2010;**5**:143-146

[21] Sofi NR, Parray GA, Bhat MA, Wani SA, Husaini AM. Participatory seed production of cold and blast tolerant high altitude rice. In: University Research in Kashmir Rice - A Historical Perspective. India: SKUAST-K; 2013. pp. 64-68

Chapter 4

Morpho-Agronomic Variation among *Phaseolus vulgaris* Landraces: A Review

Nontuthuko R. Ntuli

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.71359

Abstract

Phaseolus vulgaris L. of the family Fabaceae is widely grown for essential nutrients in its edible leaves, immature pods, and mature seeds. Landraces are local crops with wide morphological and genetic diversity. Morpho-agronomically, *P. vulgaris* landraces vary exceptionally in their vegetative and reproductive traits. These landraces vary in their germination rate and final percentage. Their growth form varies from bushy to vining type. Flowers range in their time to flowering, color, and size. Pods also vary widely in their time to pod formation; pod size, color, and shape; number of pods per plant; and time to pod maturity. Seeds also vary in their size, shape, color, and mass, as well as their number per pod and per plant. These landraces also vary in their resistance to pests and diseases from seed germination, plant growth and yield, and seed storage duration. A review on variation among *P. vulgaris* landraces forms basis for their future breeding as they are a good source of genetic diversity. This enables a possible selection for leaf, pod, and seed consumption, as well as resistance toward pests and diseases during the entire growth.

Keywords: *Phaseolus vulgaris*, traits, variability, landraces, morpho-agronomic

1. Introduction

Phaseolus vulgaris L. known as common bean is a member of the family Fabaceae [1]. It is an annual leguminous crop grown for its nutritional leaves, tender pods, and dry seeds [2]. It is a warm season legume crop and is self-pollinating with low frequency of crossing [3]. *P. vulgaris* provides protein and calories [4] as well as micronutrients such as zinc (Zn) and iron (Fe), essential vitamins, dietary fiber, and fat [5]. It is also an important legume which contains antioxidants [6, 7] and other chemically diverse components which fight against many diseases [8].

A landrace is defined as a crop with wide genetic diversity, which is usually identifiable, is known locally, has a local name, and has not undergone the proper crop improvement [4].

Landraces of *P. vulgaris* show a wide range of variation in their vegetative and reproductive traits [5]. The germination percentage among common bean cultivars ranges from 89 to 94% [9]. *P. vulgaris* landraces either show bushy, determinate, or indeterminate climbing growth form [10]. The number of branches among *P. vulgaris* landraces ranges from 17 to 57, and the number of leaves ranges from 19 to 37 [1]. Furthermore, days to flowering ranges from 26 to 40 days after sowing in *P. vulgaris* cultivars [9]. Some *P. vulgaris* landraces show white flower color, while others lilac [10]. *P. vulgaris* landraces have green and yellow mottling color of immature and matured pods, respectively [11]. Seed colors vary from black, brown, cream, green, mix red and white around the hilum, purple, and white to white/mottled [10, 12]. Pod color varies from pure green to green with purple or carmine stripes [11]. Pod length varies from 67.4 to 163.4 mm [13]. The number of pods per plant among *P. vulgaris* cultivars ranges from 5.4 to 9.9, while the number of seeds per pod ranges from 2.9 to 4.4 [9]. Seed length also ranges from 10.0 to 16.7 mm, width from 6.1 cm to 9.8 mm, and height from 4.2 to 8.2 mm [14]. Studies on the variation among *P. vulgaris* landraces are essential to select the desired traits for future breeding.

2. Taxonomy, uses, and variation among *Phaseolus vulgaris* landraces

2.1. Taxonomy, origin, and distribution of *Phaseolus vulgaris*

P. vulgaris L. belongs to subclass Rosidae, order Fabales, family Fabaceae, and subfamily Papilionoidea [5]. It is commonly known as the common bean [1], French bean, garden bean, kidney bean, snap bean, or string bean [15]. The genus *Phaseolus* contains more than 150 species [1], where the major domesticated species are *Phaseolus acutifolius* A. Gray, *P. coccineus* L., *P. lunatus* L., *P. polyanthus* Greenman, as well as *P. vulgaris* L [16]. *P. vulgaris* is the third important legume crop grown worldwide, after soya beans (*Glycine max* L.) and peanut (*Arachis hypogea* L.) [17].

Common beans are mostly annual, while others are short-lived perennial. They are cultivated in the warm climatic regions especially in tropical, semitropical, and temperate regions [18]. *P. vulgaris* is predominantly self-pollinating species with low average of cross-pollinating rate (3%) [3]. *P. vulgaris* is cultivated under various conditions in all continents and countries [19]. It is grown in a variety of soil types rich in organic matter, light loamy, sandy loam, well-drained soils with range pH of 5.7 and 7.0 neutral [20]. Fall, summer, and spring are seasons suitable for good crop production of *P. vulgaris* with optimum growth temperature ranges from 16 to 30°C [1].

P. vulgaris is native to Central and Southern America, where the world biodiversity hotspots of *P. vulgaris* are South-Central Mexico [19]. It was introduced to Africa and worldwide by Spaniards and Portuguese [2]. The African countries that are major producers of *P. vulgaris* are Burundi, Democratic Republic of Congo, Ethiopia, Kenya, Malawi, Rwanda, South Africa, Tanzania, and Uganda [20].

2.2. Uses of *P. vulgaris*

P. vulgaris is considered as a basic crop in many developing countries due to its high content of protein, micronutrients, vitamins, minerals, fiber, and carbohydrates [21]. It also serves as a

source of iron and thus is consumed as a meat substitute [7]. In some varieties, green immature pods are cooked as vegetable, while mature seeds are cooked and consumed for their high nutrient content [21]. The consumption of *P. vulgaris* is higher among both rural and urban societies with low income [20].

The consumption of common bean has health benefits by decreasing and preventing the glucose and cholesterol level [21, 22]. It also prevents stress and cancer and decreases heart diseases and obesity [8, 21]. It consists of enzyme inhibitors as well as compounds such as phenolic, phytates, and lectins, which help in metabolic functions in animal and human body systems [6].

However, *P. vulgaris* also has some problems due to the presence of certain anti-nutritional compounds such as saponins, flatulence factors, lectins, and phytic acid, and it also needs prolonged cooking [21, 22]. *P. vulgaris* fixes nitrogen to the soil through rhizobia by nitrogen-fixing bacteria [23].

2.3. Landraces and their uses

Landraces are crops with wide genetic diversity, which are usually identifiable, are known locally, and have not undergone the proper crop improvement [4]. Landraces are categorized into primary and secondary landraces [24]. Primary landraces contain their original and uncontaminated traits, whereas secondary (improved) landraces consist of foreign material that was incorporated into them through partial breeding [24]. Secondary landrace may change back to primary landrace after sometime [24]. An autochthonous landrace is a variety which is native and grown for a long period of time in a certain environment within a particular agricultural system [25]. It has specialized traits that allow biotic and abiotic stress conditions to increase and stabilize their yield [25]. Allochthonous landraces are varieties which are taken from other regions and introduced (grown) in another region and then allowed to adapt to that new region [24]. Landraces are naturally selected and are also characterized by the lack of formal genetic improvement [26].

Landraces play a significant role in agricultural production ensuring quality and well-managed crops [26]. They are varieties that have genotypes with wide specific traits [27]. These traits are adaptive to a specific environment and produces well-improved genotype, reduces the vulnerability, resistance to pests and diseases [27]. Landraces serve as a source of genetic diversity, and plant breeders often use specific traits to create new variation and maximize genetic diversity [12]. It also plays important role in ensuring food security [26]. Landraces result in high to intermediated yield, which is also stable under a low-input agricultural system in small-scale farmers [24, 27, 28]. They are a unique source of special traits which have marginal environment tolerance and nutritional quality [26]. The basis of diversity in landraces is genetic heterogeneity [29].

Common bean landraces have advantages of adaptation to cultural practices and local climatic conditions, resistance or tolerance to diseases, and early or late seed maturation, resulting high to intermediate yields under low inputs [10]. In eastern and southern Africa, farmers grow *P. vulgaris* landraces as genetic resources to be used for breeding programs [11]. *P. vulgaris*

landraces result to higher variation within the population [10]. Landraces are also much appreciated for their taste, high nutritional value, and short cooking time [28].

2.4. Germination percentage

The higher germination percentage of seed depends on the availability of environmental factors, like adequate temperature, light, salinity, moisture, and water [30]. The germination stage is the most important stage in the crop survival, which is to determine the amount of water and nutrient resources that need to be applied [31]. In Mexico the germination percentage ranges from 58.27 to 73.51% among the *P. vulgaris* landraces [31]. *P. vulgaris* landraces from Uganda show uniformity in seed germination, where after 5 days of planting all genotypes emerge from the soil [11].

2.5. Growth form, plant height, and number of branches and leaves

P. vulgaris differed in their growth habits which may be climbing or semiclimbing, erect or even bush type [1]. Their growth habit can either be determinate or indeterminate [17]. These growth habits are classified into four major classes, namely, Type 1 has determinate, upright, and bushy habit; Type 2 has indeterminate, upright, and bushy habit; Type 3 has indeterminate, prostrate with no climbing or semiclimbing, habit; and Type 4 has indeterminate and strong climbing habit [17].

The plant height of Brazilian *P. vulgaris* landraces ranges from 338 to 988 mm [27]. According to Stoilova et al. [32], plant height of landraces from Portugal and Bulgaria ranges from 195 to 1234 mm with the average of 447 mm. However, Sozen et al. [33] record plant height among Turkey landraces ranging from 200 and 3100 mm. The plant height shows wide variability among the landraces in Madeira where climbing landraces have a variation from 1086 to 1441 mm and bushy from 138 to 382 mm [5].

P. vulgaris landraces from Portugal and Bulgaria with climbing growth form have the numerous branches than bushy type [10]. The number of shoots in the main stem shows variation, with a range either from 4 to 14 among the landraces in Uganda [11] or from 17 to 57 in Nigeria [1]. The number of leaves per plant varies with a range from 45 to 96 leaves among *P. acutifolius* landraces in Botswana [34].

2.6. Days to first flower formation and flower color

Days to flowering also vary among *P. vulgaris* landraces, which generally commences from 26 to 51 days after planting in Portugal and Bulgaria landraces [10], Honduras [35], and Uganda [11]. However, a variation in days to flowering from 35 to 75 days after sowing is evident among landraces from Mexico [32]. The color of the flowers among *P. vulgaris* landraces can be white, carmine, red, purple, pink, white with lilac edges, or white with red stripes [11, 32].

2.7. Color, shape, number, and size of pods

The color of immature pods in Uganda *P. vulgaris* landraces is pure green; green with purple, carmine, or red stripes; dark purple; carmine, red, or pink, whereas physiologically matured

pods are yellow, yellow mottling, red, pink, or dark purple in color [11]. Pod shape varies from straight to slightly curved to fully curved [35]. In Portugal and Bulgaria, Stoilova et al. [10] reported a number of pods per plant as ranging from 6.4 to 20.8 among landraces. In Greece, the number of pod per plant shows wide variation between the local landraces and commercial cultivars. The numbers of pod per plant ranges from 21.5 to 51.3, among local landraces, and from 20.4 to 28.4 among cultivars [36]. The number of pods per plant ranges from 6.3 to 18.1 among *P. vulgaris* landraces from Brazil [27]. Turkey landraces have a number of pods per plant ranging from 1 and 163 [33]. The number of pods among landraces from Chrisoupoli and Nakolets in Greece ranges from 51.8 to 101.1, respectively [37]. Pod length shows wide variation among *P. vulgaris* landraces, with a range from 89 to 129 mm in Portugal and Bulgaria [10], from 123 to 309 mm in Island of Madeira [5], and from 40 to 120 mm in Uganda [11]. *P. vulgaris* landraces also show variation in yield parameters, where the number of pods per plant varies from 20.5 to 51.3 [36]. Pod length varies from 67.4 to 163.4 mm [13].

2.8. Color, shape, number, and size of seeds as well as seed maturity

Genetic variability in *P. vulgaris* landraces is sometimes indicated by seed color, size, and shape (**Figure 1**) [28, 38]. Shininess of seeds can either be shiny, intermediate, or opaque [39]. There is a wide variation in both seed coat main and secondary colors. Seed coat main color can be brown, cream, red, white, or yellow, while the secondary color can be black, red, or violet on the entire grain [28, 40]. White grain seed is commonly used by commercial farmers [41]. Seed shape can be round (circular), oval, kidney, hook, truncate, as well as cuboid (rectangular) shape [12, 40].

The number of seeds per pod among the *P. vulgaris* landraces has comparable ranges from 4.96 to 5.01 in Greece [37], from 2.8 to 6.6 in Italy [13], and from 3.60 to 5.53 in Zimbabwe [42]. Landraces from Italy that are categorized into traditional and nontraditional agro-food products vary from 3.2 to 6.3 and from 3.0 to 4.9 seeds per pod, respectively [43].

Seed size varies widely among *P. vulgaris* landraces. In Kosovo, seed length has a range of 12.8–18.3 mm, width 7.4–10.1 mm, and thickness 4.6–6.9 mm [44]. In Turkey, seed length has a variation of 11.8–23.1 mm, width 5.8–15.4 mm, and thickness 0.7–10.0 mm [38]. Seed length also ranges from 10.0 to 16.7 mm, from width 6.1 cm to 9.8 mm, and thickness from 4.2 to 8.2 mm, in Iran [14]. Consumers normally favor medium-sized to large-sized seeds probably because of their mass, taste, and easiness in hydration when cooked [45]. Seeds have certain properties such as early or late maturity, as some physiological maturity ranges from 65 to 120 days [10, 46].

2.9. Plant resistance to diseases and pests

In Tanzania, the screening of different *P. vulgaris* landraces and released varieties against *Phaeoisariopsis griseola* (Sacc) [Ferr], which causes angular leaf spot disease, shows that landraces were resistant, while varieties were either intermediate resistant or susceptible to this disease [20]. This suggests the presence of resistant genes on these landraces toward the *P. griseola*. The response of *P. vulgaris* parental lines to infestation by bean fly (*Ophiomyia phaseoli*) ranges from susceptible to resistant in Kenya [47].

Figure 1. Variation in shape, size, and color of some Italian *P. vulgaris* landraces [28].

3. Conclusion

A wide variation in growth and yield of *P. vulgaris* landraces discussed in this review will enable a possible breeding selection for leaf consumption based on bigger and soft-textured leaves. A selection for green beans can be on pod size, texture, and yield. Further, selection for dry beans can be based on seed yield, size, taste, and cooking time, to name a few. Breeding for resistance toward pests and diseases can be enhanced on landraces with resistant genes.

Author details

Nontuthuko R. Ntuli

Address all correspondence to: ntulir@unizulu.ac.za

University of Zululand, South Africa

References

[1] Nwadike C, Terkimbi V. Effects of planting date on performance of common bean (*Phaseolus vulgaris* L) landraces of the Jos plateau: A preliminary studies. International Journal of Current Research and Academic Review. 2015;**3**:309-324

[2] Adesoye AI, Ojobo OA. *Phaseolus* spp.: Valuable but underutilized genetic resource in Nigeria. International Journal of Plant Breeding and Genetics. 2015;**9**(1):1-9

[3] Kouam EB, Ndomou M, Gouado I, Pasquet RS. Assessment of the genetic diversity of cultivated common beans (*Phaseolus vulgaris L.*) from Cameroon and Kenya using allozymes markers. Journal of Experimental Biology and Agricultural Sciences. 2017;**5**:088-097

[4] Njoki NB. Breeding for durable resistance to angular leaf spot (*Pseudocercospora griseola*) in common bean (*Phaseolus vulgaris*) in Kenya [thesis]. South Africa: University of KwaZulu-Natal; 2013

[5] Freitas G, Gananca JFT, Nobrega H, Nunes E, Costa G, Slaski JJ. Morphological evaluation of common bean diversity on the island of Madeira. Genetic Resources and Crop Evolution. 2011;**58**:861-874

[6] Piergiovanni AR, Lioi L. Italian common bean landraces: History, genetic diversity, and seed quality. Diversity. 2010;**2**:837-862

[7] Doria E, Campion B, Sparvoli F, Tava A, Nielsen E. Anti-nutrient components and metabolites with health implications in seeds of 10 common bean (*Phaseolus vulgaris* L. and *Phaseolus lunatus* L.) landraces cultivated in southern Italy. Journal of Food Composition and Analysis. 2012;**26**:72-80

[8] Camara CRS, Carlos A, Urrea CA, Schlegel V. Pinto beans (*Phaseolus vulgaris* L.) as a functional food: Implications on human health. Agriculture. 2013;**3**:90-111

[9] Lima ER, Santaigo AS, Araujo AP, Texeira MG. Effect of the size sown seed on the growth and yield of common bean cultivars of different seed sizes. Brazilian Journal Plant Physiology. 2005;**17**:273-281

[10] Stoilova T, Pereira G, De Sousa MMT, Carnide V. Diversity in common bean landraces (*Phaseolus Vulgaris* L.) from Bulgaria and Portugal. Journal of Central European Agriculture. 2005;**6**:443-448

[11] Okii D, Tukamuhabwa P, Odong T, Namayanja A, Mukabaranga J, Paparu P, Gepts P. Morphological diversity of tropical common bean germplasm. African Crop Science Journal. 2014;**22**:59-67

[12] Casquera PA, Lema M, Santalla M, De Ron AM. Performance of common bean (*Phaseolus vulgaris* L.) landraces from Spain in the Atlantic and Mediterranean environments. Genetic Resources and Crop Evolution. 2006;**53**:1021-1032

[13] Scarano D, Rubio F, Ruiz JJ, Rao R, Corrado G. Morphological and genetic diversity among and within common bean (*Phaseolus vulgaris* L.) landraces from the Campania region (southern Italy). Scientia Horticulturae. 2014;**180**:72-78

[14] Marzooghian A, Valizade M, Moghaddam M, Kooshki MH. Evaluation of seed storage protein in common bean by same biplot analysis. International Journal of Bioscience. 2013;**3**:101-107

[15] Panda A, Paul A, Mohapatra P. A study on variability, character association and path analysis for pod yield in French bean (*Phaseolus vulgaris* L.). International Journal of Bioresource and Stress Management. 2016;**7**:033-039

[16] Coelho RC, Faria MA, Rocha J, Reis A, Oliveria MBPP, Nunes E. Assessing genetic variability in germplasm of *Phaseolus vulgaris* L. collected I northern Portugal. Scientia Horticulturae. 2009;**122**:333-338

[17] Gomez O. Evaluation of Nicaraguan common bean (*Phaseolus vulgaris* L.) landraces [thesis]. Uppsala: Swedish University of Agricultural Sciences; 2004

[18] Ince AG, Karaca M. Genetic variation in common bean landraces efficiently revealed by td-DAMD-PCR markers. Plant Omics Journal. 2011;**4**:220-227

[19] Lanna AC, Mitsuzono ST, Terra ST, Vianello TGR, Pereira VR, Carvalho MAF. Physiological characterization of common bean (*Phaseolus vulgaris* L.) genotypes, water stress induced with contrasting response towards drought. Australian Journal of Crop Science. 2016;**10**:1-6

[20] Fivawo NC, Msolla NS. The diversity of common bean landraces in Tanzania. Tanzania Journal of Natural and Applied Sciences. 2011;**2**:337-351

[21] Adesoye AI, Ojobo OA. Genetic diversity assessment of (*Phaseolus vulgaris* L.) landraces in Nigeria's mid-altitude agro ecological zone. International Journal of Biodiversity and Conservation. 2012;**4**:453-460

[22] Hernández-Delgado S, Muruaga-Martínez JS, Vargas-Vázquez MLP, Martínez-Mondragón J, Chávez-Servia JL, Gill-Langarica HR, Mayek-Pérez N. Advances in genetic diversity analysis of *Phaseolus* in Mexico. Molecular Approaches to Genetic Diversity. 2015:49-73

[23] Argaw A, Mekonnen E, Muleta D. Agronomic efficiency of N of common bean (*Phaseolus vulgaris* L.) in some representative soils of eastern Ethiopia. Cogent Food & Agriculture. 2015;**1**:1074790

[24] Zeven AC. Landraces: A review of definitions and classifications. Euphytica. 1998;**104**:127-139

[25] Hammer K, Diederichsen A. Evolution, Status and Perspectives for Landraces in Europe. Biodiversity Technical Bulletin No. 15. Italy: Biodiversity International, Rome; 2009

[26] Villa TCC, Maxted N, Scholten M, Lloyd BF. Defining and identifying crop landraces. Plant Genetic Resources. 2006;**3**:373-384

[27] Bertoldo JG, Coimbra JLM, Guidolin AF, de Andrade LRB, Nodari RO. Agronomic potential of gene bank landrace elite accessions for common bean genetic breeding. Science in Agriculture 2014;**71**:120-125.

[28] Lo Bianco M, Grillo O, Cremonini R, Sarigu M, Venora G. Characterization of Italian bean landraces (*Phaseolus vulgaris* L.) using seed image analysis and texture descriptors. Australian Journal of Crop Science. 2012;**9**:1022-1034

[29] Veteläinen M, Negri V, Maxted N. European Landraces on-Farm Conservation, Management, and Use. Biodiversity Technical Bulletin No. 15. Rome, Italy: Biodiversity International; 2009

[30] Cokkizgin A. Salinity stress in common bean (*Phaseolus vulgaris* L.) seed germination. Notulae Botanicae Horti Agrobotanici. 2012;**40**:177-118

[31] Aguilar-Benitez G, Pena-Valdivia CB, Vega JR, Castro-Rivera R, Ramirez-Tobias HM. Seed germination and early root growth in common bean and maize landraces and improved cultivars at different water stress levels. International Journal of Applied Science & Technology. 2014;**4**:118-127

[32] Stoilova T, Pereira G, Tavares- De- Sousa M. Morphological characterization of a small common bean (*Phaseolus vulgaris* L.) collection under different environments. Journal of Central European Agriculture. 2013;**14**:1-11

[33] Sözen Ö, Özçelĭk H, Bozoğlu H. The revealing of morphological variability and characterization of local bean populations in eastern Black Sea Region. Biyoloji Bilimleri Arastirma Dergisi. 2014;**7**:29-36

[34] Molosiwa OO, Kgokong SB, Makwala B, Gwafika C, Ramakopane MG. Genetic diversity in Tepary bean (*Phaseolus acutifolius*) landraces grown in Botswana. Journal of Plant Breeding and Crop Science. 2014;**6**:194-199

[35] Meza N, Rosas JC, JP M'n, Ortiz JM. Biodiversity of common bean (*Phaseolus vulgaris* L.) in Honduras, evidenced by morphological characterization. Genetic Resources and Crop Evolution. 2013;**60**:1329-1336

[36] Mavromatis AG, Arvanitoyannis VA, Korkovelos E, Giakountis A, Chatzitheodorou VA, Goulas CK. Genetic diversity among common bean (*Phaseolus vulgaris* L.) Greek landraces and commercial cultivars: Nutritional components, RAPD and morphological markers. Spanish Journal of Agricultural Research. 2010;**8**:986-994

[37] Vakali C, Papathanasion F, Papadopoloulos I, Tamoutiside E. Preliminary results on a comparative study evaluating landrace of common bean (*Phaseolus vulgaris*) under organic agriculture in a protected areas in Greece. In: European Summer Academics on Organic Farming ,Lednice na Morava, June 24–27, 2009. Czech Republic.

[38] Hatice B, Omer S. A sample for biodiversity in Turkey: Common bean (*Phaseolus vulgaris L.*) landraces from Artvin. African Journal of Biotechnology. 2011;**10**:13789-13796

[39] Barelli MAA, Goncalves-Vidigal MC, Filho PSV, Neves LG, da Silva HT. Genetic divergence in common bean landrace cultivars from Mato Grosso do Sul state. Semina: Ciências Agrárias, Londrina. 2009;**30**:1061–1072.

[40] Venora G, Grillo O, Ravalli C, Cremonini R. Identification of Italian landraces of bean (*Phaseolus vulgaris* L.) using an image analysis system. Scientia Horticulturae. 2009;**121**:410-418

[41] Martins SR, Vences FJ, Miera S'e d, LE BMR, Carnide V. RAPD analysis of genetic diversity among and within Portuguese landraces of common white bean (*Phaseolus vulgaris* L.). Scientia Horticulturae. 2006;**108**:133-142

[42] Musango R, Kudzai K, Mhungu S, Tibugar H. Phenotypic characterization of common bean (*Phaseolus vulgaris* L.) accessions conserved at the genetic resources and biotechnology institute. Journal of Biodiversity and Environmental Sciences. 2016;**8**:2222-3045

[43] Piergiovanni AR, Villecco D, Lioi L, Zaccardelli M. Comparison among recognized and non-recognized *Phaseolus vulgaris* L. landraces as traditional agro-food products of the Campania region (Italy). Genetic Resources Crop Evolution. 2015;**62**:1009-1019

[44] Fetahu S, Aliu S, Rusinovci I, Kelmendi B, Caka H, Maliqi N. Diversity of seeds size and weight of common beans landraces (*Phaseolus vulgaris* L.) in Kosovo. International Symposium on Agriculture, Opatija, Croatia. 2012. pp. 270-274

[45] Mavromatis AG, Arvanitoyannis LS, Chatzitheodorou V, Kaitsa A, Patsiaoura I, Nakas CT. A comparative study among landraces of *Phaseolus vulgaris* L. and *P. coccineus* L. based on molecular, physicochemical and sensory analysis for authenticity purpose. Scientia Horticulturae. 2012;**144**:10-18

[46] Rahman IU, Afzal A, Iqbal Z, Ijaz F, Manan S, Ali SA, Khan K, Karim S, Qadir G. Growth and yield of *Phaseolus vulgaris* as influenced by different nutrients treatment in Mansehra. International Journal of Agronomy and Agricultural Research. 2014;**4**:20-26

[47] Okwiri Ojwang PP, Melis R, Githiri MS, Songa JM. Genetic analysis for resistance to bean fly (*Ophiomyia phaseoli*) and seed yield among common bean genotypes in a semi-arid environment. Field Crops Research. 2011;**120**:223-229

Chapter 5

Agronomic Performance, Nutritional Phenotyping and Trait Associations of Okra (*Abelmoschus esculentus*) Genotypes in South Africa

Abe Shegro Gerrano

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.70813

Abstract

Okra, *Abelmoschus esculentus* L. (Moench), is an important fruit vegetable crop which belongs to the family Malvaceae. It is a good source of protein, carbohydrates, vitamins, minerals, and enzymes that are often consumed in small quantities in developing country. Okra is a highly nutritious underutilized fruit vegetable crop in South Africa. However, despite its importance for food, nutritional, and health benefits, the crop is rarely produced in some areas of South Africa. The study was carried out to assess the genetic diversity using agro-morphological traits and nutritional contents towards future use in the okra breeding programme. The experiment was carried out at the Roodeplaat research farm of the Agricultural Research Council in a randomized complete block design replicated three times. Agro-morphological traits and selected nutrients were determined. The analysis of variance for both showed highly significant differences for most traits recorded. The multivariate analysis showed a wide genetic diversity among the okra genotypes, which could be exploited in selecting suitable and potential parents when breeding for high yield and nutritional qualities. The present study revealed the genetic potential of the genotypes studied and their importance for use in the breeding programme aimed toward addressing malnutrition, food security, and poverty alleviation by breeding for increased yields, and nutritional contents in South Africa.

Keywords: fruit, genetic diversity, multivariate analysis, nutritional content, okra

1. Introduction

Okra (*Abelmoschus esculentus* (L.) Moench) is an annual fruit vegetable crop propagated through seed and commonly grown commercially in tropical and sub-tropical regions of the world. It is also grown in warmer temperate regions of the Mediterranean region [1]. Currently, this crop is found all over the African continent [2–5]. It is one of the most important

African indigenous fruit vegetable crops belonging to the family Malvaceae. It originated in Ethiopia [6], the former Abyssinia, and was cultivated by the ancient Egyptians. Its cultivation spread throughout Middle East and North Africa [7, 8]. In Ethiopia it is also called Kenkase (Berta), Andeha (Gumuz), and Bamia (Oromica/Amharic) [9]. Authors of Refs. [4, 9–11] reported that okra is a multipurpose crop due to its various uses of the pods, fresh leaves, buds, flowers, stems, and seeds. The immature fruits can be consumed as vegetables, in the form of salads, soups and stews, fresh or dried, fried or boiled. The plant contains mucilage in various plant parts, which is associated with other important substances including tannins [12]. The biological functions of mucilage within the plant includes aiding in water storage, decrease diffusion in plants, aid in seed dispersal and germination, and act as a membrane thickener and food reserve. Okra contains proteins, carbohydrates, and vitamins [8] that plays a substantial role in food security, human health [5], and nutritional security. Consumption of young and green immature okra fruits is very important as fresh fruits, and it can be consumed in different forms [13] such as boiled, fried, or cooked. Okra seeds contain about 20% protein and 20% oil [7]. It was reported that the seeds can be dried and the dried seeds are a nutritious material that can be used to prepare vegetable curds or roasted and ground to be used as coffee additive or substitute [14]. Moreover, okra leaves can also be used as animal feed. In similar fashion, the green leaf buds and flowers are also edible [15]. Okra mucilage is used for industrial and medicinal applications [16] in different parts of the country in the world. Industrial use of mucilage is usually for glace paper production and has a confectionery use. Okra has found medical application as a plasma replacement or blood volume expander [17, 18]. A study conducted in China suggested that an alcohol extract of okra leaves can eliminate oxygen free radicals, alleviate renal tubular-interstitial diseases, reduce proteinuria, and improve renal function [19].

Okra is a traditional crop, which requires relatively low agronomic input, but can contribute substantially to sustainable agricultural production and productivity in South Africa and beyond. This species is under-exploited and have potential for contributing toward food, nutritional, and health security for current alarmingly growing population, contributes a vital role in income generation and poverty alleviation. It is a valuable source of nutrients [20] with important medicinal properties [21]. Its wide range of biodiversity contributes to food, nutritional security, health benefit, and income diversification in the subsistence farming system that predominates in the different parts of the world. Therefore, improving the genetic potential of indigenous fruit vegetables like okra species is of paramount importance for yield, and nutritional quality. Evaluation and characterization of germplasm is important and the first step to the breeders who desire sources of genes for novel traits. It was reported that characterization of genetic resources refers to the process by which accessions are identified, differentiated, or distinguished according to their morphological and/or nutritional quality traits [22]. Currently, there is no clear record on genetic characterization and evaluation of the genetic resources of this crop under South African condition. Okra production and productivity is negatively due to the use of low yielding local landraces and use of poor agronomic management practices. Furthermore, production technology, development of new cultivar, and okra management practices are very limited in South Africa. To date, there are no reports of any improved cultivars developed in South Africa for high yield, nutritional contents as well as disease and pest tolerance. In addition,

http://dx.doi.org/10.5772/intechopen.70813

the variation in the agro-morphological and nutritional composition of okra has not been determined in the country. Therefore, it is important to profile okra genotypes using agro-morphological traits and nutritional contents in the immature fruits of okra for future breeding purposes in South Africa.

2. Agronomic performance and nutritional quality of okra

For field evaluation, 50 genotypes of okra (**Table 1**) were obtained from the AVRDC (World Vegetable Center), Taiwan. In this study, the field experiment was conducted in the Gauteng province of South Africa under rain-fed conditions during the 2015 and 2016 summer growing seasons at Roodeplaat (25°59′S; 28°35′E) research farm of the Agricultural Research Council. It is situated at an altitude of 1168 m above sea level. Roodeplaat has annual maximum average temperature ranged from 15.38 to 30.36°C and receives an average annual rainfall of 584.21 mm during the cropping seasons. The experimental site has loam clay type of soil. Two seeds of each okra genotype was planted in three rows of 4 m length spaced at 0.85 m between rows and 0.4 m between the plants. The seedlings were thinned into one when fully establishment in the field. A randomized complete block design with three replications was applied. Trial management such as plot preparation, and hand weeding were done when required and supplementary irrigation was employed when rainfall is not enough for the growth and development of the crop under research.

2.1. Agronomic characterization

Morphological phenotypic traits were evaluated and recorded using the International Plant Genetic Resources [23] okra descriptor list. The agro-morphological traits record includes plant height (PH), number of fruits per plant (NFP), number of branches per plant (NB), number of leaves per plant (NL), number of internodes (NI), internode length (IL), stem diameter (SD), leaf length (LL), leaf width (LW), days to 50% flowering (D50%F), fruit length (FL), fruit diameter (FD), fresh fruit yield (FYLD), number of seeds per fruit (NSF), 1000 seed weight (TSwt), shell weight (Swt), fruit harvest index (FHI) and grain yield per plant (GY) (**Table 1**).

2.2. Nutritional characterization

The fresh and immature fruits of the okra genotypes were harvested and analyzed for total protein content and selected mineral elements (calcium, copper, iron, potassium, magnesium, manganese, sodium, phosphorus, aluminum, boron and zinc) at the analytical laboratory of the Agricultural Research Council in Pretoria, South Africa. Fruits of okra were collected from each replicate in the field for analysis of mineral elements and protein content. Laboratory analysis were performed in triplicate and the results were expressed as mean for analysis (**Table 2**).

Protein analysis: A dry oxidation method was used to determine the total nitrogen and the crude protein contents (N × 6.25) of the samples [24, 25].

Genotypes	Morphological phenotypic traits																	
	PH	NFP	NB	NL	NI	IL	SD	LL	LW	D50%F	FL	FD	FYLD	GY	NSF	TSwt	Swt	FHI
VI033778	84.50	22.33	6.89	37.83	8.50	69.18	17.60	17.06	7.89	56.33	27.91	25.47	136.80	61.21	65.07	47.36	110.73	35.51
VI041763	124.82	22.06	7.11	43.72	11.61	86.44	16.38	18.21	7.96	48.33	23.66	21.22	127.30	53.65	61.30	43.05	79.42	39.13
VI037993	81.52	31.00	8.78	55.56	8.00	61.08	16.06	16.32	8.98	56.33	21.85	25.17	130.76	53.20	67.78	38.63	81.95	33.16
VI060803	150.31	13.50	9.28	47.39	17.72	83.84	28.18	17.33	7.09	69.00	15.10	17.56	85.71	45.90	57.92	43.00	61.79	54.47
VI046567	79.01	26.17	7.72	41.50	6.78	47.02	17.83	17.73	7.26	56.33	23.64	24.07	262.88	63.58	63.87	42.76	116.50	20.59
VI055219	114.29	14.33	9.61	53.72	8.06	60.14	17.31	17.78	8.56	53.00	17.40	18.93	87.67	42.27	61.80	42.23	66.07	50.54
VI060802	89.68	26.28	7.56	49.28	7.11	38.22	20.76	13.28	5.69	54.17	22.18	23.95	126.96	65.90	74.02	43.02	101.19	41.35
VI055996	107.96	10.33	7.93	21.00	8.52	55.72	21.67	18.44	7.55	55.00	25.03	24.27	113.44	44.06	68.00	38.46	75.97	34.23
VI037996	77.49	13.11	8.94	46.67	6.61	61.20	16.10	16.58	8.75	74.50	14.58	23.74	253.58	39.02	69.44	42.21	156.03	23.04
VI060824	102.64	13.56	7.83	59.61	8.56	73.47	22.66	15.07	8.08	52.50	14.30	20.46	198.34	33.36	61.82	39.86	103.69	23.78
VI047672	121.06	25.78	5.17	38.17	6.83	57.54	20.49	13.61	7.09	53.00	21.19	22.89	105.86	65.04	65.58	43.16	91.81	38.39
VI033803	78.71	4.83	7.44	39.78	10.11	48.94	22.80	17.74	8.80	53.83	18.21	17.73	94.29	45.00	60.45	38.58	75.28	45.35
VI033796	91.75	12.72	5.50	34.44	8.17	87.80	19.77	20.83	9.36	53.67	23.44	17.57	110.27	37.13	64.32	35.19	80.03	31.20
VI033777	87.73	13.39	8.61	48.50	5.67	60.91	17.57	15.93	7.51	55.67	32.32	18.44	91.20	42.07	56.58	43.37	68.09	53.73
VI060679	96.77	20.61	4.17	24.44	8.11	67.51	13.36	14.33	7.08	53.33	22.25	22.70	115.59	62.75	60.35	42.23	91.18	67.39
VI033775	126.96	14.78	6.83	33.56	12.00	70.37	14.02	16.77	8.94	48.33	18.98	17.49	95.42	48.07	56.27	37.67	68.69	40.96
VI050958	121.87	12.06	6.44	40.11	8.50	72.05	21.94	14.99	8.26	49.17	15.89	15.27	90.92	42.46	59.20	40.69	66.82	43.52
VI055110	93.30	14.06	5.56	41.72	7.22	51.05	19.85	16.48	7.20	51.67	15.46	18.59	94.61	36.02	61.72	43.98	71.36	49.33
VI049632	69.02	15.83	5.67	51.61	9.61	55.56	24.95	15.70	7.72	56.33	14.10	18.98	79.15	42.44	56.95	37.28	65.70	48.16
VI056069	110.68	11.17	4.33	34.89	6.11	60.84	18.91	18.79	8.12	67.17	15.25	18.65	84.43	37.89	59.65	39.10	66.27	49.73
VI056457	75.88	27.61	6.28	37.39	8.61	86.20	14.22	12.34	5.28	49.67	17.86	23.79	105.96	59.66	55.23	38.51	65.03	60.03
VI033797	90.07	9.11	8.89	37.22	10.39	40.91	20.56	15.70	7.11	48.33	17.81	20.26	97.37	41.32	55.04	42.31	75.54	48.15
VI060131	75.59	11.22	6.94	57.42	6.00	72.39	24.05	15.58	9.81	55.26	14.91	17.66	78.69	37.19	51.57	37.66	64.43	55.98
VI060817	87.28	10.39	6.00	40.67	7.89	62.44	14.06	14.47	8.01	56.67	14.32	22.86	93.18	41.72	64.90	37.42	69.70	43.06
VI050150	90.61	15.61	6.00	39.61	12.28	51.19	15.52	16.13	8.75	53.00	14.17	17.58	84.10	41.80	53.42	34.73	80.40	49.31

Genotypes	Morphological phenotypic traits																	
	PH	NFP	NB	NL	NI	IL	SD	LL	LW	D50%F	FL	FD	FYLD	GY	NSF	TSwt	Swt	FHI
VI039652	111.73	27.11	4.61	26.56	5.78	47.60	16.59	14.69	6.60	48.00	21.78	25.47	108.13	52.64	55.22	48.50	71.67	64.70
VI050957	78.80	3.44	8.33	47.83	12.22	52.10	25.89	15.92	8.89	67.17	12.48	18.62	88.92	31.67	57.40	34.98	53.64	36.05
VI060678	75.84	19.72	4.17	37.83	7.17	50.05	14.94	12.85	5.43	48.00	24.12	24.45	123.21	56.86	67.06	45.90	78.92	41.90
VI055220	94.83	12.39	5.33	38.72	6.78	84.58	20.00	15.61	7.13	55.83	16.08	18.14	122.06	42.04	58.80	37.60	63.17	30.86
VI039618	113.47	27.17	5.11	31.78	10.50	50.51	10.98	14.36	4.96	49.33	22.91	24.54	132.82	62.65	58.63	32.08	96.62	24.19
VI060313	122.90	2.39	10.44	55.56	10.78	70.84	19.18	16.94	8.22	68.33	12.47	15.01	78.23	32.86	48.72	31.36	40.17	39.22
VI046561	72.54	7.61	5.47	39.78	7.59	42.81	14.74	15.93	8.30	56.33	15.10	22.08	130.58	41.00	55.82	45.12	97.95	34.71
VI055119	111.20	16.17	6.67	56.22	11.06	51.42	21.48	15.19	7.52	49.17	12.99	13.32	68.00	37.33	44.40	38.50	54.51	74.69
VI060823	107.59	2.78	9.00	46.56	7.72	72.87	23.58	16.86	5.84	66.83	11.73	14.26	52.41	25.88	50.62	28.66	31.57	55.64
VI056449	76.63	11.61	5.50	34.28	6.83	70.99	18.67	16.67	6.49	56.33	17.90	16.24	88.55	41.01	56.38	35.76	64.74	43.17
VI060822	77.78	3.92	6.78	40.57	6.72	46.53	19.22	15.82	9.01	69.67	8.06	21.14	65.91	33.76	64.72	30.17	38.44	46.08
VI041215	73.36	8.44	5.50	40.78	6.50	71.14	17.97	13.43	5.69	49.17	15.67	15.13	95.29	46.08	59.68	39.35	68.90	43.36
VI055421	90.93	8.28	5.89	26.28	7.28	73.64	21.88	16.80	8.28	53.33	13.36	14.39	50.36	25.33	34.79	32.08	24.64	71.22
VI056450	75.69	15.28	6.72	38.11	6.17	58.47	12.54	14.21	6.63	52.50	12.61	18.06	75.82	37.19	52.74	42.96	59.90	69.45
VI039651	86.17	23.17	4.39	27.94	7.33	35.07	11.52	9.77	5.74	51.83	19.26	22.39	111.85	41.58	56.35	33.26	72.58	39.02
VI055423	89.17	12.11	3.83	33.89	5.61	51.65	17.64	18.18	5.59	54.17	15.25	17.06	81.64	36.94	51.92	36.03	61.89	49.48
VI039638	112.89	14.58	5.39	21.67	8.11	93.21	13.68	12.43	6.16	47.50	13.95	16.36	95.93	35.92	47.75	38.72	58.86	40.98
VI056079	88.43	11.72	4.72	29.83	4.56	49.11	11.60	17.45	6.60	55.83	16.08	15.67	82.29	34.27	50.81	36.69	61.46	53.27
VI041210	58.63	4.33	7.89	53.56	6.67	47.36	12.68	15.96	7.75	56.33	13.68	15.89	69.70	28.86	40.81	32.26	39.69	44.81
VI050960	88.79	4.06	5.94	38.17	7.17	41.82	17.43	15.24	7.00	68.17	8.98	21.13	64.07	19.85	62.72	26.36	32.41	40.33
VI055884	47.15	1.56	5.67	39.61	12.50	38.48	13.95	15.43	7.94	74.33	13.41	11.91	77.61	30.93	47.10	30.57	33.83	38.32
VI056081	70.67	4.22	4.94	47.83	4.78	34.08	18.16	14.82	7.52	68.33	11.29	12.77	51.65	28.32	39.63	26.57	41.49	48.64
VI050956	57.26	2.06	8.33	60.67	5.50	52.15	17.74	15.57	7.57	46.83	12.51	11.42	40.49	19.58	34.71	17.75	31.69	43.86
VI050959	57.20	2.61	7.39	42.56	5.94	27.33	20.85	16.09	7.12	69.33	10.04	10.40	68.61	19.51	31.82	19.24	24.39	28.08
VI060790	45.81	4.00	4.50	36.50	4.22	23.93	12.13	17.23	8.38	66.17	9.03	16.77	120.07	19.07	31.80	27.37	61.17	24.32

Genotypes	Morphological phenotypic traits																	
	PH	NFP	NB	NL	NI	IL	SD	LL	LW	D50%F	FL	FD	FYLD	GY	NSF	TSwt	Swt	FHI
MS Genotype (G)	2851.36 ns	380.18**	16.20**	564.85**	37.77**	1581.03**	97.02*	21.74 ns	8.08	358.03**	157.33**	96.42**	10797.12**	907.67**	570.52**	261.47**	3779.95**	947.40**
MS Season (S)	715.43**	2062.66**	641.63**	126.00 ns	37.02 ns	15599.48**	9490.11**	10045.17**	3991.76**	128.76 ns	1937.33**	16.18 ns	54306.32**	32019.06**	7945.44**	1791.23**	125752.76**	46954.33**
MS G x S	7787.95**	223.42**	18.23**	283.97**	22.84**	1453.34**	199.89**	120.91**	48.12**	180.32**	156.01**	103.28**	10195.02**	1221.83**	688.83**	235.34**	5114.72**	1395.51**
CV (%)	7.91	22.29	25.87	7.76	22.27	8.17	16.15	16.16	27.18	3.73	6.77	5.65	7.57	7.02	5.32	5.05	7.83	7.07
LSD (0.05)	57.48	4.63	2.95	3.87	2.69	19.98	9.31	7.77	4.92	2.36	7.25	5.68	50.2	20.28	15.55	8.52	41.91	22.67

*, ** significant at 0.05 and 0.01, respectively; MS: mean squares; CV: coefficient of variation; LSD: least significant difference; **: highly significant at the 0.01 probability level; G: genotype; S: season; PH: plant height; NFP: number of fruits per plant; NB: number of branches per plant; NL: number of leaves per plant; NI: number of internodes; IL: internode length; SD: stem diameter; LL: leaf length; LW: leaf width; D5%F: days to 50% flowering; FL: fruit length; FD: fruit diameter; FYLD: Fresh fruit yield; NSF: number of seeds per fruit; TSwt: 1000 seed weight; Swt: shell weight; FHI: fruit harvest index and GY: grain yield per plant.

Table 1. Mean, mean squares, and least significant differences for morphological phenotypic traits of okra genotypes.

Genotypes	Concentration of mineral elements (mg kg^{-1}) and total protein content (%) in dry basis											
	K	Ca	P	Mg	Na	Fe	Al	B	Zn	Mn	Cu	Protein[a]
VI033775	21777.4667	4919.5200	3784.0067	3113.6567	566.6250	305.4650	66.4567	34.1087	33.8938	22.1936	6.9753	16.2625
VI033777	20199.5500	4049.4267	3364.8200	3049.9750	365.9400	243.1507	63.6460	26.5820	32.9281	15.0279	7.3294	14.9563
VI033778	20976.1833	4039.0500	3728.2700	2848.3783	427.8367	587.3322	248.2663	29.1645	39.4484	21.5503	9.5204	11.7188
VI033796	24781.3500	5659.1100	4910.4033	3933.0633	316.2057	580.2120	286.3238	27.1184	48.0576	33.4163	8.6937	13.1500
VI033797	19496.4000	5459.1100	3628.6400	3283.7150	465.3840	319.8758	87.9778	35.3847	39.5059	21.5607	5.7575	13.8500
VI033803	29732.9667	5251.9117	5086.1233	3859.9633	360.7653	244.4440	86.3564	33.0417	43.5501	23.3683	9.0871	14.6688
VI037993	25140.4000	5251.9967	4474.0133	3384.7550	346.3737	196.7373	50.2557	30.8226	41.6551	21.2625	9.1599	18.2500
VI037996	19549.6333	3625.8133	3758.7700	3042.0667	465.9550	314.0100	52.6698	27.7749	39.8695	18.1510	7.7360	19.1750
VI039618	20532.6667	6145.3900	4137.7267	3254.7100	473.0543	327.1998	134.5552	35.0339	37.3512	19.3329	7.1605	20.1375
VI039638	19143.2500	6434.1967	3634.2800	3239.2133	541.3173	237.0652	58.1016	33.6771	35.9769	21.1262	7.0842	18.0125
VI039651	18187.3667	5633.1733	3561.0367	3248.9417	445.8867	543.2582	198.0592	32.4005	35.5071	25.3214	8.8091	15.4750
VI039652	20332.4500	6799.6683	4230.5233	3538.3867	458.9537	229.6398	53.5867	38.2204	35.1689	21.8068	6.1323	16.1688
VI041210	22947.1500	5933.6300	3882.7467	3427.8833	358.7303	364.1095	122.3338	38.2099	34.9044	31.9749	8.3166	21.4531
VI041215	26819.2833	5868.2317	5023.2233	4119.1067	334.1373	450.2562	167.4110	31.2438	46.4874	25.0930	8.3696	15.8063
VI041763	21640.4833	7854.9067	4073.6100	3443.6567	647.6400	300.9452	123.0660	34.3759	36.9295	26.4952	7.4171	17.0000
VI046561	22492.4000	5654.5283	3447.9233	3464.2400	407.2767	229.2835	53.4742	33.3572	33.5706	23.6109	7.0620	14.6813
VI046567	21929.6667	5400.8167	4072.6400	3506.6067	354.7370	343.1148	97.5733	28.1606	42.7555	21.9402	9.2944	16.4313
VI047672	21399.2833	7307.9933	4188.6933	4308.0267	580.9880	163.1217	48.7502	38.8493	41.6830	23.0591	6.6877	16.3625
VI049632	26532.9667	7061.0317	4871.0333	4544.9117	565.4510	219.7217	46.9941	37.2903	46.2990	20.4599	9.7918	20.4813
VI050150	19739.6833	4545.3967	3529.1033	2866.0400	375.8160	349.8102	131.2670	36.4835	36.8622	22.2440	6.6769	13.7375
VI050956	25174.1333	5645.9150	3523.6000	3844.0583	550.9303	267.0100	39.8715	32.6595	33.1479	22.7143	7.6115	16.8250
VI050957	18911.7167	6711.7317	3341.8667	3097.9567	605.7137	213.1867	73.3786	48.9588	35.9209	18.8477	7.8291	13.2563

Genotypes	Concentration of mineral elements (mg kg^{-1}) and total protein content (%) in dry basis											
	K	Ca	P	Mg	Na	Fe	Al	B	Zn	Mn	Cu	Protein[a]
VI050958	30384.6500	7192.0783	4773.4067	4264.1850	371.8990	311.4003	88.2743	29.7895	46.0846	29.7776	7.8219	12.7063
VI050960	22861.0000	4190.5433	3760.5800	3129.9400	244.8727	214.4395	42.9363	31.1971	36.5446	21.9904	6.8090	16.6438
VI055110	22807.7500	5744.4700	3869.0600	3454.7217	390.3277	524.4017	235.1390	33.6967	38.6857	23.5905	8.1426	13.6375
VI055119	19443.2833	5290.2550	3633.6933	3186.2433	410.1533	305.7237	98.7495	37.6476	36.2588	21.0558	7.2133	17.1688
VI055219	23070.6500	4870.4083	3764.9300	3397.6450	602.0480	320.4245	69.9319	27.1459	31.2486	26.2729	6.6490	13.1188
VI055220	22039.1167	4959.4950	3930.5167	3313.6267	470.6367	331.5973	170.5438	31.1861	37.3141	24.7236	8.7069	13.9250
VI055421	22250.2333	5170.7733	4055.0100	3382.7017	541.3853	329.6240	154.7592	36.9024	37.6627	25.2132	7.3251	16.9313
VI055423	26855.7833	8242.4900	5246.9500	4346.0583	387.0487	269.0805	77.1162	37.6389	46.8939	38.4287	9.8883	10.5063
VI055884	26798.8667	6911.8150	4562.8467	4272.4200	378.2300	211.1957	43.7957	35.2969	43.8937	25.6543	9.8447	15.1625
VI055996	23505.8000	7743.9483	4134.1000	4332.8850	487.7183	312.7035	62.6846	34.3182	44.9328	27.8276	8.7146	0.0000
VI056069	28706.5167	6461.0400	4559.9067	4027.2917	458.7717	568.4727	248.5053	32.3390	41.9579	26.1344	7.8264	13.5438
VI056079	25276.4667	5577.5167	4487.2500	4071.8733	611.6603	248.1673	200.1460	37.7277	40.0674	25.4336	7.1736	14.5063
VI056081	24805.6667	7059.3000	4299.2600	4213.8217	606.2797	138.0773	61.8873	38.9608	43.3255	20.0586	8.8108	14.1344
VI056449	22553.7000	5359.5200	4707.8667	3515.1033	440.1860	275.8338	110.4897	28.1297	29.1162	19.2992	7.3364	15.8969
VI056450	20130.4667	5227.5867	3508.4967	2993.3767	435.4760	615.2870	217.4698	27.4518	33.7007	30.4491	8.6698	0.0000
VI056457	19639.6500	4436.3000	2992.7300	2935.7233	688.4357	184.4547	55.1700	34.1046	29.9832	15.7108	4.7011	18.3469
VI060131	18126.8833	6714.2217	3888.1533	2935.2733	753.2457	231.2972	55.9428	34.7359	40.4822	15.7953	8.1000	13.1188
VI060313	22134.4167	6565.3033	3187.5967	3537.5250	778.8260	258.7387	36.2318	31.8077	26.9255	22.8010	7.0514	15.7750
VI060678	20507.9667	4557.0733	3709.7167	2944.3133	454.7623	339.4737	136.0335	37.2839	36.7775	22.4670	6.3063	0.0000
VI060679	17983.3167	6812.6650	3331.1167	3023.8683	630.4910	524.9860	250.7027	35.8244	34.4461	22.9563	7.3806	13.3813
VI060802	23664.0333	3282.7067	4423.4400	3438.6433	338.7350	626.8430	315.8953	26.7719	45.3539	23.1645	7.8676	14.1031
VI060803	20822.4333	5370.3683	3266.7200	3014.2350	505.6263	292.1103	66.6540	34.0651	35.9979	21.4352	6.7664	13.2188

Genotypes	Concentration of mineral elements (mg kg^{-1}) and total protein content (%) in dry basis											
	K	Ca	P	Mg	Na	Fe	Al	B	Zn	Mn	Cu	Protein[a]
VI060822	23264.7500	6857.1450	4022.7600	3973.4783	505.8723	118.2778	40.3196	36.9215	35.6317	29.4834	6.6292	16.8844
VI060824	24137.4500	4564.6567	3796.4967	3204.3767	346.6313	746.5115	462.5675	31.2123	38.4833	29.4813	5.9867	16.6813
G. mean	22591.4630	5748.1354	4003.6012	3507.1444	475.1095	333.2189	121.5722	33.5885	38.3313	23.6905	7.7005	14.5054
LSD (0.01)	950.10	95.05	79.43	60.04	29.61	15.39	8.26	1.50	2.79	0.69	1.10	0.33
M. squares	25960000.00**	3993160**	901596.00**	690456.00**	43058.20**	63194.89**	24920.75**	56.08**	76.06**	63.46**	4.04**	60.65**
CV(%)	2.60	1.00	1.20	1.10	3.80	2.80	4.20	2.80	4.50	1.80	8.80	1.40

G. mean: grand mean; M. square: mean squares; CV: coefficient of variation; LSD: least significant difference; **: highly significant at the 0.01 probability level; [a] conversion factor of 6.25; K^+: potassium; calcium: Ca^{2+}; phosphorus: P^{3-}; magnesium: Mg^{2+}; sodium: Na^+; iron: Fe^{2+}; manganese: Mn^{2+}; B^{3+}; aluminum: Al^{3+}, Zn^{2+} and copper: Cu^+.

Table 2. Mean values for the concentration of selected mineral elements and crude protein content in the immature fruits of okra genotypes.

http://dx.doi.org/10.5772/intechopen.70813

Mineral analysis: K^+, Ca^{2+}, P^{3-}, Mg^{2+}, Na^+, Fe^{2+}, Mn^{2+}, B^{3+}, Al^{3+}, Zn^{2+} and Cu^+ contents in the samples of immature fruits were determined using the inductively coupled plasma-optical emission spectrometric method [26].

2.3. Data and analysis

The morphological phenotypic and nutritional data were subjected to analysis of variance using Agronomix computer software [27]. The means of all okra genotypes were compared by the least significance difference (LSD) at 0.05 probability level. The mean data were standardized and subjected to multivariate analysis [28] using principal component analysis (PCA). The correlation coefficients were also computed to determine the degree of trait association [28].

3. Results and discussion

3.1. Variation in agronomic traits

Characterization and evaluation of crop species is essential in crop improvement programme [29] to identify potential parents according to their traits [22]. In the present study, 50 okra genotypes were characterized for agro-morphological and nutritional traits. The mean squares for the analysis of variance for most of the agro-morphological traits and nutritional values recorded showed highly significant differences (**Tables 1** and **2**) among the genotypes indicating that there were the existence of wide genetic and phenotypic and nutritional variability among the 50 okra genotypes evaluated. Furthermore, the existence of significant genotype by season interaction showed the influence of the growing season on the agronomic performance of traits. The maximum variation was observed in leaf width, branches, number of fruits closely followed by number internodes. Other characters also showed considerable variability. Even though plant height and leaf length was not significantly different, the genotype VI060803 appeared to have the tallest plant compared to the rest of the genotypes (150 cm) and the shortest genotypes was VI060790 (45.81 cm) (**Table 1**) and these traits were influenced by the gene factor. It is usually observed that tall and thin plants easily lodge due to environmental factors such as excessive flooding, rain, wind and when they produce high fruit during favorable growing seasons, however, they are also essential for firewood, construction of houses, fences and for livestock feed. Furthermore, selection for tallness gene might be important when the yield performance is low during unfavorable environmental conditions. This okra genotype may produce high dry biomass compared to the rest of the genotypes and can be used as livestock feed during the dry season. Plant height and leaf length were not significantly different which might be due to these traits are controlled by dominant gene. The present values were double higher than the values (17.96–76.65 cm) reported in 21 okra genotypes in Ghana [30]. This might be due to the variation in genetic and environmental factors which prevailed during the growth period. According to Ref. [31], the plant height is controlled by gene. He also reported that it is closely associated with number of flowering nodes, average fruits per plant and number of internodes.

The number of fruits per plant varied from 2.00 in genotype VI055884, VI060313, and VI050956 to 31 in genotype (VI037993) followed by genotypes VI056457 (27) and VI039618 (27). Ref. [30]

reported the average mean value of 6.00, which was lower than the values reported in the current study. They also reported 20.00 fruits per plant, which was lower than the values found in the current investigation. Furthermore, [32] reported the number of fruits per plant that varied from 3.22 to 5.67 in Nigeria. The immature fruits of okra are consumed as a vegetable and should be fresh, tender, and green without indication of coloration. Selection of potential parents based on this phenotypic trait would be essential in okra breeding programmes to develop new cultivar in the country.

Number of branches ranged from 3.83 to 10.44 and the highest value was recorded in genotype VI060313 and closely followed by VI055219. Moreover, the number of leaves ranged from 21.00 in genotype VI055996 to 60.67 in genotype VI050956 followed by VI060131, VI055119 and VI037993 during flowering. Leaves are the primary sources of photosynthesis to produce better yield and yield-related traits in okra genotypes. The values currently reported were higher than the values what [32] reported in Nigeria. Highly significant variation was also observed in the number of internodes, internode length, stem diameter, leaf length and width. The genotype that had tallness gene had highest number of internodes, internode length, and thick stem. The thicker the stem, resists the environmental influences from lodging and can withstand high fruit yield. Days to 50% flowering varied from 46.63 to 74.50 and influenced by genotype and genotype by season interaction. The genotype VI050956 was the first to flower, which is significantly flowered early compared to the rest of the genotypes. Some of the genotypes were expected to be similar in early flowering (**Table 1**). This genotype could be selected for earliness trait. Early maturing in okra genotypes can be useful to escape drought condition and can be cultivated as climate change crop in drought prone areas of South Africa. Therefore, this trait is potentially very important in okra improvement programmes for earliness and drought escaper genes. Depending on the traits of interest, the user of this crop can select the genotypes for early maturity or late maturity groups for future use. These values reported in the present study were lower than the values reported in Nigeria for days to 50% flowering among the genotypes [30]. Fruit length significantly varied from 10.41 to 32.32 in which the highest value was recorded in genotype VI055777. This trait is the most economically important trait, which affects the yield of okra. As a fruit vegetable crop, the longer pods are very important for consumption and preferred by the consumers in the South Africa, therefore, this trait is important as selection criteria for the improvement programme of okra in the county. Ref. [30] reported that fruit length is the most important determinants in okra production. The wider the size and the longer the fruit is associated with higher number of seeds in the fruit per plant. In this study, the widest fruit was recorded in the genotypes VI037993 and VI039652 with the highest number of seed (67.78) per plant and thousand seed weight, which were the primary determination of the ultimate yield in okra. Hence, widest fruit, highest number of seeds and thousand seed weight were considered as selection criteria for the breeding of okra genotypes for yield and yield related traits. Ref. [33] reported that seed weight is largely a function of seed components such as protein, fat, ash, and nitrogen free extracts. The highest fresh fruit yield, number of seeds per fruit and grain yield per plant were found in the genotypes VI046567 and VI060802. The values reported in the present study were higher than the values reported by [30] in 21 okra genotypes. The genotypes that produced the heaviest pod wall (fruit shell wall) could be selected for parental lines to produce high fodder

yield for animal feed compared to the rest of the genotypes where the highest pod wall were recorded in VI037996, VI046567 followed by VI033778 and VI060802. This could help to provide the type of okra which is useful as fodder during the dry season. In this experiment, it is clearly seen that the late maturing genotypes produced significantly higher fodder yield compared to other genotypes. Fruit harvest index is one of the most important trait for drought tolerance and results in yield gains in both drought, irrigated or rain fed environmental conditions. In the present study, the highest fruit harvest index was recorded in the genotypes VI055421 and VI056450 and could be used as parental lines in the development of new cultivar for drought condition in the South Africa.

3.2. Variation in nutritional traits

Vegetables are major sources of essential minerals and vitamins and are often low in calories, fat, and sugar that are an important addition to any diet consumed, particularly for resource poor community. Deficiency of mineral elements and crude protein content is a wide problem in alarmingly growing human populations in the world. The mineral elements play an important role in the development of the human body [34]. The existence of wide genetic differences among the crop plants for the nutritional quality would assist the improvement of the crop of interest for high quality through breeding in the available germplasm collection/gene pool [34–36]. Identification of okra genotypes based on selected mineral elements and crude protein content will help in the selection of the best parents for breeding nutritionally enhanced okra for food and nutritional security in South Africa [36]. Calcium and phosphorus are very important in the formation of strong bones and teeth, for growth, blood clotting, heart function and cell metabolism [37, 38]. In the present study, the mean squares for the analysis of variance for the concentration of mineral elements and protein content recorded showed highly significant differences for all the nutritional traits (**Table 2**) in the immature fruits of okra genotypes indicating that there was a wide genetic variability among the genotypes evaluated. It was reported that potassium is the major cation of intracellular fluid, which helps to regulate the acid base balance, osmotic pressure and water balance [39, 40]. The concentration of potassium varied from 17983.32 to 30384.65 mg kg^{-1} and the highest concentration and uptake of potassium was observed in the genotype VI050958 closely followed by genotype VI033803, while significantly the lowest uptake was observed found in the genotype VI060679, respectively (**Table 2**). The overall mean value of the genotypes recorded in this study was much higher than the values what [41] reported in okra in Cameroon. Okra is a good source of potassium for human health. The predominant mineral element in the current study was potassium and the concentration of this mineral element is superior to that of the rest of the elements evaluated. It is a primary mineral element found in the body and plays an important role in maintaining fluid balance.

Genotypes VI055423 (8242.49 mg kg^{-1}), VI041763 (7854.91 mg kg^{-1}) and VI055996 (23505.800 mg kg^{-1}) were significantly higher in calcium content compared to all other genotypes (**Table 2**), respectively. The mineral element, calcium plays a significant function in the growth and development of plant meristems, root hairs and root tips as well as for bone development

and strength [42]. Calcium deficiency in plants leads to stunted growth and development of roots [38]. The values detected in this study were higher than the values reported by [43, 44].

The highest concentration of phosphorus was 5246.95 mg kg^{-1} and found in genotype VI055423; while significantly lowest concentration was recorded in genotypes VI060313 (3187.60 mg kg^{-1}) and VI056457 (2992.73 mg kg^{-1}) (**Table 2**). The concentration of phosphorus in immature and green fruits of okra determined in this study (expressed in mg kg^{-1}) were higher than the values reported by [44] in the previous study. Furthermore, the genotype VI055423 showed the highest concentration of micronutrients such as zinc, manganese and copper are the essential micro-elements that plays a great role in the human growth and development. Ref. [45] reported that zinc is one of the essential trace mineral nutrient for human nutrition. The values reported in current study for zinc content in okra fruits was higher than the values reported previously by [43]. This genotype also contributed to the substantial concentration of magnesium next to the genotype VI049632 (4544.91 mg kg^{-1}) with the concentration of 4346.06 (mg kg^{-1}). The overall concentration of mean value of magnesium (3506.61 mg kg^{-1}) determined in the genotype VI046567 mg kg^{-1}) was higher than the values reported by [43]. The sodium concentration varied from 244.87 to 778.83 mg kg^{-1} (**Table 2**). The highest significant concentration was found in VI060313 and VI060131, respectively, compared to other okra genotypes.

The iron content of genotype VI060822 (118.28 mg kg^{-1}) was significantly lower than that of the other genotypes studied, while genotypes VI060824, VI060802 and VI056450 had the highest concentration compared to the rest of the genotypes evaluated for this mineral element which is higher than the values reported by [43, 44]. In similar fashion, the genotype VI060824 had showed significantly highest concentration of aluminum (462.57 mg kg^{-1}). Among all genotypes, significantly highest concentration of boron was recorded in the genotype VI050957 (48.96 mg kg^{-1}) and the lowest in VI060802 (26.77 mg kg^{-1}). Okra had a mineral concentration mean values in the order of K > Ca > P > Mg > Na > Fe > Al > Zn > B > M > Cu, that magnifies the importance of macro- and micro-elements in health, growth and development of human body. The existence of mineral elements such as iron, zinc, manganese and nickel has been reported in the immature pods of okra [44, 46]. Okra provides an important source of vitamins and mineral elements which are often lacking in the diet in developing countries [44].

Protein is an essential component of the diet of animals and human and supplies the required amino acids [47]. Protein played a significant role in growth, development and replacement of lost tissues in the human body. It is an important nutritional component in the body to build and repair body tissues that is a building block of bones, muscles, cartilage, skin and blood. Protein is a macronutrient, which the human body needs relatively in a large amount. In this experiment, the crude protein content of immature fruits of okra was analyzed and it varied from negligible amount, of which nitrogen was not detected in genotypes VI055996, VI056450 and VI060678 to 21.45% in VI041210 followed by VI049632 and VI039618 (**Table 2**) indicating that there was a high significant variation among the genotypes due to the effect of genes they carry. Similarly, [48] reported 21.40% of crude protein in the fruits of okra. Ref. [49] also found 21% protein content in okra pods. Moreover, the current value was also almost similar to the

values what [41] reported in the fruits of okra; while higher than the values reported by [43] in Nigeria. Ref. [50] reported 19.5% crude protein content in the fruits of okra, which is lower than the values reported in the current study. This difference might be due to genetic and environmental conditions prevailed during the growth period. Ref. [51] reported 23.40% of crude protein in the fruits of okra in dry basis in Nigeria, which is relatively higher than the current study. Similarly, [52] reported that 23.68% of protein content in two okra genotypes in Pakistan. The green okra fruit showed slightly lower protein content, when compared to the protein contents reported in the immature green pods and fresh young leaves of cowpea genotypes [53, 54] evaluated in South Africa, and hence, okra could be considered as protein fruit vegetable in the dietary requirements. It was earlier reported that okra is a good source of protein among a few protein vegetables such as spinach, cauliflower, broccoli, asparagus and others. The genetic variation existed in this study would assist the breeders in selection of potential parental okra lines for the development of new okra cultivars with high protein content. Okra pre-breeding for nutritional quality would start with the selection of potential parents based on their individual nutritional values determined in the current evaluation. Due to the current prevalence of malnutrition in the world, particularly in sub-Saharan Africa and South Africa, breeding for higher nutritional quality which is suitable for human are so quite important for end users. Therefore, quantification and identification of the nutritional composition in the immature and green fruits of collection of okra genotypes is important to develop new okra cultivar with high nutritional composition of interest in the South Africa.

3.3. Trait association

3.3.1. Morphological phenotypic traits

The results of the association analysis for 18 morphological phenotypic traits are presented in **Table 3**. A strong positive and highly significant association was observed between grain yield and number of fruit per plant (r = 0.85), fruit length (r = 0.75), and fruit diameter (r = 0.73). Grain yield was also positively and significantly associated with number of seeds per plant as well as thousand seed weight, which indicates that all yield components are important for the improvement of grain yield in okra genotypes. Number of fruit was positively and significantly associated with yield contributing traits such as fruit length, fruit diameter, fruit yield, grain yield, number of seeds per fruit, thousand seed weight as well as shelled fruit weight without seeds. Moderately positive and significant association was also observed between grain yield and fruit yield (r = 0.48) and plant height (r = 0.34). Plant height was moderately and positively significantly associated with number of fruits per plant, number of internodes, number of seeds per fruit, internode length, and thousand seed weight. Moreover, fruit length was also positively and significantly associated with grain yield and its related traits. The strong positive association between the different phenotypic traits would allow the breeder for simultaneous selection and improvement of these traits. In this study, plant height, number of fruits per plant, fruit length, fruit diameter, fruit yield, grain yield, number of seeds per fruit, and thousand seed weight were identified as selection criteria for obtaining potential and good parents for the development of new cultivar in the okra breeding programme in the South African condition.

Phenotypic traits	PH	NFP	NB	NL	NI	IL	SD	LL	LW	D50%F	FL	FD	FYLD	GY	NSF	TSwt	Swt
PH	1.00																
NFP	0.33*	1.00															
NB	0.17	−0.18	1.00														
NL	−0.14**	−0.18	0.60**	1.00													
NI	0.47**	0.05	0.35**	0.10	1.00												
IL	0.52**	0.20	0.15	−0.03	0.29*	1.00											
SD	0.25	−0.24	0.41**	0.37**	0.31*	0.21	1.00										
LL	0.12	−0.30*	0.27	0.08	0.10	0.13	0.30*	1.00									
LW	−0.06**	−0.33**	0.33**	0.34**	0.16	0.06	0.27	0.53**	1.00								
D50%F	−0.23	−0.52**	0.22	0.19	0.05	−0.27	0.20	0.27	0.28	1.00							
FL	0.26	0.66**	0.00	−0.23	0.04	0.21	−0.13	0.04	−0.16	−0.46**	1.00						
FD	0.18	0.71**	−0.07	−0.24	0.00	0.03	−0.20	−0.18	−0.13	−0.22	0.58	1.00					
FYLD	0.03	0.48**	0.10	−0.02	−0.03	0.06	−0.14	0.03	0.02	−0.03	0.40**	0.61**	1.00				
GY	0.34**	0.85**	−0.10	−0.20	0.20	0.22	−0.13	−0.22	−0.28*	−0.44	0.75**	0.73**	0.48**	1.00			
NSF	0.34**	0.51**	0.08	−0.08	0.16	0.24	0.09	0.00	−0.01	−0.12	0.53**	0.76**	0.51**	0.67**	1.00		
TSwt	0.40**	0.61**	−0.03	−0.16	0.16	0.31*	0.02	−0.04	−0.06	−0.40	0.60**	0.61	0.47**	0.71**	0.65**	1.00	
Swt	0.15	0.61**	0.00	−0.07	0.03	0.10	−0.14	0.00	0.02	−0.21	0.55**	0.70	0.86**	0.67**	0.67**	0.67**	1.00
FHI	0.17	0.02	−0.11	−0.07	0.02	0.15	0.09	−0.14	−0.13	−0.24	−0.10	−0.22	−0.58	−0.04	−0.24	0.15	−0.39**

*, **: significant at 0.05 and 0.01, respectively; PH: plant height; NFP: number of fruit; NB: number of branch; NL: number of leaves; NI: number of internodes; IL: internode length; SD: stem diameter; LL: leaf length; LW: leaf width; D50%F: days to 50% flowering; FL: fruit length; FD: fruit diameter; FYLD: fruit yield; GY: grain yield; NSF: number of seed per fruit; TSwt: thousand seed weight; Swt: shell weight; FHI: fruit harvest index.

Table 3. Association analysis for 18 morphological phenotypic traits of okra genotypes.

3.3.2. Nutritional traits

Moderate to highly significant association was observed for the mineral and protein content determined in the okra genotypes (**Table 4**). Highly to moderate significant positive associations were observed between K and Ca, P, Mg, Zn, Mn, and Cu; while negative and moderate correlation was observed between K and Na. Similarly, there were significantly positive association between Ca and P, Mg, Na, B, Zn, Mn and Cu. It was also observed that there were highly significant associations between P and Mg, Zn, Mn, and Cu. These results suggested that high P content might be accompanied with increased concentration of Mg, Zn, Mn, and Cu contents of okra fruits and vice-versa. Na was negatively and significantly associated with all micronutrients except B indicating that high Na content associated with the low contents of the Fe, Al, Zn, Mn, and Cu. An extremely strong association was observed between Al and Fe (0.91) which might be the indication of the existence of genetic control compared to the rest of the traits evaluated. Strong association was also found between K and P as well as K and Mg. Significantly, negative association was observed between protein and Fe, and between protein and Mn. In general, in the current study most of the traits evaluated showed highly and significantly positive and moderate associations among them indicating that there were some functional interaction existed among the mineral elements and protein content. Ref. [55] reported the positive correlation among the mineral elements in rice was due to the interaction between ions whose chemical properties were sufficiently similar, and they compete for site of absorption, transport, and function in plant tissues. Hence in the present study, positive association between and among the mineral elements and protein contents showed that

Mineral elements and protein	**Concentration of mineral elements (mg kg^{-1}) and total protein content (%) in dry mass basis**										
	K	**Ca**	**P**	**Mg**	**Na**	**Fe**	**Al**	**B**	**Zn**	**Mn**	**Cu**
K	1.00										
Ca	0.24**	1.00									
P	0.75**	0.34**	1.00								
Mg	0.75**	0.63**	0.74**	1.00							
Na	−0.32**	0.35**	−0.36**	−0.06	1.00						
Fe	−0.01	−0.33**	−0.00	−0.23**	−0.34**	1.00					
Al	0.05	−0.30**	0.09	−0.16	−0.29**	0.91**	1.00				
B	−0.14	0.50**	−0.07	0.14	0.36**	−0.42**	−0.27**	1.00			
Zn	0.55**	0.26**	0.73**	0.58**	−0.37**	0.14	0.18**	0.001	1.00		
Mn	0.46**	0.35**	0.42**	0.45**	−0.32**	0.35**	0.34**	−0.05	0.33	1.00	
Cu	0.37**	0.23**	0.48**	0.38**	−0.25**	0.111	0.039	−0.110	0.54**	0.27**	1.00
Protein	0.03	−0.06	0.05	0.02	0.04	−0.26**	−0.16	0.064	−0.11	−0.27**	−0.11

** significant at the 0.01 probability level.

Table 4. The correlation coefficients between mineral elements and total protein contents evaluated in immature fruits of okra genotypes.

selecting and improving the primary traits of interest would have a positive effect on the secondary traits in the breeding programme for nutritional quality.

3.4. Multivariate analysis

3.4.1. Morphological phenotypic traits

The principal component analysis (PCA) was used for the reduction of data set and transforming the available raw data set into principal components or component factors, which are equal to the number of evaluated morphological phenotypic traits (**Table 5**). From the current experiment, the PCA transformed 18 raw set of data into 18 factors loadings or principal components with the pattern that the first principal component (PC1) contributed the most variability and the last principal component (PCn) contributed the lowest variability, which accounted for the entire (100%) variability. However, the PC1, PC2, PC3, and PC4 showed high significant variability compared to the rest of the PCs (**Table 5**) with the eigenvalues greater than one and cumulatively accounted for 68.49% of the total variation among the okra genotype. These PCs had eigenvalue more than 1 [56, 57]; while the rest of the PCS had eigenvalue less than 1 [58], and would not be considered in the interpretation of the results obtained and removed, as they were not significantly influencing and contributing to the variability among the genotypes. Morphological phenotypic traits showed different pattern of contribution to the variability in the principal components loading suggesting the existence of genetic variability that would be used in the okra improvement programme. The current cumulative variation explained by the first four PCs was comparable with what [59, 60] reported in contributing to the variations among different okra genotypes. In the first principal component, grain yield, number of fruits, fruit diameter, shelled pod weight, thousand seed weight, number of seeds per fruit and fruit length, respectively, contributed high variability with positive loading compared to the rest of the traits. This principal component alone explained 33.10% of the total variability among the okra genotypes with the eigenvalue of 5.96. The PC2 accounted for 16.10% of the total variation and was mainly influenced by vegetative growth traits such as number of branches, stem diameter, leaf width, leaf length, number of internodes and number of leaves with positive loading. The PC3 with 12.72% variance distinguished the okra genotypes based on fruit harvest index, plant height, fruit yield, and internode length with all positive loading except fruit yield. Similarly, the PC4 was associated and dominantly influenced by the number of leaves, leaf length and number of branches with all positive loadings except leaf length and this PC accounted 6.58% variances. The remaining phenotypic traits had no any significant contribution to the variation in the four PCs and hence were of minor importance in the characterization of okra genotypes.

Biplot analysis was carried out based on the first two PCs. The genotypes and morphological phenotypic traits were shown on a biplot to clearly visualize their associations and differences (**Figure 1**). This PCA biplot more explained the 49.20% of total variability among the genotypes, displaying that number of branches, number of seeds, grain yield, number of fruit, fruit harvest index, and days to 50% flowering were considered as the most discriminating parameters (**Figure 1**). The genotypes that were positioned on the right top quadrant were closely associated and characterized by longest fruit, largest seed size, heavy fruit shell, highest fruit yield, highest number of seeds per fruit, tallest plant, longest internode, and highest number of

Phenotypic traits	Factor loadings																	
	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18
PH	0.16	0.22	0.40	−0.15	0.29	0.09	−0.06	0.04	−0.18	−0.16	0.67	0.08	0.06	0.32	−0.02	−0.05	−0.10	0.13
NFP	0.35	−0.10	0.08	0.16	−0.05	0.06	0.10	0.01	0.00	−0.04	0.14	0.56	−0.25	−0.14	0.10	0.00	0.49	−0.40
NB	−0.04	0.44	−0.02	0.36	−0.05	0.17	0.13	0.21	−0.48	−0.07	−0.01	−0.38	−0.28	−0.09	−0.16	0.21	0.10	−0.18
NL	−0.11	0.31	−0.10	0.59	−0.30	0.08	−0.03	−0.07	0.01	0.01	0.09	0.35	0.44	0.03	0.19	−0.15	−0.16	0.14
NI	0.05	0.32	0.25	0.07	0.50	0.08	0.54	−0.17	0.30	−0.04	−0.26	−0.06	0.10	−0.19	0.09	−0.19	0.03	0.02
IL	0.11	0.21	0.35	−0.19	−0.04	0.44	−0.47	−0.28	−0.19	0.23	−0.40	0.14	−0.06	−0.07	0.06	−0.02	−0.11	−0.05
SD	−0.07	0.39	0.17	0.11	0.00	−0.34	−0.42	0.31	0.54	−0.01	−0.01	0.00	−0.35	−0.02	0.01	−0.02	−0.02	−0.03
LL	−0.07	0.34	−0.07	−0.56	−0.26	−0.01	0.13	0.34	−0.07	−0.21	0.03	0.20	0.25	−0.44	0.05	0.08	−0.03	−0.02
LW	−0.10	0.36	−0.15	−0.25	−0.35	−0.10	0.31	−0.47	0.11	0.34	0.16	0.06	−0.28	0.27	−0.09	0.07	0.07	0.05
D50%F	−0.19	0.18	−0.29	−0.09	0.41	−0.37	−0.08	−0.02	−0.38	−0.13	−0.31	0.38	−0.05	0.32	0.05	−0.02	−0.01	−0.09
FL	0.32	0.01	0.04	−0.09	−0.24	0.10	0.23	0.51	−0.04	0.14	−0.28	−0.03	−0.06	0.46	0.09	−0.40	−0.01	0.11
FD	0.35	−0.02	−0.15	0.04	0.08	−0.26	0.04	−0.03	−0.23	0.29	0.16	−0.03	−0.29	−0.38	0.39	−0.15	−0.46	0.09
FYLD	0.27	0.11	−0.38	−0.01	0.05	0.20	−0.17	−0.15	0.05	−0.41	−0.04	−0.03	−0.21	−0.08	0.01	−0.09	0.28	0.61
Grain yield	0.37	−0.02	0.08	0.11	0.03	−0.06	0.15	0.10	0.09	0.05	−0.18	0.27	0.03	0.06	−0.35	0.64	−0.29	0.25
NSF	0.32	0.17	−0.07	−0.01	0.12	−0.31	−0.20	0.01	−0.11	0.45	0.04	−0.16	0.43	−0.12	−0.29	−0.10	0.42	0.06
TSwt	0.33	0.08	0.12	−0.04	−0.19	−0.28	−0.04	−0.22	0.01	−0.32	−0.13	−0.31	0.22	0.21	0.51	0.30	0.09	−0.18
Swt	0.34	0.09	−0.26	−0.02	−0.06	0.03	−0.06	−0.19	0.13	−0.30	0.01	−0.07	0.07	0.03	−0.45	−0.31	−0.37	−0.46
FHI	−0.06	−0.12	0.49	0.08	−0.30	−0.43	0.09	−0.19	−0.26	−0.27	−0.14	0.05	−0.11	−0.16	−0.28	−0.28	0.02	0.23
Eigenvalue	5.96	2.90	2.29	1.18	0.97	0.81	0.67	0.63	0.50	0.45	0.39	0.36	0.28	0.23	0.13	0.10	0.08	0.06
Variability (%)	33.10	16.10	12.72	6.58	5.41	4.52	3.73	3.48	2.77	2.50	2.18	1.99	1.55	1.26	0.73	0.58	0.47	0.33
Cumulative %	33.10	49.20	61.91	68.49	73.91	78.42	82.15	85.63	88.41	90.91	93.09	95.08	96.63	97.89	98.62	99.20	99.67	100.00

Table 5. Principal component factors, eigenvalues, individual, and cumulative variability for the morphological phenotypic traits.

http://dx.doi.org/10.5772/intechopen.70813

Figure 1. Plant characteristics.

internodes. The genotypes demarcated on the top left quadrant were associated with highest number of branches and leaves, widest stem and leaves, longest leaves as well as late maturing genotypes. Furthermore, the biplot demarcated the genotypes on the left bottom quadrant based on derived traits called fruit harvest index. This trait is the most important trait to select the genotypes for drought tolerance and these traits were suggested to have drought tolerance traits. Similarly, the right bottom quadrant consists of genotypes with highest grain yield, fruits per plant and widest fruits. The genotypes concentrated around the origin had similar genetic characteristics, while the genotypes that were found far from the origin are discriminated from the rest of the group due to their peculiar genes/alleles and considered as unrelated genotypes. Therefore, selection of these genotypes as potential parents would result in successful hybridization to develop heterotic groups in the okra-breeding programme (**Figure 3**).

3.4.2. Nutritional traits

The data set of all the mineral elements and crude protein contents were subjected to principal component analysis (PCA), which removed the highly inter-correlated and redundancy nature of the prevalent variations among the okra genotypes (**Table 6**). The PCA grouped the mineral

Mineral elements and protein	Factor loadings											
	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12
K	0.410	−0.037	0.199	0.128	−0.320	0.102	−0.181	0.597	0.056	−0.201	−0.479	−0.017
Ca	0.215	−0.392	−0.399	0.089	0.039	0.144	0.196	−0.392	−0.394	−0.245	−0.424	−0.155
P	0.441	−0.048	0.176	0.080	0.028	0.002	−0.214	−0.133	−0.380	0.725	0.017	0.176
Mg	0.407	−0.250	−0.003	0.115	−0.176	0.142	−0.130	−0.017	−0.082	−0.389	0.726	0.073
Na	−0.208	−0.322	−0.321	0.106	0.095	0.669	−0.268	0.130	0.339	0.268	0.009	0.081
Fe	0.070	0.522	−0.257	0.186	0.168	0.166	−0.010	−0.023	−0.131	−0.218	−0.087	0.703
Al	0.089	0.475	−0.257	0.388	0.229	0.080	−0.179	0.133	−0.151	0.032	0.126	−0.633
B	−0.035	−0.376	−0.349	0.230	0.364	−0.579	−0.048	0.419	−0.023	0.050	0.061	0.176
Zn	0.402	0.021	0.057	−0.113	0.394	−0.171	−0.389	−0.375	0.546	−0.143	−0.160	−0.038
Mn	0.325	0.124	−0.368	0.172	−0.376	−0.123	0.506	−0.063	0.467	0.277	0.055	−0.001
Cu	0.313	0.020	0.056	−0.453	0.520	0.270	0.481	0.324	−0.027	0.027	0.088	−0.062
Protein	−0.058	−0.146	0.521	0.677	0.276	0.121	0.348	−0.106	0.132	−0.043	−0.031	0.041
Eigenvalue	4.043	2.867	1.374	0.895	0.774	0.643	0.541	0.307	0.200	0.185	0.119	0.051
Variability (%)	33.693	23.893	11.454	7.455	6.449	5.359	4.508	2.560	1.669	1.540	0.993	0.426
Cumulative %	33.693	57.586	69.040	76.495	82.944	88.303	92.811	95.371	97.040	98.581	99.574	100.000

Table 6. Principal component analysis of mineral elements and protein traits in 46 okra genotypes showing eigenvectors, eigenvalues, individual, and cumulative percentage of variation explained by the first three PC axes.

Figure 2. Fruit characteristics.

http://dx.doi.org/10.5772/intechopen.70813

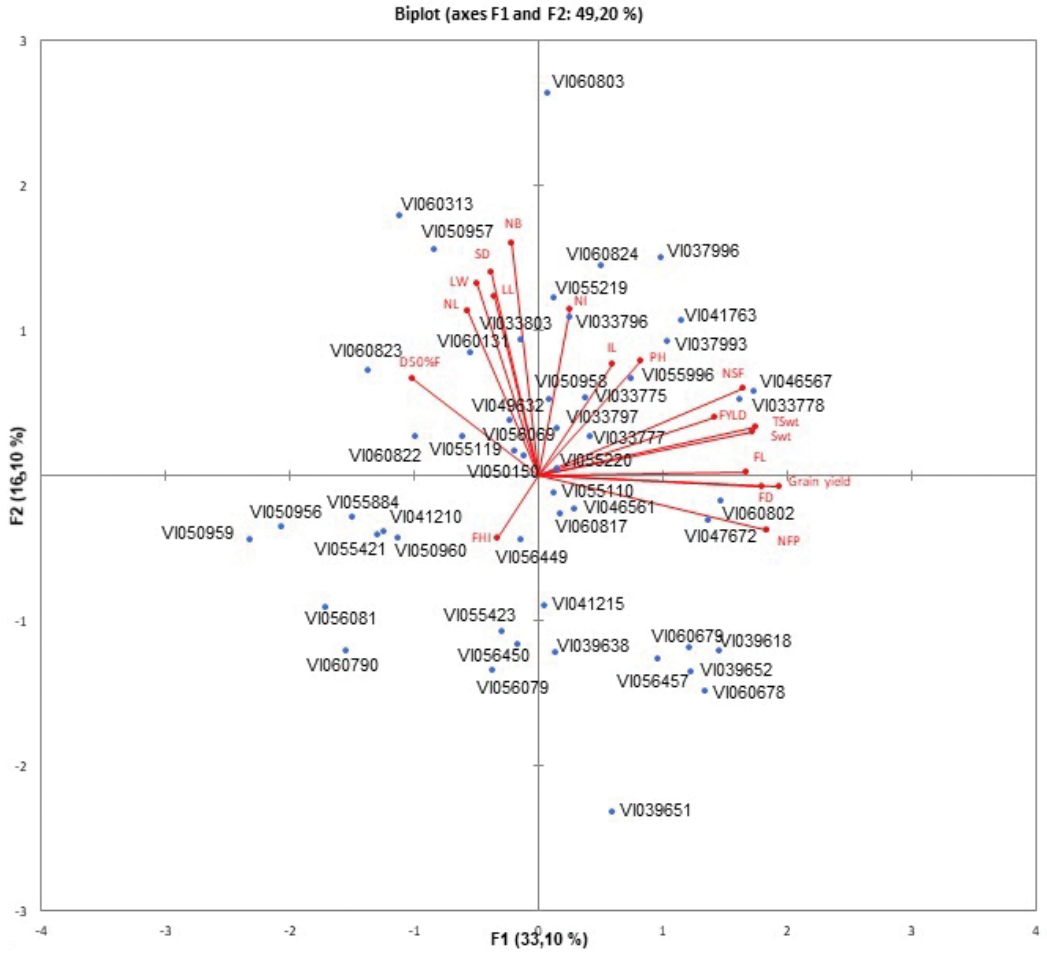

Figure 3. Scatter plot of the first and second principal component analysis for morphological phenotypic traits and okra genotypes.

elements and protein traits into 12 components, which accounted for the entire (100%) genetic variability among the evaluated okra genotypes. According to Chatfied and Collins [58, 61], components with an eigenvalue of less than one should be removed so that fewer components with significant meanings are considered. Furthermore, Ref. [56] suggested that eigenvalues greater than one are considered significant and component loadings greater than ± 0.3 were considered meaningful. Hence, from this study, as it can be seen clearly that only the first three eigenvectors which had eigenvalues greater than one and cumulatively explained about 69.04% of the total variation by the first, second and third principal components in the whole data set for the genotypes and provide discriminatory information in respective to the mineral elements and protein. The first principal component, that is the PC1 alone describes and explains 33.69% of the total variability among the okra genotypes, which was mainly contributed by the variances due to K, P, Mg, Zn, Mn and Cu (**Table 6**) with positive loading. The second principal component

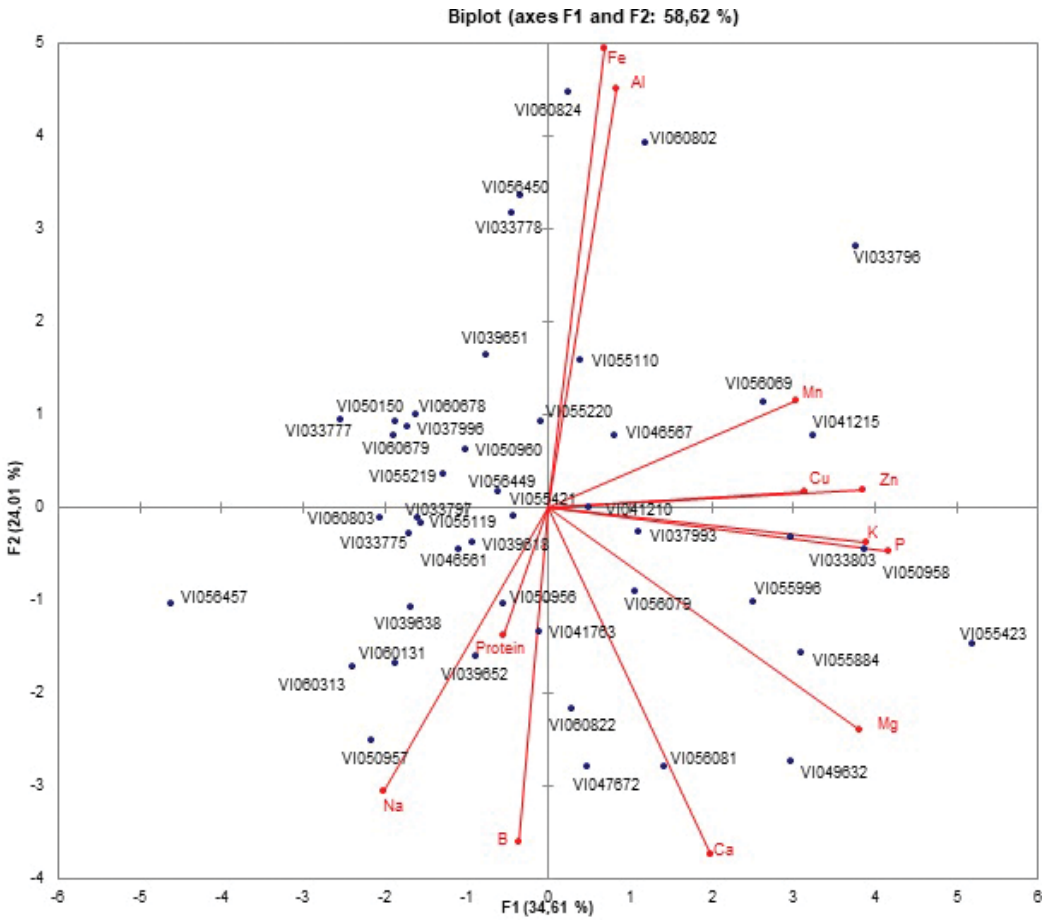

Figure 4. Biplot generated using the concentration of mineral elements and protein content data set of okra genotypes.

(PC2) represents and accounted for 23.89% with eigenvalue of 2.87 had dominantly influenced by the mineral elements such as Ca, Na, Fe, Al, and B with the highest loading vector of Fe followed by Al with positive loading. The nutritional trait that contributed great variability among the genotypes showing 11.45% of variation were protein with the highest positive loading. The mineral elements Ca, Na, B, and Mn also contributed differences in this PC.

The existence of wider nutritional variability among okra genotypes studied was further described by the PCA biplot (**Figure 2**) using multivariate technique. The PCA biplot provided important information regarding the similarities as well as the pattern of differences among the nutritional traits of the different okra genotypes and of the interrelationships between the quantified nutritional traits. The PCA clustered the okra genotypes into different groups over the four quadrants based on the nutritional traits determined (**Figure 2**). The okra genotypes scattered in all four quadrants on the axes, indicating that there were a wide genetic variability for the traits studied. Accessions that overlapped and closer to each other in the principal component axes had similar genetic relationships in the nutritional traits. However, genotypes which are far from each other could be considered as genetically

distinct [54]. The okra genotypes in the top right quadrant were closely associated with the mineral elements such as Al, Fe, Cu, Zn, and Mn (**Figure 2**). The right bottom quadrant consists of the okra genotypes that are closely related with the mineral elements such as K, Mg, P, and Ca. Those genotypes that found on the left bottom quadrant were mostly associated with the low concentration of mineral elements such as Na and B and protein content. In the present study, the genotypes VI056457, VI033796, VI060824, VI060802, VI055423, and VI049632 stand out clearly as the most genetically divergent okra genotypes for the nutritional traits evaluated. This indicated that they might have a peculiar gene/allele that separated them from the group of the genotypes assessed for the nutritional composition and could be used as parental genotypes for hybridization to develop new cultivar for the traits of interest in our breeding programme (**Figure 4**).

4. Conclusion

In the present study, the existence of genetic variability in the morphological, phenotypic and nutritional traits would help the breeder in selection of the okra genotypes for the improvement for these traits, which would help to increase the frequency of favorable genes in the pre-breeding programme. This is a first pre-breeding programme of okra in South Africa established recently as a prerequisite for the development of new cultivar in the country and beyond for yield and nutritional quality. The okra genotypes in this study showed enormous phenotypic and nutritional variations that would help in the okra improvement programme. The significant positive association between grain yield and yield traits as well as nutritional quality traits could be used as selection criteria for potential and good parental lines in okra breeding programme in South Africa. Understanding and the knowledge of variability and trait association in this study is important in the okra-breeding programme as an initial step to develop new cultivar for the traits of interest. To my best knowledge, this is the first study on this under-utilized fruit vegetable crop species in South Africa that would contribute to food, nutritional and health security.

Acknowledgements

Agricultural Research Council (ARC) and National Research Foundation (NRF) for research and funding opportunity, South Africa. The author would also like to acknowledge the Genetic Resources and Seed Unit, World Vegetable Center (AVRDC), Taiwan for providing the okra germplasm for the study.

Author details

Abe Shegro Gerrano

Address all correspondence to: agerrano@arc.agric.za

Agricultural Research Council, Vegetable and Ornamental Plants, Pretoria, South Africa

References

[1] Dhankhar BS, Mishra JP. Origin, history and distribution. In: Dhankhar BS, Singh R, editors. Okra Hand Book of Global Production, Processing, and Crop Improvement. New York, NY: HNB Publishing; 2009. p. 10024

[2] Kochhar SL. Okra (Lady's finger). In: Kochhar SL, editor. Tropical Crops, a Textbook of Economic Botany. 1986. pp. 263-264

[3] Thakur MR, Arora SK. Okra. In: Bose TK, Som MG, editors. Vegetable Crops in India. Calcutta: Naya Prokash; 1986. pp. 610-618

[4] Schippers RR. African Indigenous Vegetable an Overview of the Cultivated Species. London, United Kingdom, ISBN: 9780859545150: University of Greenwich, National Resources Institute; 2000. p. 214

[5] Saifullah M, Rabbani MG. Evaluation and characterization of okra (*Abelmoschus esculentus* L. Moench.) genotypes. SAARC Journal of Agriculture. 2009;**7**:92-99

[6] Sathish D, Eswar A. A review on: *Abelmoschus esculentus* (okra). International Research Journal of Pharmaceutical and Applied Sciences. 2013;**3**:129-132

[7] Tindall HD. Vegetables in the Tropics. London, UK: Macmillan Education Limited; 1983

[8] Lamont WJ. Okra-A versatile vegetable crop. Horticultural Technology. 1999;**9**:179-184

[9] Gemede HF, Ratta N, Haki GD, Woldegiorgis AZ, Beyene F. Nutritional quality and health benefits of okra (*Abelmoschus esculentus*): A review. Journal of Food Processing and Technology. 2015;**6**:458

[10] Gemede WF, Haki GD, Beyene F, Woldegiorgis AZ, Rakshit SK. Proximate, mineral, and antinutrient compositions of indigenous okra (*Abelmoschus esculentus*) pod accessions: Implications for mineral bioavailability. Food Science & Nutrition. 2016;**4**(2):223-233

[11] Habtamu FG, Negussie R. Antinutritional factors in plant foods: Potential health benefits and adverse effects. International Journal of Nutrition and Food Sciences. 2014;**3**:284-289

[12] Woolfe ML, Chaplin MF, Otchere G. Studies on the mucilages extracted from okra fruits (*Hibiscus esculentus* L.) and baobab leaves (*Adansonia digitata* L.). Journal of the Science of Food and Agriculture. 1977;**28**(6):519-529

[13] Ndunguru J, Rajabu AC. Effect of okra mosaic virus disease on the above-ground morphological yield components of okra in Tanzania. Scientia Horticulturae. 2004;**99**:225-235

[14] Moekchantuk T, Kumar P. Export Okra Production in Thailand. Bangkok, Thailand: Inter-country programme for vegetable IPM in South and SE Asia phase II Food and Agriculture Organization of the United Nations; 2004

[15] Doijode SD. Seed Storage of Horticultural Crop. New York, USA: Food Product Press; 2001

[16] Akinyele BO, Temikotan T. Effect of variation in soil texture on the vegetative and pod characteristics of okra (*Abelmoschus esculentus* (L.) Moench). International Journal Agriculture Research. 2007;**2**:165-169

[17] Savello PH, Martin FW, Hill J. Nutritional composition of okra seed meal. Journal of Agricural and Food Chemistry. 1980;**28**:1163-1166

[18] Kumar S, Dagnoko S, Haougui A, Ratnadass A, Pasternak D, Kouame C. Okra (*Abelmoschus spp*.) in west and Central Africa: Potential and progress on its improvement. African Journal of Agricultural Reserch. 2010;**5**:3590-3598

[19] Kumar P, Nenes A, Sokolik IN. Importance of adsorption for CCN activity and hygroscopic properties of mineral dust aerosol. Geophysical Research Letters. 2009;**36**(24)

[20] Yang RY, Keding GB. Nutritional contributions of important African indigenous vegetables. In: Shackleton CM, Pasquini MW, Drescher AW, editors. African Indigenous Vegetables in Urban Agriculture. London: Earthscan; 2009. pp. 105-143

[21] Hilou A, Nacoulma OG, Guiguemde TR. In vivo antimalarial activities of extract from *Amaranthus spinosus* L., and *Boerhaavia erecta* L. Journal of Ethnopharmacology. 2006; **103**:236-240

[22] Merriam-Webster. Webster's Ninth New Collegiate Dictionary. Springfield, Massachusetts, USA: Merriam-Webster Inc., Publisher; 1991

[23] IBPGR. International Board for Plant Genetic Resources, Genetic Resources of Abelmoschus Genetic Resources. Rome, Italy: Crop Genetic Resources Center; 1984

[24] Matejovic I. Total nitrogen in plant material determined by means of dry combustion: A possible alternative to determination by Kjeldahl digestion. Communications in Soil Science and Plant Analysis. 1995;**26**:2217-2229

[25] Jimenez RR, Ladha JK. Automated elemental analysis: A rapid and reliable but expensive measurement of total carbon and nitrogen in plant and soil samples. Communications in Soil Science and Plant Analysis. 1993;**24**:1897-1924

[26] Zasoski RJ, Burau RG. A rapid nitric-Perchloric acid digestion method for multi-element tissue analysis. Communications in Soil Science and Plant Analysis. 1977;**8**:425-436

[27] Agronomix. Agrobase Generation II. Agromix Software. 2008; Winnipeg, Manitoba, R2N054, Canada

[28] NCSS. Number Cruncher Statistical Systems. 2004; Jerry L. Hintze, 329 North 1000 East, Kaysville, Utah 84037, Canada

[29] De Vicente MC, Guzmán FA, Engels J, Ramanatha Rao V. Genetic Characterization and its Use in Decision Making for the Conservation of Crop Germplasm. The Role of Biotechnology. 2005. 63 p; Villa Gualino, Turin, Italy 5-7

[30] Asare AT, Asare-Bediako E, Agyarko F, Taah K, Osei EO. Phenotypic traits detect genetic variability in okra (*Abelmoschus esculentus*. L. Moench). African Journal of Agricultural Research. 2016;**11**:3169-3177

[31] Verma VD. Collecting eggplant and okra in Madhya Pradesh and Maharashtra. IBPGR Newsletter for Asia Pacific and Oceania. 1993;**13**:14-15

[32] Eshiet AJ, Brisibe EA. Morphological characterization and yield traits analysis in some selected varieties of okra (*Abelmoschus Esculentus* L. Moench). Advances in Crop Science and Technology. 2015;**3**:197. DOI: 10.4172/2329-8863.1000197

[33] Henshaw FO. Varietal differences in physical characteristics and proximate composition of Cowpea (*Vigna unguiculata*). World Journal of Agricultural Sciences. 2008;**4**:302-306

[34] Ng'uni D, Geleta M, Hofvander P, Fatih M, Bryngelsson T. Comparative genetic diversity and nutritional quality variation among some important southern African sorghum accessions [Sorghum bicolor (L.) Moench]. African Journal Crop Science. 2012;**6**:56-64

[35] Shegro A, Shargie NG, Van Biljon A, Labuschagne MT. Diversity in starch, protein and mineral composition of sorghum landrace accessions from Ethiopia. Journal of Crop Science and Biotechnology. 2012;**15**:275-280

[36] Gerrano AS. Preliminary study on variability and heritability estimates of micronutrient composition in the immature fruits of okra (*Abelmoschus esculentus*) genotypes in South Africa. Cogent Food and Agriculture. 2017;(3):1408253

[37] Rolfes SR, Pinna K, Whitney E. Understanding Normal and Clinical Nutrition. 9th ed. Vol. 2009. Cengage Learning. p. 1184

[38] Jimoh WLO, Abdullahi MS. Analysis of some mineral elements in selected sorghum cultivars. Bayero Journal of Pure and Applied Sciences. 2011;**4**:56-58

[39] Yellen G. The voltage gatted potassium channels and their relatives. Nature. 2002:35-42

[40] Olatunji EMA, Temitope BM. Comparative study of mineral elements distribution in sorghum and millet from Minna and Bida, North Central Nigeria. International Journal of Food Nutrition and Safety. 2013;**3**:33-55

[41] Kamga RT, Kouamé C, Atangana AR, Chagomoka T, Ndango R. Nutritional evaluation of five African indigenous vegetables. Journal of Horticultural Research. 2013;**21**:99-106

[42] Samia MA, Hagir BE, Wisal IH, Elfadil EB, Abdullahi HE. Proximate composition, antinutritional factors and mineral content and availability of selected legumes and cereals grown in Sudan. Journal of Food Technology. 2005;**3**:511-515

[43] Adetuyi FO, Osagie AU, Adekunle AT. Nutrient, antinutrient, mineral and zinc bioavailability of okra *Abelmoschus esculentus* (L) Moench variety. American Journal of Food and Nutrition. 2011;**1**:49-54

[44] Singh P, Chauhan V, Tiwari BK, Chauhan SS, Simon S, Bilal S, Abidi AB. An overview on okra (*Abelmoschus esculentus*) and it's importance as a nutritive vegetable in the world. International Journal of Pharma and Bio Sciences. 2014;**4**:227-233

[45] Kono K, Yoshida Y. Current aspect of trace element research in environmental health. Bulletin of the Osaka Medical College. 1989;**35**:1-9

[46] Moyin-Jesu EI. Use of plant residues for improving soil fertility pod nutrients root growth and pod weight of okra *Abelmoschus esculentum* L. Bioresource Technology. 2007;**98**:2057-2064

[47] Okareh OT, Adeolu AT, Adepoju OT. Proximate and mineral composition of plantain (Musa paradisiaca) wastes flour; a potential nutrients source in the formulation of animal feeds. African Journal of Food Science and Technology. 2015;**6**:53-57

[48] Aremu MO, Opaluwa OD, Bamidele TO, Nweze CC, Ohale IM, Ochege MO. Comparative evaluation of nutritive value of Okro (*Abelmoschus esculentus*) and bush mango (*Irvingia gabonensis*) fruits grown in Nasarawa state, Nigeria. Food Science and Quality Management. 2014;**27**:63-70

[49] Kalra CL, Pruthi JS, Teotia MS, Raina BL. Influence of variety, size, grade and number of ridges of okra (*Abelmoschus esculentus* (L.) Moench) on the canned okra. Indian Food Packer. 1982;**36**:53-57

[50] Sohal BS. Biochemical constituents. In: Dhankhar BS, Singh R, editors. Okra Hand Book of Global Production, Processing, and Crop Improvement. New York, NY: HNB Publishing; 2009. p. 10024

[51] Ogungbenle HN, Omosola SM. The comparative assessment of nutritive values of dry Nigerian okra (*Abelmoschus esculentus*) fruit and oil. International Journal of Food Science and Nutrition Engineering. 2015;**5**:8-14

[52] Anwar F, Rashid U, Mahmood Z, Iqbal T, Sherazi TH. Inter-varietal variation in the composition of okra (*Hibiscus esculentus* L.) seed oil. Pakistan Journal of Botany. 2011;**43**:271-280

[53] Gerrano AS, Adebola PO, Jansen van Rensburg WS, Venter SL. Genetic variability and heritability estimates of nutritional composition in the leaves of selected cowpea genotypes [*Vigna unguiculata* (L.) Walp.]. HortScience. 2015;**50**:1435-1440

[54] Gerrano AS, Jansen van Rensburg WS, Adebola1 PO. Nutritional composition of immature pods in selected cowpea [*Vigna unguiculata* (L.) Walp.] genotypes in South Africa. Australian Journal of Crop Science. 2017;**11**:134-141

[55] Jiang SL, JG W, Feng Y, Yang XE, Shi CH. Correlation analysis of mineral element contents and quality traits in milled rice (*Oryza stavia* L.). Journal of Agricultural and Food Chemistry. 2007;**55**:9608-9613

[56] Hair JF, Andrson JR, Tatham RE, Black WC. Multivariate Data Analysis. 5th ed. London: Prentice-Hall international, Inc; 1998

[57] Aslam M, Maqbool MA, Zaman QU, Shahid M, Akhtar MA, Rana AS. Comparison of different tolerance indices and PCA biplot analysis for assessment of salinity tolerance in lentil (Lens culinaris) genotypes. International Journal of Agriculture and Biology. 2017; **19**:470-478

[58] Chatfied C, Collin AJ. Introduction to Multivariate Analysis. New York, NY: Published in the USA by Chapman and Hall in Association with Methuen, Inc.; 1980. p. 733

[59] Campos ET, Espinosa MAG, Warburton ML, Monter AV. Characterisation of mandarin (Citrus spp) using morphological and AFLP markers. Interciencia. 2005;**30**:1-14

[60] Ogunbayo SA, Ojo DK, Guei R, Oyelakin OO, Sanni KA. Phylogenetic diversity and relationships among 40 rice accessions using morphological and RAPDs techniques. African Journal of Biotechnology. 2005;**4**(11):1234-1244

[61] Kaiser HF. The application of electronic computers to factor analysis. Educational and Psychological Measurement. 1960;**20**:141-151

Chapter 6

Genetic Potential and Usefulness of Native Maize Populations in Developing Novel Germplasm for Current and Upcoming Goals

Froylan Rincon-Sanchez and Norma A. Ruiz-Torres

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.71360

Abstract

Traditional agricultural system is referring to the maize production based on indigenous or farmers knowledge and practices that have been developed through many generations. In the area of study, genetic maize diversity was explored by the expression of quantitative traits of the ear and the race classification approach. Evaluation results indicated that the native populations adapted to the transition and highland (above 2000 masl) areas, showed a contrasting yield response when they were evaluated at the intermediate environment; whereas, those populations adapted to the lowland and intermediate altitudes showed a satisfactory yield performance in both environments. The above performance pattern is essential because it may be useful to identify favorable alleles that, in a local population *per se* or through genetic combination, results in population changes in allele frequencies that could mitigate the effects of climate changes, particularly in maize populations adapted to highland altitudes. Selection procedures applied to a local adapted population can be managed attending different goals, including the conservation of genetic diversity (*per se* selection), and to develop novel germplasm. The introgression of foreign germplasm into a local population and the application of three selection cycles resulted in a novel variety (JAGUAN) adapted to a regional northeast Mexico environmental conditions.

Keywords: *Zea mays* L., native maize populations, genetic diversity, genetic by environmental interaction, selection procedures

1. Introduction

The maize (*Zea mays* L.) is a native crop of Mexico adapted to the most diverse environmental conditions; planted around the country in altitudes ranging from sea level to altitudes

greater than 2550 m. In Mexico, 78.3% of the maize production is sown under rain-fed conditions, mainly with native (landrace) adapted varieties [1]. Those varieties are usually grown in non-optimum agronomic conditions, therefore they are adapted to variable rainfall, and are in some extent, tolerant to biotic and abiotic stresses. These varieties are available to the farmers for sowing, due to their flexible response to adverse situations, and are frequently used for seed exchange among farmers, within the same community or with other communities. In the Coahuila state, 24,900 ha of maize for grain production were sown during 2016, 84.7% was sown in the southeast region, mainly with local adapted populations (landrace populations), and 94.8% of these were sown under rain-fed conditions [1]. Typically, the native maize production is mainly for local consumption, both human and livestock as forage.

The area of study is located in the southeast of the Coahuila state in Mexico, and it is represented by five counties (Arteaga, General Cepeda, Parras, Ramos Arizpe, and Saltillo). Coahuila state is situated in the central part of the North of Mexico, with a territorial area of 151,571 km^2. The climate in the state is dry to very dry, semi warm (75% of total area), average temperature ranging from 18 to 20°C; annual average precipitation of 316 mm. Based on the environmental conditions, the region of study is considered as critical, determined by an average annual precipitation ranging between 350 and 450 mm; average temperature of 16.8°C; with presence of drought and frost seasons in the year. Moreover, to the environmental and ecological conditions in the region, the maize diversity is determined by the adaptation, and genetic combinations among race complexes allowed by seed exchange within and among different communities [2, 3].

Exploring and understanding the genetic potential of adapted cultivars on traits of interest may determine and guide further research for a particular environment or crop system, as well as the efficient use of both economic and human resources. Thus, the objectives of research work were to describe the regional maize genetic diversity, determine the genetic potential of locally adapted maize populations and to identify strategies for crop improvement to resolve current and future aims.

2. Regional maize genetic diversity

In any crop system, the genetic diversity is determined—among other factors—by the use of diverse types of local varieties. Conceptually, two types of genetic materials are commonly developed, those obtained from preferences (color, flavors, crop type, etc.) and those selected for adaptability to biotic or abiotic micro-environments [4]. In addition to the native population for a particular environment, maize genetic diversity is associate to other factors that may change the genetic structure in a population such as the seed exchange among farmers within a community or among different communities, and, depending on the migration index and introgression of foreign germplasm, would contribute to the genetic variation in a native population [5, 6].

http://dx.doi.org/10.5772/intechopen.71360

The maize (*Z. mays* L.) is a native crop of Mexico, and it is the place where the highest genetic diversity is found. This crop is adapted to the most diverse environmental conditions; thus, the specific local adapted populations (landrace) have been developed, with particular attributes that differentiate each other, within and among regions, circumstances that make possible to recognize Mexico as center of origin and diversification [7]. Commonly, the maize diversity has been described by the racial classification approach, which allowed to identify the first 25 races of maize in Mexico, based on morphological data (plant, ear, and tassel) [8].

In the region of study, the maize diversity has been documented by the presence of representative race populations in Coahuila state, such as Tuxpeño [8], Raton and Tuxpeño Norteño [9], Celaya, Conico Norteño, Elotes Conicos, and Olotillo [10]. A case study carried out in native populations from Coahuila State in Mexico, indicated that genetic diversity shows a continuous pattern among racial complexes, and is associated to the altitudinal and ecological regions [11].

2.1. Relationship among local populations

In this section, several quantitative traits of the ear and grain were considered to analyze the maize genetic diversity and the relationship among local populations within the southeast region of the Coahuila state in Mexico. The racial classifications and relationship among the native adapted populations (landraces) were studied with a sample of 77 maize populations that were collected in the region from altitudes ranged from 774 to 2557 masl. A total of 51 of these populations were collected in 2008 [10] and 26 during 2010 (unpublished data), which represents the maize genetic diversity in the region of study.

Sample sizes of 10 representative ears were first used for a visual classification of the maize populations based on the primary race classification [8]. In addition to the race classification, a set of quantitative traits from the ear and grain were used to analyze the relationship among the native maize populations. Several authors have emphasized that the reproductive organ traits such as the ear traits, are the most useful for race classification in maize [12, 13]. Thus, eight racial complexes were identified: Celaya, Conico Norteño, Elotes Conicos, Elotes Occidentales, Olotillo, Raton, Tuxpeño, and Tuxpeño Norteño. At the same time, maize populations were grouped by an altitudinal stratum: lowland (0–1000 m), intermediate (1001–1800 m), transition (1801–2000 m), and highland (above 2000 m) (**Table 1**).

Ten quantitative ear and grain traits were obtained from the collected sample to analyze the maize diversity. Five ear traits: ear and cob diameter (EAR_DIAM, COB_DIAM) (cm), ear length (EAR_LENG) (cm), ear rows (EAR_ROWS), shelling percent (SHELL_PCT), and five kernel traits: Kernel measurements such as kernel length (KER_LENG), width (KER_WIDTH) and thickness (KER_THICK) (mm), kernel per row (KER_PER_ROW), weight of 100 dry kernel (WT_100_KER) (g) [14]. Data were explored by the analysis of variance to test adaptation groups and racial complexes differences. In both cases, populations within groups and populations within races were analyzed using the PROC GLM procedure of SAS [15]. Data means were used to explore maize diversity by principal component analysis using the quantitative traits as testers [16].

Race classification	Race ID	Lowland (<1000)	Intermediate (1001–1800)	Transition (1801–2000)	Highland (>2000)	Total
Celaya	C		4			4
Cónico Norteño	CN		3	8	23	34
Elotes Cónicos	EC		1	1	1	3
Elotes Occidentales	EO		2			2
Olotillo	O		1			1
Ratón	R	4	16	1		21
Tuxpeño	T		1			1
Tuxpeño Norteño	TN	1	9	1		11
Total		5	37	11	24	77

Table 1. Racial classification of local native populations from the Southeast of Coahuila State in Mexico.

The races Conico Norteño and Elotes Conicos (the ear conical type) are adapted from the transition to the highland areas, in altitudes above 1700 m; whereas, Raton, Tuxpeño, and Tuxpeño Norteño (the ear cylindrical type), the adaptation area is widely: Raton (84–1300 m), Tuxpeño (0–1950 m), and Tuxpeño Norteño (1400–1701) [17]. The maize diversity in the Coahuila state is represented mainly by three racial complexes: Conico Norteño, Raton, and Tuxpeño Norteño [10].

The eight racial complexes and the four adaptation groups were statistically different ($P \leq 0.01$) for most traits, indicating relative differences among the race type and the adaptation of landrace populations in the region; the populations within racial complexes and populations within adaptation groups were significant ($P \leq 0.01$), indicating the variation associated within race groups, the genetic combination among populations and race complexes (**Table 1**), and the specific adaptation to the different ecological environments within the region.

The scatter plot of the interaction among the 77 native maize populations with the 10 quantitative traits is presented in **Figure 1**.

The maize populations and the traits studied are all distributed along **Figure 1**, where individual points (maize population or traits) reached a vector from the origin indicates the joint association that makes the distinctions among maize populations and the relationship with the associated traits. The group of populations indicated by the dashed oval corresponds to the Conico Norteño and Elotes Conicos, two races adapted to highland altitude, characterized by a conical ear type. Populations outside the oval show the cylindrical ear type represented mainly by the races Raton, Tuxpeño, and Tuxpeño Norteño (**Table 1**). By the exploration of dispersion of these population × racial groups in **Figure 1**, it is possible to detect a continuous pattern among the races Raton and Tuxpeño Norteño. Similar genetic variation pattern was found among racial groups through the landraces analyses in the state of Coahuila [11].

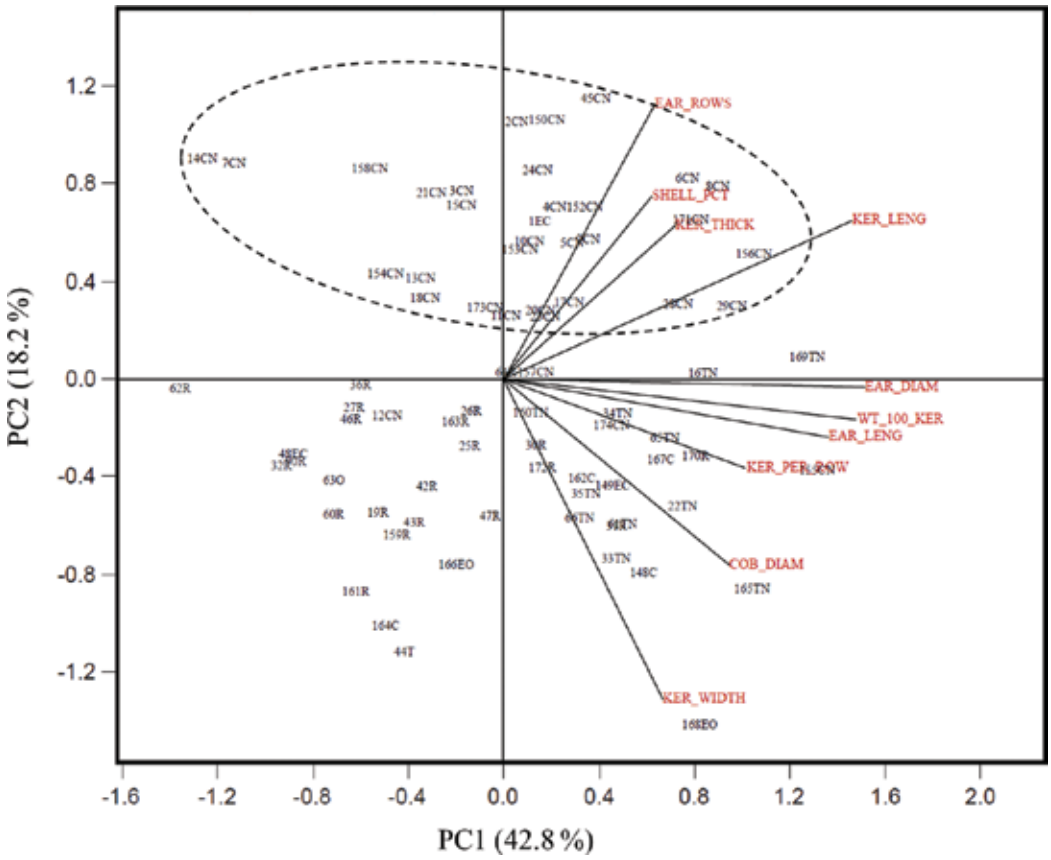

Figure 1. Scatter plot of the first two principal component scores showing the dispersion of the 77 maize populations and relationship among traits. Populations are identified by the population number and combinations of letters that indicate the race classification: C = Celaya; CN = Conico Norteño; EC = Elotes Conicos; EC = Elotes Occidentales; O = Olotillo; R = Raton; T = Tuxpeño; and TN = Tuxpeño Norteño.

Dispersion in **Figure 1** shows the identification of two main groups: the group of maize populations with the conical ear type (transition and highland altitudes) and those with the cylindrical ear type (lowland to intermediate altitudes) (**Table 1**). This indicates the usefulness of the ear and grain traits in exploring the association among maize populations as suggested by [12, 13]. The continuous pattern among the different races within the two main groups is explained by the adaptation of the local populations (**Table 1**), and the genetic combination among races as a consequence of seed exchange by farmers within the community and with other communities [5, 6]. The seed exchange among farmers is a common practice in the area of study, is evident by the presence of some maize populations located at a different adaptation area (**Table 1**), explained by the genetic combinations among different races, and further adaptation to micro-environments as was verified in several regional studies [3, 11].

3. Genetic potential of maize populations

Maize genetic diversity accounted by the locally adapted populations (landraces) within a traditional agricultural system is not static; it varies from 1 year to the another as a consequence of many factors such as migration (seed, pollen), selection, genetic drift [18], and adaptation to changes and interactions with external factors within the ecosystem development, as part of the evolutionary process and selection [19]. Thus, genetic variation is closely related to the environmental and production conditions and to the different uses of the crop, in particular the grain (color and flavor). The knowledge and understanding of the genetic variation, the environmental interaction and its potential use, may determine both the genetic conservation strategy and the possible utilization in breeding programs to improve local populations or for developing novel germplasm for particular goals.

3.1. Yielding potential and environmental response

The maize genetic diversity in the Coahuila state was determined initially by the description of 90 native maize populations that were collected during 2008 [10]. At the same time, those populations were established on field experiments for agronomic evaluation for 2 years (2008–2009), at two contrasted locations to determine the grain yielding potential. The agronomic evaluation was conducted in: El Mezquite, Galeana, Nuevo Leon (1890 masl), and General Cepeda, Coahuila (1350 masl). These locations are representative of both, the highland and intermediate environmental conditions in the area of study. The combination of two locations and 2 years of evaluation was named as four different environments. To analyze the environment response, the native populations were grouped based on the adaptation altitude: lowland (0–1000), intermediate (1001–1800), transition (1801–2000), and highland (greater than 2000 masl).

The local population × environment interaction analysis allowed to identify three groups that describe the specific adaptation of the maize populations [20]: the first one with adaptation to El Mezquite (33.3%), the second adapted to General Cepeda (42.2%), and a third group (24.4%) with an average yielding potential across environments, indicating a form of stability [16, 21]. In this study, it was shown that the racial types are associated to the locations of evaluation: the race Conico Norteño to El Mezquite (highland) and the races Raton, Tuxpeño, and Tuxpeño Norteño to the General Cepeda site (intermediate), which are also associated to the adaptation origin as indicated in **Table 1**. In Mexico, the maize genetic diversity is associated to the agro-ecological conditions that determine the different races types and their ear and grain distinctiveness and uses [17]. Results of the study determined the yielding potential and population response to the environmental evaluation, outstanding the racial groups Tuxpeño, Tuxpeño Norteño, and Raton with the highest yielding potential.

The assigned groups of native maize populations and the groups × environments were statistically different ($P \leq 0.01$), explained by the diversity in altitudinal origin of maize populations and the differential response on the evaluation environments. A relative comparison of the average grain yield of the groups of native populations evaluated in the two contrasted environments during 2 years is presented in **Figure 2**.

http://dx.doi.org/10.5772/intechopen.71360

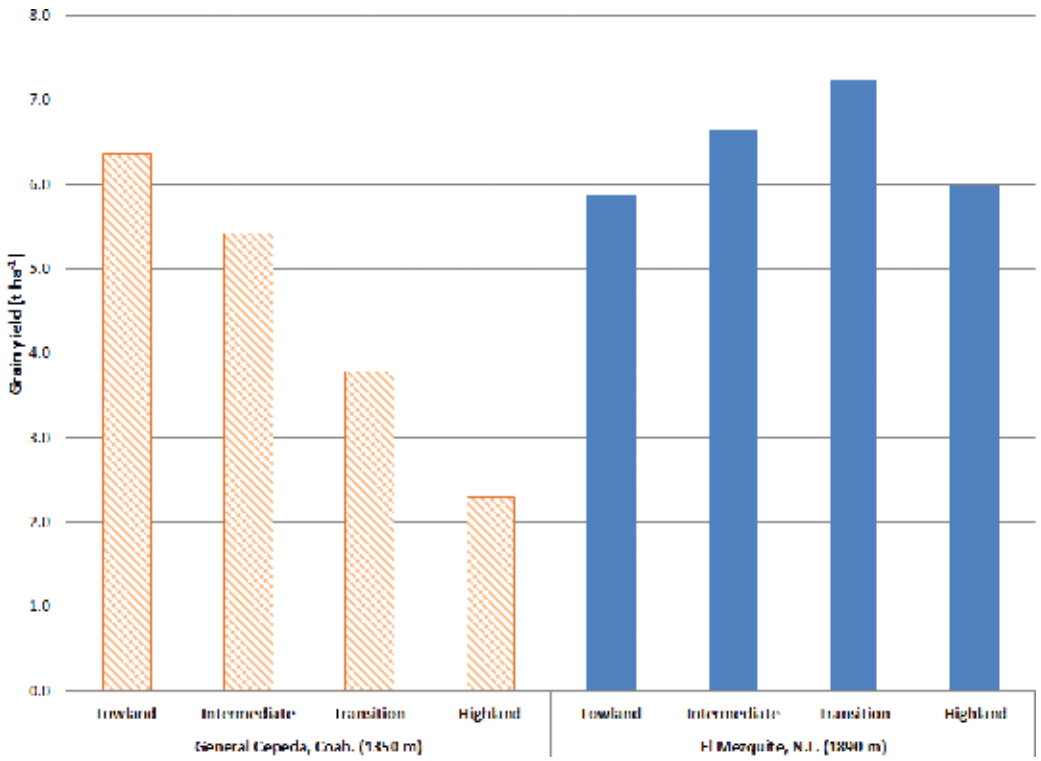

Figure 2. Groups of native maize populations by the environment interaction based on grain yield.

The maize populations represented in the transition and highland groups, showed a contrasted yielding response when evaluated at the General Cepeda area (intermediate altitude), in comparison with El Mezquite environment; whereas those populations adapted to the lowland and intermediate altitudes showed an adequate yield performance in both environments. Similar response pattern has been reported by [22], who also mentioned that populations adapted to highland areas have a difficult performance when established at lowland altitudes; on the other hand, those populations with an adaptation to lowland to intermediate areas have an acceptable agronomic performance. Thus, populations from lowland and intermediate altitudes have better adaptation range with satisfactory yielding potential. The above performance pattern is important because it may be useful to identify favorable alleles that, in a local population *per se* or through genetic combination among different racial complexes, resulting in population changes in allele frequencies controlling traits of interest, consequently, it could mitigate the effects of climate changes, particularly in maize populations adapted to highland altitudes [23].

In a different study, carried out in the southeast of the Coahuila state, an agronomic evaluation of native maize populations was performed in 2013 (unpublished data). The objectives of the research work were to determine the agronomic performance and yield potential of local maize populations, and to define the area of adaptation using two contrasting and

representative environments of the southeast of the Coahuila state in Mexico. The agronomic evaluation of 63 maize populations and 7 improved checks was carried at 2 locations and 2 replications (blocks) under irrigation conditions: El Mezquite, Galeana, N. L. (1890 masl) and General Cepeda, Coah. (1350 masl). The combination of two locations and two replications was named as four different environments (GC1, GC2, MEZ1, and MEZ2). In both locations, replications were established independently, and in General Cepeda, the two replications represented two planting dates. The genetic diversity was represented by eight racial complexes: Celaya (3), Conico Norteño (26), Elotes Conicos (4), Elotes Occidentales (1), Olotillo (3), Raton (16), Tuxpeño (6), and Tuxpeño Norteño (4). The improved materials used as checks have variability on maturity and grain type: an experimental variety (POBAM), two improved varieties (VAN210 and JAGUAN), and four synthetic populations (6221, 6222, Pool31, and Pool32). The yield potential was analyzed across environments and the genotype × environment interaction based in the model of the additive main effects and multiplicative interaction (AMMI) [24].

The analysis of variance showed differences ($P \leq 0.01$) among environments, genotypes (populations and checks) and genotype × environment interaction. Among the 25 outstanding populations, the racial groups with higher yield grain potential correspond mainly to the Raton races (9 populations), Tuxpeño (6 populations), and Tuxpeño Norteño (4 populations). Also, there were five native maize populations adapted to intermediate areas: three of the Tuxpeño race (I38T, I52T, and I54T) and two of the Raton race (I13R and I40R) with similar yields to the best improved check. In a previous study, the races Raton, Tuxpeño, and Tuxpeño Norteño had also the highest yield potential [20].

In the environmental response analysis based on the AMMI model, (**Figure 3**) shows the main effects for grain yield (Genotype and Environment) on the abscissa axis, and the first interaction principal component (IPC1) on the ordinate axis.

The AMMI model (**Figure 3**) allowed identifying genotypes with specific adaptation to the two contrasting environments, and the average response across environments. For instance, using an approximate range of $-0.25 \leq 0 \leq 0.25$ of the IPC1 as criteria to identify those genotypes with an average performance across environments, represents a form of stability stability [21]. There were five populations identified with good yield potential and a type of stability across environments: two Conico Norteño (H28CN and H27CN), one of Tuxpeño Norteño (I59TN), and two of Raton race (I56R and I40R). Likewise, in addition to the two Conico Norteño populations (H43CN and T08CN), there were four populations with adaptation to intermediate areas, with good potential in high altitude valleys: Tuxpeño race (I52T and I54T), a population of Elotes Occidentales (I33EO), and another Raton population (I35R). Most of the populations with the highest grain yield on **Figure 3** (positive values of main effects) have adaptation to intermediate areas, and have a potential performance in the two contrasting environments and the stability as well. This pattern agrees with the results presented in **Figure 2**, suggesting that those native populations may be used in a breeding strategy, individually, or as a combination with populations adapted to highland altitudes to identify useful alleles for developing novel germplasm that could mitigate the climate change.

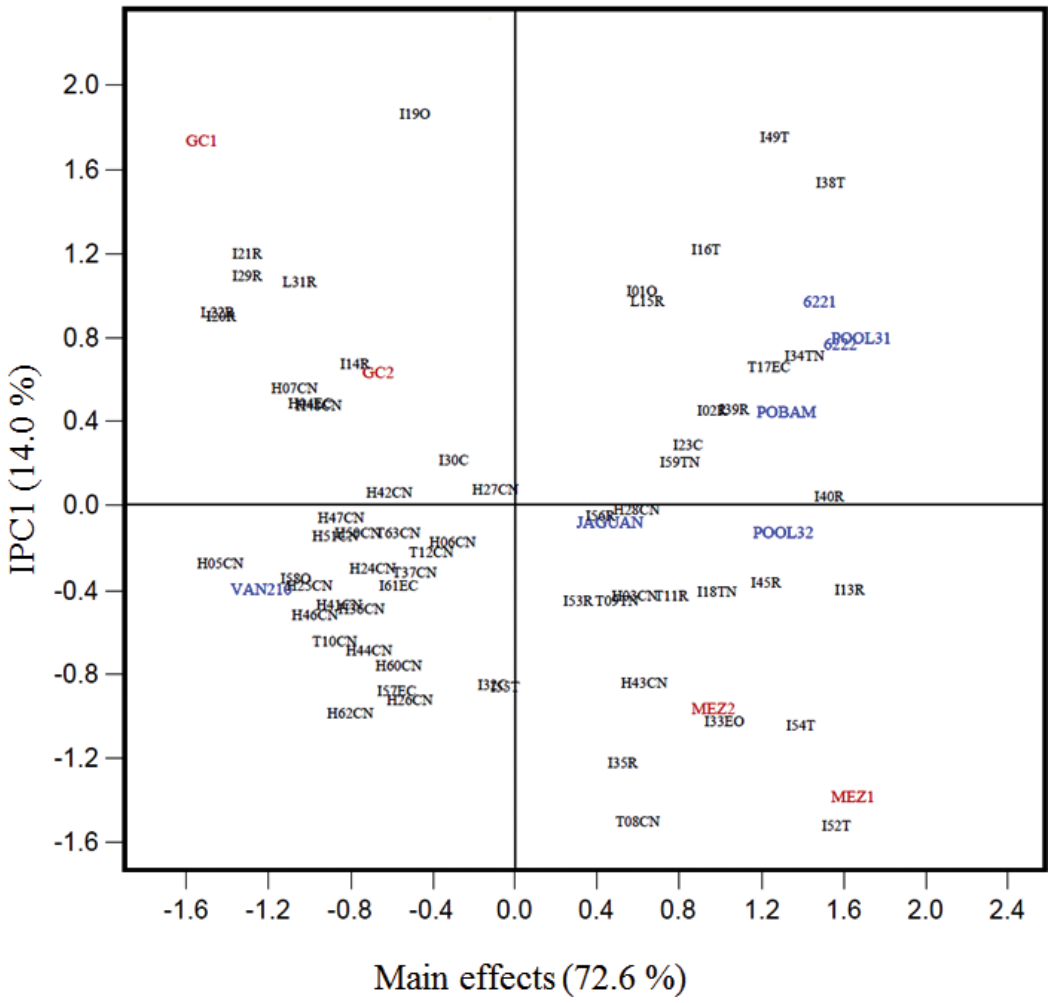

Figure 3. Scatter plot of AMMI model for grain yield of 70 maize genotypes evaluated at 2 contrasting environments (GC1, GC2, MEZ1, and MEZ2). Population's data points are indicated by combinations of letters and numbers. The first character indicates the adaptation area L = lowland, I = intermediate, T = transition and H = highland; followed by the population number and finally, the racial group: C = Celaya, CN = Conico Norteño, EC = Elotes Conicos, EO = Elotes Occidentales, O = Olotillo, R = Raton, T = Tuxpeño, and TN = Tuxpeño Norteño.

4. Strategies for crop improvement

In a traditional agricultural system, the native maize populations are developed and maintained by farmers through multiple cycles of empirical mass selection. Those populations are commonly the only source of genetic variation available for sowing, due to their flexible response to adverse situations, and usually two types of local varieties may be distinguished: the local varieties that are planted in a very small area for special uses, basically for consumption, and

represent the diversity of the crop; whereas, in the second case, varieties are planted in a larger areas, widely distributed, and are frequently used for seed exchange among farmers within the same community or with other communities. In either case, strategies for the efficient use of the germplasm need to be determined. For instance, in the first case, the recurrent selection strategy applied is basically to improve it; while, the second group, in addition to *per se* selection, those varieties are eligible for any form of genetic combination with external exotic germplasm.

4.1. Selection strategies applied to a locally adapted maize population

In the southeast of the Coahuila state, a wide adapted local maize population was identified to apply different selection strategies [25]. The native variety named JAGUEY, representative of the race Conico Norteño, is adapted to Jagüey de Ferniza, Saltillo, Coahuila (2100 masl). This variety was exposed to different management and selection procedures where four populations were obtained: (1) the original adapted population (OP); (2) the first generation from the local population (G1), obtained through a seed production scheme (detasseled rows); (3) and (4) two populations generated by the combination of the original population with an improved population, using a divergent selection for early (EM) and late (LM) maturities, respectively. After the populations were developed, a set of 25 half sib families was randomly obtained from each of the 4 populations for evaluation in 2 locations during 2003: El Mezquite, Galeana, Nuevo Leon (1890 m) and Jagüey de Ferniza, Saltillo, Coahuila (2100 m), being the irrigated and rain-fed environments, respectively. The four populations were compared to analyze the effects of selection procedures on agronomic traits using the site and the local population as references. Data were recorded for days to anthesis, plant height (m), husk cover (%), stalk and root lodging (%), moisture content of seed (%), number of ears per plant, and ear yield (t ha^{-1}) adjusted to a 15% moisture content.

Results showed significant differences ($P \leq 0.01$) among populations for most traits. A relative comparison among pairs of populations from a multivariate analysis, based on the agronomic traits evaluated is presented in **Table 2**.

A relative comparison among pairs of populations based on a multivariate analysis, showed significant differences ($P \leq 0.01$) among the OP and G1, with the two populations obtained by the introgression with improved germplasm, in the two environments evaluated (EM and LM), indicating the contribution of the improved material to the original population. On the other hand, there was not any evidence of a difference among the OP and the G1; at the same time, they were comparatively more diverse than EM and LM, as an effect of the selection methodology. Thus, populations showed significant differences in the agronomic traits, determined by both, procedures and selection applied criteria, which determine the selection strategy and management for the conservation and use of genetic diversity.

Contribution of selection methodologies after the first selection cycle, indicated by the average difference in grain yield between G1 and OP, was 1.7%; whereas, the contribution associated to the germplasm combination, the EM against OP was in the order of 24.0%. In both cases, the first cycle of selection was associated with reduction in root and stalk lodging percentages, asynchrony silk interval, husk cover, and plant and ear height, in reference to the original

Populations†	G1	EM	LM
	Jagüey, Saltillo, Coahuila (Rain-fed)		
OP	0.338	3.636**	11.396**
G1		2.850**	12.267**
EM			9.844**
	El Mezquite, Galeana, Nuevo Leon (Irrigation)		
OP	0.609	3.219**	11.258**
G1		2.806**	10.628**
EM			6.462**

†**, Significant at 0.01 probability level; OP = Original local population; G1 = Fist generation obtained through a seed production scheme; EM and LM = Early and Late maturity populations obtained through the local × improved germplasm; adapted from Rincón and Ruiz [25].

Table 2. Squared distances among pairs of maize populations based on agronomic traits evaluated at two environments.

population. These results indicate that genetic variation of local populations can be managed attending different goals including conservation of genetic variation and crop improvement by the introgression of exotic germplasm.

4.2. Potential of local × improved combination to enhance a locally adapted maize population

Based on the results of the research paper carried out by [25], plants of the local population JAGUEY were crossed with an improved population to determine the value of a breeding material to enhance a local adapted maize population. Introgression of exotic germplasm to adapted material in maize has been a powerful tool to increase genetic variability in the local population, as well as to transfer favorable alleles, such as insect or disease resistance. The proportion of the exotic germplasm has been addressed in several studies to determine the usefulness of the foreign material on the foundation of breeding base populations [26–28]. The relevance of the introgression of foreign or exotic germplasm to an adapted population may change depending on the particular objectives. For instance, the improvement of a local farmer population by the introgression of exotic or foreign germplasm requires the identification of a good source donor and the establishment of the selection strategy. However, the application of breeding techniques to local maize populations may change their genetic structure, being the level of change related to the breeding methodology and the selection pressure. Besides the improvement of the local material, it is essential to preserve as much as possible the genetic variation accounted by the local material. In this case, the maize population JAGUEY adapted to Jagüey de Ferniza, Saltillo, Coah., Mexico, was considered the local material (L), and used to assess the contribution of an improved material introgressed to the local population [25]. This local population is a white dent type of the race Conico Norteño, adapted to rainfall conditions, such as low fertilizer inputs and limited water supply, and it is maintained by farmers through an empirical mass selection. The improved material (I), considered the foreign material, was an

early flowering experimental population (CPRE), previously chosen, based on its combining ability performance when crossed with local populations [29]. Initially, full sib (FS) families were obtained throughout plant-to-plant crosses between plants from the local (JAGUEY) and the improved (CPRE) maize populations [29]. Derived families evaluation data included days to flowering (anthesis and silking), plant and ear heights (m), husk cover (%), stalk and root lodging (%), number of ear per plant, and grain yield (t ha^{-1}) adjusted to a 15% moisture content. Best families (10%) were selected based on the evaluation data under irrigated and the rain fall conditions, with special emphasis on the performance under the rainfall environment [30]. After the first genetic recombination of selected families (50:50 of local and improved germplasm), three full sib selection cycles were applied to develop an improved native maize variety named JAGUAN [31]. In addition to grain yield, the selection procedure included number of ears per plant, husk cover, with special attention in keeping the same flowering date as the original native population using selection indices [32]. The difference in grain yield between JAGUAN and the local variety JAGUEY, the original population was 24.5%. JAGUAN is a variety developed for rain-fed conditions, with intermediate biological cycle (83–90 days flowering), plant height of 2.5 m, selected by high planting densities (50,000–60,000 plants ha^{-1}), adapted to the transition to highland areas (above 1800 m), and with phenotypic expression as the original native population [29].

5. Conclusions

Regional genetic maize diversity was explored based on the race classification approach using selected ear and grain traits. Eight race complexes were identified that represent the maize diversity in the region of study. There are local populations with a wide and specific adaptation; for those wide adapted populations a crop improvement selection scheme was applied, whereas, on the specific adapted populations, a strategy for conservation and use could be implemented. Genetic introgression to a native population and further selection criteria has been useful to develop a novel variety JAGUAN adapted to regional environmental conditions. Genetic combinations among selected populations that represent the genetic diversity within the region have been identified as potential allele donors to improve genetic materials to mitigate the alterations associated to the climate change effect, particularly in populations adapted to highland altitudes.

Author details

Froylan Rincon-Sanchez[1]* and Norma A. Ruiz-Torres[2]

*Address all correspondence to: frincons@prodigy.net.mx

1 Departamento de Fitomejoramiento, Universidad Autonoma Agraria Antonio Narro (UAAAN), Buenavista, Saltillo, Coah, Mexico

2 Centro de Capacitacion y Desarrollo en Tecnologia de Semillas, Universidad Autonoma Agraria Antonio Narro (UAAAN), Buenavista, Saltillo, Coah, Mexico

References

[1] SIAP. Anuario estadístico de la producción agrícola [Internet]. 2016. Available from: http://nube.siap.gob.mx/cierre_agricola/ [Accessed: July 19, 2017 [Updated: 2016]

[2] Herrera C BE, Macias L A, Diaz R R, Vazquez R M, Delgado A A. Uso de semilla criolla y caracteristicas de mazorca para la seleccio de semilla de maiz de Mexico. Revista Fitotecnia Mexicana. 2002;**25**(1):17-23

[3] Aguirre M VJ, Rincon S F, Ramirez S R, Colom A OG, Razo M MG. Modelo para la conservacion de maices criollos en el sureste de Coahuila, Mexico. Vicente Javier Aguirre Moreno: Mexico; 2011. p. 61

[4] Eyzaguirre P, Iwanaga M. Farmer's contribution to maintaining genetic diversity in crops. In: Eyzaguirre O, Iwanaga M, editors. Participatory Plant Breeding. Wageningen, The Netherlands: IPGRI; 1996. p. 9-18

[5] Louette D. Seed exchange among farmers and gene flow among maize varieties in traditional agricultural systems. In: Serratos JA, Willcox MC, Castillo F, editors. Gene Flow among Maize Landraces, Improved Maize Varieties, and Teosintle: Implications for Transgenic Maize. Mexico: CIMMYT; 1997. p. 56-66

[6] Louette D, Smale M. Genetic Diversity and Maize Seed Management in a Traditional Mexican Community: Implications for In Situ Conservation of Maize. CIMMYT: Mexico; 1996. p. 22

[7] Kato Y TA, Mapes S C, Mera O LM, Serratos H JA, Bye B RA. Origen y Diversificacion del Maiz: Una Revision Analitica. 1st ed. Mexico: Universidad Nacional Autonoma de Mexico, Comision Nacional para el Conocimiento y Uso de la Biodiversidad; 2009. p. 115

[8] Wellhausen EJ, Roberts LM, Hernandez X E, Mangelsdorf PC. Razas de maiz en Mexico. su origen caracteristicas y distribucion. 1st ed. Mexico: Oficina de Estudios Especiales, Secretaria de Agricultura y Ganaderia; 1951. p. 237

[9] Ortega P R. Descripcion de algunas razas mexicanas de maiz poco estudiadas. In: Ortega P R, editor. Variedades y razas mexicanas de maiz y su evaluacion en cruzamientos con lineas de clima templado como material de partida para fitomejoramiento. 1st ed. URSS; 1985. p. 1-22

[10] Rincon S F, Hernandez P CJ, Zamora C F, Hernandez C JM. Recoleccion de maices nativos de Coahuila 2008. In: Rincon S F, Castillo G F, Ruiz T NA, editors. Diversidad y distribucion de los maices nativos en Coahuila, Mexico. 1st ed. Mexico: SOMEFI; 2010. p. 4-12

[11] Rincon S F, Castillo G F, Ruiz T NA, Illescas P CN, Ramon G L. Diversidad de los maices nativos de Coahuila con base en caracteres de la mazorca. In: Rincon S F, Castillo G F, Ruiz T NA, editors. Diversidad y distribucion de los maices nativos en Coahuila, Mexico. 1st ed. Mexico: SOMEFI; 2010. p. 13-25

[12] Goodman MM, Paterniani E. The races of maize: III. Choices of appropriate characters for racial classification. Economic Botany. 1969;**23**:265-273

[13] Sanchez G JJ, Goodman MM, Rawlings JO. Appropriate characters for racial classification in maize. Economic Botany. 1993;**47**:44-59

[14] CIMMYT - IBPGR. Descriptors for Maize. 1st ed. Rome, Italy: Centro Internacional de Mejoramiento de Maiz y Trigo, International Board for Plant Genetic Resources; 1991 88 p

[15] SAS Institute. SAS/STAT 9.1 User's Guide. USA: SAS Institute Inc.; 2004. p. 5121

[16] Yan W, Kang MS. GGE Biplot Analysis. A Graphical Tool for Breeders, Geneticists, and Agronomists. 1st ed. New York, USA: CRC Press L.L.C; 2003. p. 271

[17] Ruiz C JA, Duran P N, Sanchez G JJ, Ron P J, Gonzalez E DR, Holland JB, Medina G G. Climatic adaptation and ecological descriptors of 42 Mexican maize races. Crop Science. 2008;**48**:1502-1512. DOI: 10.2135/cropsci2007.09.0518

[18] Falconer DS, Mackay TFC. Introduction to Quantitative Genetics. 4th ed. England: Longman Group Ltd.; 1996. p. 464

[19] Harlan J. Crops and Man. 2nd ed. Madison, WI, USA: ASA-CSSA; 1992. p. 284

[20] Najera C LA, Rincon S F, Ruiz T NA, Castillo G F. Potencial de rendimiento de poblaciones criollas de maiz de Coahuila, Mexico. Revista Fitotecnia Mexicana. 2010;**33**(4):31-36

[21] Yan W, Tinker NA. Biplot analysis of multi-environment trial data: Principles and applications. Can. J. of Plant Sci. 2006;**86**(3):623-645. DOI: org/10.4141/P05-169

[22] Mercer K, Martinez V A, Perales RH. Asymmetrical local adaptation of maize landraces along an altitudinal gradient. Evolutionary Applications. 2008;**1**(3):489-500. DOI: 10.1111/j.1752-4571.2008.00038.x

[23] Mercer K, Perales RH. Evolutionary response of landraces to climate change in centers of crop diversity. Evolutionary Applications. 2010;**3**(5-6):480-493. DOI: 10.1111/j.1752-4571.2010.00137.x

[24] Zobel R W, Wright MJ, Gauch HG. Statistical analysis of yield data. Agronomy Journal. 1988;**80**:388-393

[25] Rincon S F, Ruiz T NA. Comparacion de estrategias e seleccion y manejo aplicadas a una poblacion criolla de maiz. Revista Fitotecnia Mexicana. 2004;**27**(1):33-37

[26] Albrecht B, Dudley JW. Evaluation of fourmaize populations containing different proportions of exotic germplasm. Crop Science. 1987;**27**:480-486

[27] Bridges WC, Gardner CO. Foundation populations for adapted by exotic crosses. Crop Science. 1987;**27**:501-506

[28] Crossa J, Gardner CO. Introgression of an exotic germplasm for improving an adapted maize population. Crop Science. 1987;**27**:187-190

[29] Rincon-Sanchez F, Ruiz-Torres NA, Leon-Castillo H, Herrera-Ayala JL. Advances on the use of maize germplasm under the participatory plant breeding approach. In: Chavez-Servia JL, Tuxill J, Jarvis DI, editors. Manejo de la diversidad de los cultivos en los agro-ecosistemas tradicionales. 1st ed. Cali, Colombia: IPGRI; 2004. p. 177-186

[30] Rincon-Sanchez F, Ruiz-Torres NA. Enhancement of a farmer's variety with an improved population in Saltillo, Coahuila, Mexico. In: Taba S, editor. Latin American Maize Germplasm Conservation: Regeneration, In Situ Conservation, Core Subsets, and Prebreeding. 1st ed. Mexico: CIMMYT; 2005. p. 58-62

[31] Rincon-Sanchez F, Ruiz-Torres NA, Cuellar-Flores R, Zamora-Cancino F. 'JAGUAN' variedad criolla mejorada de maiz para areas de temporal del sureste de Coahuila, Mexico. Revista Fitotecnia Mexicana. 2014;**37**(4):403-405

[32] Barreto HJ, Bolaños JA, Cordova HS. Programa indices de seleccion. guia para la operacion del software. 1st ed. Mexico: CIMMYT; 1991. p. 27

Chapter 7

Lessons from Common Bean on How Wild Relatives and Landraces Can Make Tropical Crops More Resistant to Climate Change

Andrés J. Cortés and Matthew W. Blair

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.71669

Abstract

Warming is expected to lead to drier environments worldwide, especially in the tropics, and it is unclear how crops will react. Drought tolerance often varies at small spatial scales in natural ecosystems, where many of the wild relatives and landraces of the main crops have been collected. Through a series of examples, we will show that collections of wild relatives and landraces, many of those deposited at germplasm banks, may represent this desired source of variation, as they are genetically diverse and phenotypically variable. For instance, using a spectrum of genotyping and phenotyping approaches, we have studied the extent of genetic and phenotypic diversity for drought tolerance in wild and landraces of common bean (*Phaseolus vulgaris* L.) and compared it with the one available at cultivated varieties. Not surprisingly, most of the naturally available variation to cope with drought in the natural environments was lost through domestication and recent plant breeding. It is therefore imperative to exploit the reservoir of wild relatives and landraces to make crops more tolerant. Yet, it remains to be seen if the rate at which this naturally available variation can be incorporated into the cultivated varieties may keep pace with the rate of climate change.

Keywords: drought tolerance, environmental adaptation, genomic signatures of selection, agro-ecological models, divergent selection

1. Introduction: Common bean – a model to explore the usefulness of wild relatives and landraces as a resource for the future

In this chapter, we review the utility of genome-environment association approaches to infer the potential of wild accessions and landraces to make tropical crops more resistant to climate change, using the food crop common bean (*Phaseolus vulgaris* L.) as a model.

Figure 1. Geographic distribution of wild relatives (light gray) and landraces (dark gray) of common bean and its diversity in terms of seed size, colors, and patterns. Modified from Cortés *et al.* [12].

http://dx.doi.org/10.5772/intechopen.71669

Wild beans are thought to have diversified and adapted locally in South and Central America from an original range in Central America [1, 2], after which domestication in the southern and northern ends of each region gave origin to Andean and Mesoamerican domesticates, respectively [3–7]. Both gene pools followed somewhat parallel pathways of dissemination through the world, generating new secondary centers of diversity in Africa and Asia [8].

Common bean is a source of nutrients and protein for over 500 million people in Latin America and Africa, and more than 4.5 out of 23 million hectares are grown in zones where drought is severe, such as in northeastern Brazil, coastal Peru, the central and northern highlands of Mexico, and in Eastern and Southern Africa [9, 10]. This situation may worsen as increased drought due to climate change will reduce global crop production in greater than 10% by 2050 [11]. Increasing drought tolerance in common bean varieties is therefore needed. Characterizing geo-referenced landraces and wild accessions of common bean at the genetic level (e.g., **Figure 1**) and quantifying single-nucleotide polymorphism (SNP) allelic associations with a bioclimatic-based drought index offer an efficient path to identify adaptive variations suitable to breed new drought-tolerant varieties.

In the following two sections, we fist explain the theoretical bases behind genome-environment associations, as well as its caveats (Section 2), and later we exemplify it with concrete cases that used geo-referenced landraces and wild accessions of common bean to infer naturally available adaptive variations (Section 3).

2. Strategies to infer adaptability of wild relatives and landraces to their natural habitats

Understanding the genomic signatures associated with environmental variation provides insights into how species adapt to their environment [13–15]. Recent genomic studies in wild populations have demonstrated that genome-environment associations, which are associations between SNP alleles and accessions' environment of origin, can indeed be used to identify adaptive loci and predict phenotypic variation. For instance, Turner *et al*. [16] predicted genetic adaptive variation to serpentine soils in *Arabidopsis lyrata;* Hancock *et al*. [17] identified climate-adaptive genetic loci among a set of geographically diverse *Arabidopsis thaliana;* Fischer *et al*. [18] predicted adaptive variation to topoclimatic factors in *Arabidopsis halleri;* Pluess *et al*. [19] predicted genetic local adaptation to climate at a regional scale in *Fagus sylvatica;* and Yeaman *et al*. [20] detected convergent local adaptation in two distantly related species of conifers.

This genome-environment association approach has also been explored in some crop accessions as a prospection strategy of germplasm, alternative to traditional phenotyping. For example, Yoder *et al*. [21] was able to capture adaptive variation to thermal tolerance, drought tolerance, and resistance to pathogens in *Medicago truncatula;* Lasky *et al*. [22] predicted genotype-by-environment interactions to drought stress and aluminum toxicity in *Sorghum bicolor;* and Berthouly-Salazar *et al*. [23] uncovered genomic regions involved in adaption to abiotic and biotic stress on two climate gradients in *Cenchrus americanus*.

Nonetheless, since genomic signatures associated with habitat heterogeneity can result from causes other than adaptation and selection [24, 25], for example, random genetic drift (**Figure 2**), and are also influenced by differences in ancestral variation and recombination in the genome [27–29], some further approaches need to be undertaken to clarify the truthful nature of the divergent regions. For instance, the origin of habitat-associated variants from novel or standing genetic variation leads to distinctively different patterns of genomic divergence [30–32]. One approach that can help to distinguish these underlying causes of divergence is comparing summary statistics (i.e., Tajima's D) from different genomic sections because demographic processes usually leave genome-wide signatures while selection tends to imprint more localized regions [33]. Specifically, habitat-mediated purifying selection is associated with localized low values of nucleotide diversity (π) [34] and Tajima's D [35] and high scores of the Watterson's theta (θ) estimator [36] because only low-frequency polymorphisms can avoid being eliminated by widespread directional selection. Although recent population bottlenecks tend to achieve the same reduction in nucleotide variation, this pattern is expected at a more genome-wide level. Similarly, local adaptation tends to homogenize haplotypes within the same niche, fix polymorphisms in different populations, and eliminate low-frequency polymorphism. Consequently, few haplotypes with high frequency are retained,

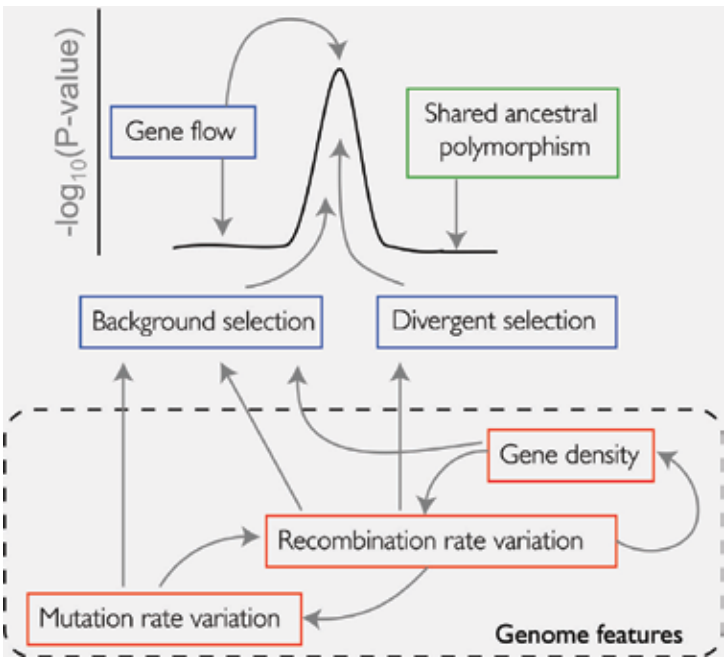

Figure 2. Multiple causes explain genome-environment associations. External processes, such as divergent selection, which is the main focus when assessing adaptation in wild relatives and landraces of crops, is only one of many possible causes. At the same time, the genomic background may be homogenized by gene flow. Similarly, background selection and genomic features in regions of reduced recombination rate and shared ancestral polymorphism (more prone to genetic drift due to their reduced effective population size) could induce hotspots of spurious genome-environment associations. Therefore, besides external processes driven by natural selection, both inherent properties of the genome and the demographic and evolutionary history of the crop influence the extent of the genome-environment associations. Modified from Ravinet, Faria [26].

corresponding to high values of nucleotide diversity (π) and Tajima's D and low scores of the Watterson's theta (θ) estimator [33]. While independent domestication events, extensive population structure, and population expansions after bottlenecks can produce the same patterns, these demographic processes also imprint genomes at a more genome-wide level.

In the following two sub-sections, we explain how to implement genome-environment associations in order to infer the adaptability of wild relatives and landraces to their natural habitats (Section 2.1) and discuss ways to account for causes, other than adaptation and selection, that may be shaping the genomic landscape of signatures associated with habitat heterogeneity (Section 2.2).

2.1. Using genome-environment association scans to identify loci associated with bioclimatic-based indexes

First of all, in order to account for possible demographic effects, subpopulation structure must be determined in geo-referenced landraces and wild accessions using principal coordinates analysis (PCoA) implemented in the software Trait Analysis by aSSociation, Evolution and Linkage, TASSEL v.5 [37]. The same dataset and software can be used to perform association analyses between the SNP markers and bioclimatic-based indexes (e.g. [12, 38, 39]).

As a rule of thumb, a total of ten generalized (GLM) and mixed (MLM) linear models should be compared [40]. Within each model family, five models are usually built as follows: (1) models with the gene pool identity and the first two PCoA axes scores as covariates; (2) models with the within-gene pool subpopulation identity (e.g. [41]) and the first two PCoA axes scores as covariates; (3) models with the first two PCoA axes scores as covariates; (4) models with the within-gene pool subpopulation identity (e.g. [41]) as covariates; and (5) models with the gene pool identity as covariates. All five MLMs usually use a centered IBS kinship matrix as a random effect to control for genomic background implementing the EMMA and P3D algorithms to reduce computing time [42]. QQ-plots of the P-values should be inspected to assess whether excessive numbers of false positives are generated and choose in this way the optimum model. Significant associations are determined using strict Bonferroni corrections of P-values at alpha = 0.001, leading, for example, to a significance threshold of 4.4×10^{-8} in a usual dataset of ca. 23,000 SNP markers (0.001 divided by the number of markers) or $-\log_{10}(4.4 \times 10^{-8}) = 7.36$. The construction of customized PCoA and Manhattan diagrams can be carried out with the software R v.3.3.1 (R Core Team).

Finally, candidate genes for habitat adaptation can be identified within the 1000 bp sections, flanking each SNP marker that is associated with a bioclimatic-based index by using the corresponding reference genome (e.g. [5]) and the PhytoMine and BioMart tools in Phytozome v.12 (phytozome.jgi.doe.gov).

2.2. Accounting for genomic constrains by inspecting genome-wide patterns of variation

In order to identify causes other than adaptation and selection that may be shaping the genomic landscape of signatures associated with habitat heterogeneity (i.e., genomic constrains and genetic drift), sliding window approaches (e.g., window size = 1×10^6 bps, step size = 200 kb) can be implemented to describe patterns of variation and overall divergence across the genome. For instance, SNP density, nucleotide diversity as measured by π [34],

Watterson's theta (θ) estimator [36], and Tajima's D [35] can be computed using the software TASSEL v.5 [37] and customized R scripts. Results of all windowed analyses are usually plotted against window midpoints in millions of base pairs (Mb) in the software R v.3.3.1 (R Core Team). The centromeres can be marked to visualize the extent of the centromeric repeats and its correlation with overall patterns of diversity and divergence.

It is advisable to calculate bootstrap-based means and 95% confidence intervals around the mean for some summary statistics (i.e., SNP density, π, θ, and Tajima's D) when computed in sliding windows that contained or did not contain at least one marker that was associated with a bioclimatic-based index. For this, each summary statistic of windows containing and not containing associated SNPs should be randomly resampled with replacement (bootstrapping) across windows within grouping factors (associated vs. not associated). The overall mean is then stored for each grouping factor. This step should be iterated at least 1000 times using customized R scripts. Bootstrapping must be performed independently for each summary statistic in order to eliminate correlations among these.

3. The adaptive potential of wild relatives and landraces in common bean

In common bean, ecological gradients related with drought stress are associated with divergent selection at the genetic level, after accounting for gene pool and subpopulation structure. This divergent selective pressure might be a consequence of local-level rainfall patterns. Specifically, in tropical environments near the equator with bimodal rainfall, a mid-season dry period occurs that can last 2–4 weeks. In contrast, in the sub-tropics, a dry period of 3 or more months can occur. In response to this mid-cycle drought of the sub-tropics, *P. vulgaris* enters a survival mode of slow growth and reduced physiological activity until rainfall resumes and flowering occurs [43]. Beans growing in wetter conditions on the other hand are less frequently subjected to these environmental pressures and have a fitness advantage to mature in a shorter length of time. Given these ecological differences, and consistent with genomic signatures of divergent selection, the reaction typically associated with drought tolerance, although favorable under dry conditions, seems detrimental under more humid conditions. The awareness about this trade-off may aid the breeding of new drought-tolerant varieties specifically adapted to unique microenvironments (e.g. [44]) and local regions rather than varieties eventually obsolete, originally intended for a wider range of environments.

In the next two sub-sections, we summarize the concrete evidence supporting these statements (Section 3.1) and explain how we can discard other fortuitous causes that may also explain the same pattern (Section 3.2), based on the approaches that we introduced in the previous section (Section 2).

3.1. The signatures of adaptation in common bean are widespread throughout the genome

SNP markers are good at recovering the well-described Andean and Mesoamerican gene pool structure and the five within-gene pool subpopulations observed in wild common bean [41]. Because of this, in a previous research by us with more than 22,000 SNP markers, QQ-plots,

from the association analyses between those SNP markers and a bioclimatic-based drought index [12], indicated that GLM analyses likely had excessive rates of false positives whereas MLM models controlling for population structure and using a kinship matrix reduced more effectively the false positive rate.

In that particular case, the MLM model with the first two PCoA axes scores used as covariates was the best at controlling for false positives. This model yielded a total of 115 SNP markers associated with the bioclimatic-based drought index at a Bonferroni-corrected significance threshold of 7.36 $-\log_{10}$(P-value). These markers explained on average 51.3% ± 0.4 of the variation in the bioclimatic-based drought index. The 115 SNPs were clustered in 90 different regions, defined as overlapping 1000 bp sections that flanked associated markers (**Figure 3**). Associated SNPs and regions were widespread in all 11 common bean chromosomes.

Following the previous example, chromosomes Pv3 and Pv8 had the highest number of associated SNPs with 21 and 32 SNPs clustered in 16 and 21 different regions, respectively. Chromosomes Pv1, Pv2, Pv4, Pv5, Pv6, and Pv9 contained an intermediate number of associated SNPs with 11, 6, 11, 7, 12, and 9 SNPs clustered in 11, 6, 8, 6, 8, and 9 different regions, respectively. Chromosomes Pv7, Pv10, and Pv11 had the fewest number of associated SNPs with 3, 2, and 1 SNPs clustered in 3, 1, and 1 different regions, respectively. Chromosome Pv8 had more regions with at least 2 associated SNPs than any other chromosome, and these regions had more associated SNPs than in any other chromosome for a total of 5 regions with an average number of associated SNPs of 3.2. The single region that contained more associated SNPs was also situated in chromosome Pv8 with 6 SNPs explaining on average 51.1% ± 0.3 of the variation in the bioclimatic-based drought index. After chromosome Pv8, Pv3 was also outstanding, having 4 regions (with at least 2 associated SNPs) with an average number of associated SNPs of 2.5. Therefore, a total of 75 regions, comprising 99 SNP markers associated with the bioclimatic-based drought index, contained at least 1 gene for a total of 77 genes. Most genes were in chromosomes Pv1, Pv3, and Pv8 with 11, 14, and 16 genes. Only two regions, at chromosomes Pv1 and Pv8 and containing a total of seven different SNPs, spanned two or more genes. The one in Pv8 was the region with more associated SNPs (six in total). One of the two genes in this region encoded an Ankyrin repeat-containing protein, which was associated with osmotic regulation via the assembly of cation channels in the membranes [45]. Among other identified candidate genes, there was a phototropic-responsive NPH3 gene [46] in Pv3.

3.2. Rampant divergent selection: interpreting genomic signatures of adaptation in common bean beyond genomic constrains

As a follow-up of the previous example, associated genomic windows were enriched for SNP density and positive Tajima's D scores. This conclusion was achieved after implementing a sliding window analysis to explore the patterns of genome-wide diversity (**Figure 3**). Marker density decayed drastically toward the centromeres. This decay in diversity proportional to the decay in the rate of recombination was first described in *D. melanogaster* and has been confirmed in many organisms since then. The correlation was initially understood as an effect of genetic hitchhiking, but background selection has been increasingly appreciated as a contributing factor [28], perhaps in many cases the dominating one.

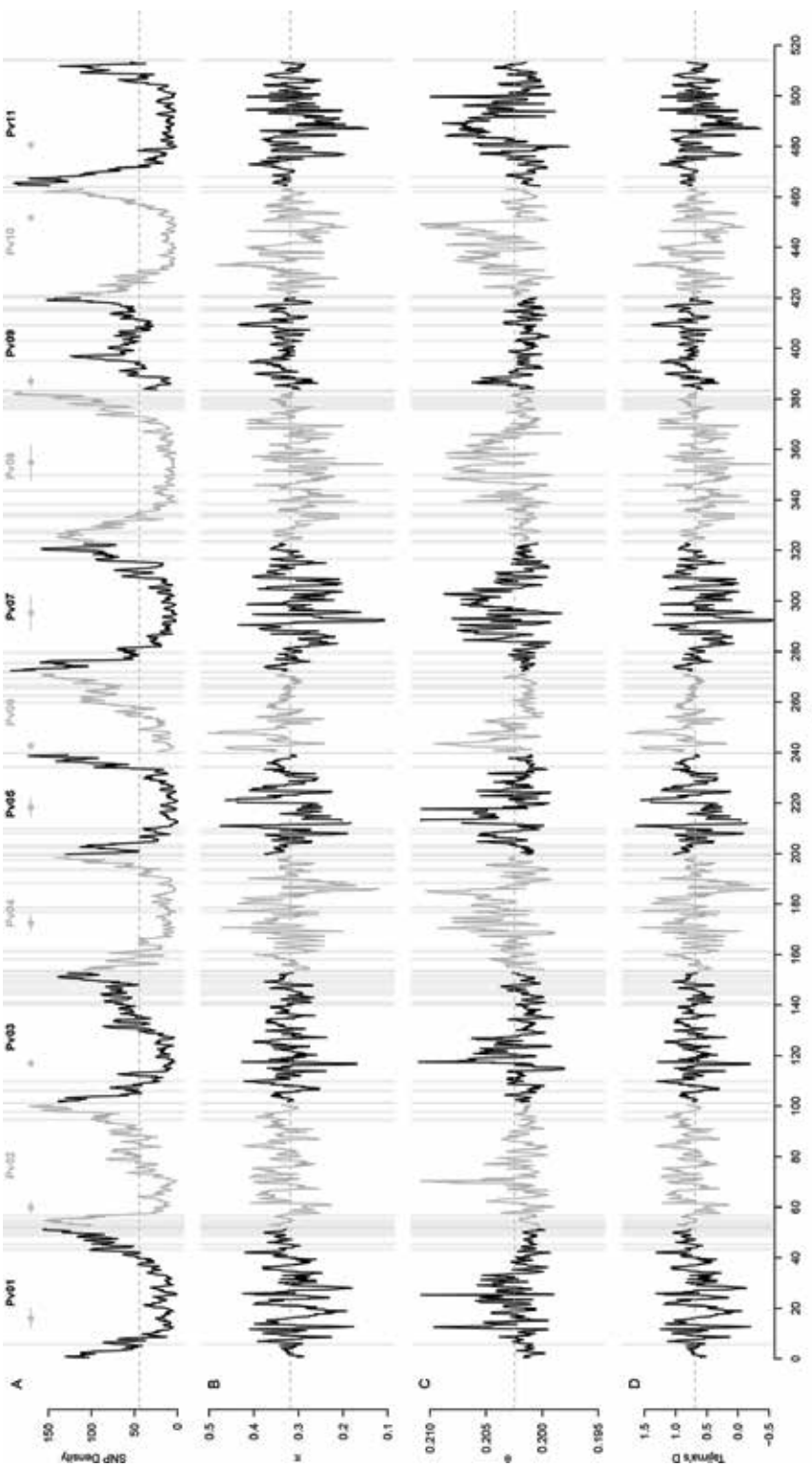

Figure 3. Patterns of genome-wide diversity in common bean accessions based on SNP markers. A sliding window analysis was used to compute (A) SNP density, (B) nucleotide diversity as measured by π, (C) Watterson's theta estimator (θ), and (D) Tajima's D. Results of all windowed analyses are plotted against window midpoints in millions of base pairs (Mb). Black and gray colors highlight different common bean (Pv) chromosomes. Gray dashed horizontal lines indicate genome-wide averages. Gray vertical boxes indicate the 1000 bp flanking region of each marker that was associated with the bioclimatic-based drought index. Horizontal gray lines with a central filled gray dot at the top of the figure mark the centromeres. Modified from Cortés & Blair [66].

The average marker density was 44 SNPs per million base pairs (95% CI, 4–143). The average nucleotide diversity as measured by π was 0.3 per million base pairs (95% CI, 0.2–0.4). The average Watterson's theta (θ) was 0.20 per million base pairs (95% CI, 0.19–0.21). The average Tajima's D was 0.68 per million base pairs (95% CI, 0.05–1.22). These very same statistics were compared between 1 Mb sliding windows that contained (associated) or did not contain (no associated) at least one marker that was associated with the bioclimatic-based drought index. Genomic windows containing at least one associated SNP had an overall higher SNP density (79 ± 6 vs. 39 ± 2), lower values for Watterson's theta (θ) scores (0.2016 ± 0.0001 vs. 0.2026 ± 0001), and more positive Tajima's D scores (0.71 ± 0.02 vs. 0.678 ± 0.009) than windows without associated markers. Nucleotide diversity, as measured by π, was slightly elevated in associated windows when compared with no associated windows (0.322 ± 0.006 vs. 0.317 ± 0.003).

Selective process, such as purifying selection and local adaptation (divergent selection), differentially imprint regions within the same genome, causing a heterogeneous departure of genetic variation from the neutral expectations and from the background trend [28]. Divergent selection tends to homogenize haplotypes within the same niche, fix polymorphisms in different populations, and eliminate low-frequency polymorphism. Consequently, few haplotypes with high frequency are retained, corresponding to high values of nucleotide diversity and Tajima's D and low scores of the Watterson's theta (θ) estimator [33]. We have identified these signatures in the various genomic regions associated with a bioclimatic-based drought index. Therefore, it is unlikely that independent domestication events, extensive population structure, and population expansions after bottlenecks are responsible for these patterns because the mixed linear model that we used to identify the genome-environment associations accounted for population structure, while demographic processes would leave genome-wide signatures in both, associated and no-associated windows.

4. Conclusions

Wild accessions and landraces of common bean occupy more geographical regions with extreme ecologies [2] and extensive drought stress [12] than cultivated accessions. Those regions include the arid areas of Peru, Bolivia and Argentina, and the valleys of northwest Mexico. Hence, a broad habitat distribution for wild common bean has exposed these genotypes to both dry and wetter conditions, while cultivated common bean has a narrower distribution and is traditionally considered susceptible to drought. These differences in the ecologies of wild and cultivated common bean have been associated with higher genetic diversity in the former group when surveying candidate genes for drought tolerance such as the ASR [47], DREB [48], and ERECTA [49] gene families, once the population structure [41] and the background distribution of genetic diversity have been accounted for.

Also, as identified through the genome-environment association approach that was illustrated in this chapter, there are notorious differences between the adaptations of wild accessions and landraces found in arid and more humid environments, in congruence with natural divergent selection acting for thousands of years. Several of these differences might be valuable for plant breeding. Therefore, we reinforce, as was envisioned by Acosta [50], that wild

accessions and landraces of common bean be taken into account to exploit naturally available divergent variations for drought tolerance. We envision that this lesson from common bean will inspire the exploitation of wild relatives and landraces of other crops to face the threats imposed by current climate change.

5. Prospects

This chapter ultimately illustrates that genomic signatures of environmental adaptation (e.g. [51]) are useful for germplasm characterization, potentially enhancing future marker-assisted selection and crop improvement. We envision that genome-environment association studies coupled with estimates of genome-wide diversity will become more common in the upcoming years. These types of studies will likely go beyond estimates of drought tolerance, as exemplified here, to also include estimates regarding frost stress (i.e. [52–54]), nutrient limitation [55, 56], as well as other threats imposed by climate change [57, 58] in different types of ecosystems (e.g. [59]) and screened by a variety wide range of genotyping techniques [60–63]. Genomic selection models [64] could also incorporate at some point environmental variables in order to improve the prediction of phenotypic variation and the estimation of the genotype-by-environment interactions as well as phenotypic characterizations through novel high-throughput methods such as remote sensing and image analysis [65], and novel models such as genome-environment associations [66] in the light of linkage disequilibrium (LD) [67] and various stochastic approaches [68, 69].

Acknowledgements

We are grateful with Daniel G. Debouck and the Genetic Resources Unit at the International Center for Tropical Agriculture for donating the seeds and sharing the geographic coordinates of the plant material that was referred to in this study. This research was supported by the Lundell and Tullberg grants to AJC. The Geneco mobility fund from Lund University is acknowledged for making possible the synergistic meeting between AJC and MWB in the Spring of 2015 at Nashville, Tennessee. Writing time was possible thanks to the grants CTS14:55 from the Carl Trygger Foundation to S. Berlin and 4.1-2016-00418 from Vetenskapsrådet (VR) to AJC.

Author details

Andrés J. Cortés[1,2]* and Matthew W. Blair[3]

*Address all correspondence to: acortes@corpoica.org.co

1 Department of Biological and Environmental Sciences, University of Gothenburg, Gothenburg, Sweden

2 Colombian Corporation for Agricultural Research (Corpoica), Rionegro, Colombia

3 Department of Agricultural & Environmental Science, Tennessee State University, Nashville, TN, USA

References

[1] Bitocchi E, et al. Mesoamerican origin of the common bean (*Phaseolus vulgaris* L.) is revealed by sequence data. Proceedings of the National Academy of Sciences. 2012;**109**(14):E788-E796

[2] Rodriguez M, et al. Landscape genetics, adaptive diversity and population structure in *Phaseolus vulgaris*. The New Phytologist. 2016;**209**(4):1781-1794

[3] Gepts P, Debouck D. Origin, domestication and evolution of the common bean (*Phaseolus vulgaris* L.). In: Van Shoonhoven A, Voysest O, editors. Common Beans: Research for Crop Improvement. Commonwealth Agricultural Bureau: Wallingford; 1991. p. 7-53

[4] Kwak M, Gepts P. Structure of genetic diversity in the two major gene pools of common bean (*Phaseolus vulgaris* L., *Fabaceae*). Theoretical and Applied Genetics. 2009;**118**(5):979-992

[5] Schmutz J, et al. A reference genome for common bean and genome-wide analysis of dual domestications. Nature Genetics. 2014;**46**(7):707-713

[6] Bitocchi E, et al. Molecular analysis of the parallel domestication of the common bean (*Phaseolus vulgaris*) in Mesoamerica and the Andes. New Phytologist. 2013;**197**(1):300-313

[7] Cortés AJ. On the origin of the common bean (*Phaseolus vulgaris* L.). American Journal of Plant Sciences. 2013;**4**(10):1998-2000

[8] Asfaw A, Blair MW, Almekinders C. Genetic diversity and population structure of common bean (*Phaseolus vulgaris* L.) landraces from the East African highlands. Theoretical and Applied Genetics. 2009;**120**(1-12)

[9] Broughton WJ, et al. Beans (*Phaseolus* spp.) – Model food legumes. Plant and Soil. 2003; **252**(1):55-128

[10] Darkwa K, et al. Evaluation of common bean (*Phaseolus vulgaris* L.) genotypes for drought stress adaptation in Ethiopia. The Crop Journal. 2016;**4**(5):367-376

[11] Tai APK, Martin MV, Heald CL. Threat to future global food security from climate change and ozone air pollution. Nature Climate Change. 2014;**4**:817-821

[12] Cortés AJ, et al. Drought tolerance in wild plant populations: The case of common beans (*Phaseolus vulgaris* L.). PLoS One. 2013;**8**(5):e62898

[13] Hoffmann AA, Sgro CM. Climate change and evolutionary adaptation. Nature. 2011; **470**(7335):479-485

[14] Stinchcombe JR, Hoekstra HE. Combining population genomics and quantitative genetics: Finding the genes underlying ecologically important traits. Heredity. 2008;**100**:158-170

[15] Savolainen O, Lascoux M, Merilä J. Ecological genomics of local adaptation. Nature Reviews Genetics. 2013;**14**(11):807-820

[16] Turner TL, et al. Population resequencing reveals local adaptation of *Arabidopsis lyrata* to serpentine soils. Nature Genetics. 2010;**42**(3):260-263

[17] Hancock AM, et al. Adaptation to climate across the *Arabidopsis thaliana* genome. Science. 2011;**334**(6052):83-86

[18] Fischer MC, et al. Population genomic footprints of selection and associations with climate in natural populations of *Arabidopsis halleri* from the Alps. Molecular Ecology. 2013;**22**(22):5594-5607

[19] Pluess AR, et al. Genome-environment association study suggests local adaptation to climate at the regional scale in *Fagus sylvatica*. The New Phytologist. 2016;**210**(2):589-601

[20] Yeaman S, et al. Convergent local adaptation to climate in distantly related conifers. Science. 2016;**353**(6306):1431-1433

[21] Yoder JB, et al. Genomic signature of adaptation to climate in *Medicago truncatula*. Genetics. 2014;**196**(4):1263-1275

[22] Lasky JR, et al. Genome-environment associations in sorghum landraces predict adaptive traits. Science Advances. 2015;**1**(6):e1400218

[23] Berthouly-Salazar C, et al. Genome scan reveals selection acting on genes linked to stress response in wild pearl millet. Molecular Ecology. 2016;**25**(21):5500-5512

[24] Nei M. The neutral theory of molecular evolution in the genomic era. Annual Review of Genomics and Human Genetics. 2010;**11**:265-289

[25] Cortés AJ, et al. Genome-wide patterns of microhabitat-driven divergence in the alpine dwarf shrub *Salix herbacea* L. In: On The Big Challenges of a Small Shrub: Ecological Genetics of *Salix herbacea* L. Uppsala: Acta Universitatis Upsaliensis; 2015

[26] Ravinet M, et al. Interpreting the genomic landscape of speciation: A road map for finding barriers to gene flow. Journal of Evolutionary Biology. 2017;**30**(8):1450-1477

[27] Strasburg JL, et al. What can patterns of differentiation across plant genomes tell us about adaptation and speciation? Philosophical Transactions of the Royal Society B: Biological Sciences. 2011;**367**(1587):364-373

[28] Ellegren H, Galtier N. Determinants of genetic diversity. Nature Reviews. Genetics. 2016;**17**(7):422-433

[29] Wolf JB, Ellegren H. Making sense of genomic islands of differentiation in light of speciation. Nature Reviews Genetics. 2017;**18**:87-100

[30] Barrett RD, Schluter D. Adaptation from standing genetic variation. Trends in Ecology & Evolution. 2008;**23**(1):38-44

[31] Pritchard JK, Pickrell JK, Coop G. The genetics of human adaptation: Hard sweeps, soft sweeps, and polygenic adaptation. Current Biology. 2010;**20**(4):R208-R215

[32] Hermisson J, Pennings PS. Soft sweeps: Molecular population genetics of adaptation from standing genetic variation. Genetics. 2005;**169**(4):2335-2352

[33] Wakeley J. Coalescent Theory: An Introduction. Cambridge: Harvard University; 2008

[34] Nei M. Molecular Evolutionary Genetics. New York: Columbia University Press; 1987

[35] Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 1989;**123**(3):585-595

[36] Watterson GA. Number of segregating sites in genetic models without recombination. Theoretical Population Biology. 1975;**7**:256-276

[37] Bradbury PJ, et al. TASSEL: Software for association mapping of complex traits in diverse samples. Bioinformatics. 2007;**23**(19):2633-2635

[38] Cortés AJ. Environmental Heterogeneity at a Fine Scale: Ecological and Genetic Implications in a Changing World. Vol. 60. Saarbrücken, Germany: LAP Lambert Academic Publishing; 2016

[39] Cortés AJ. Local scale genetic diversity and its role in coping with changing climate. In: Bitz L, editor. Genetic Diversity. Rijeka: InTech; 2017. 140 p

[40] Cortés AJ, et al. Maintenance of female-bias in a polygenic sex determination system is consistent with genomic conflict. In: On The Big Challenges of a Small Shrub: Ecological Genetics of *Salix herbacea* L. Uppsala: Acta Universitatis Upsaliensis; 2015

[41] Blair MW, Soler A, Cortés AJ. Diversification and population structure in common beans (*Phaseolus vulgaris* L.). PLoS One. 2012;**7**(11):e49488

[42] Zhang Z, et al. Mixed linear model approach adapted for genome-wide association studies. Nature Genetics. 2010;**42**(4):355-360

[43] Beebe S, et al. Selection for drought resistance in common bean also improves yield in phosphorus limited and Favorable environments. Crop Science. 2008;**48**:582-592

[44] Cortés AJ, et al. Small-scale patterns in snowmelt timing affect gene flow and the distribution of genetic diversity in the alpine dwarf shrub *Salix herbacea*. Heredity. 2014;**113**:233-239

[45] Voronin DA, Kiseleva EV. Functional role of proteins containing ankyrin repeats. Cell and Tissue Biology. 2008;**2**(1):1-12

[46] Pedmale UV, Liscum E. Regulation of phototropic signaling in Arabidopsis via phosphorylation state changes in the phototropin 1-interacting protein NPH3. Journal of Biological Chemestry. 2007;**282**:19992-20001

[47] Cortés AJ, et al. Molecular ecology and selection in the drought-related Asr gene polymorphisms in wild and cultivated common bean (*Phaseolus vulgaris* L.). BMC Genetics. 2012;**13**:58

[48] Cortés AJ, et al. Nucleotide diversity patterns at the drought-related DREB2 encoding genes in wild and cultivated common bean (*Phaseolus vulgaris* L.). Theoretical and Applied Genetics. 2012;**125**(5):1069-1085

[49] Blair MW, Cortés AJ, This D. Identification of an ERECTA gene and its drought adaptation associations with wild and cultivated common bean. Plant Science. 2016;**242**:250-259

[50] Acosta JA, Kelly JD, Gepts P. Prebreeding in common bean and use of genetic diversity from wild germplasm. Crop Science. 2007;**47**(S3):S44-S59

[51] Cortés AJ. On the Big Challenges of a Small Shrub: Ecological Genetics of *Salix herbacea* L. Uppsala: Acta Universitatis Upsaliensis; 2015

[52] Wheeler JA, et al. The snow and the willows: accelerated spring snowmelt reduces performance in the low-lying alpine shrub *Salix herbacea*. Journal of Ecology. 2016;**104**(4):1041-1050

[53] Wheeler JA, et al. Increased spring freezing vulnerability for alpine shrubs under early snowmelt. Oecologia. 2014;**175**(1):219-229

[54] Wheeler JA, et al. With a little help from my friends: Community facilitation increases performance in the dwarf shrub *Salix herbacea*. Basic and Applied Ecology. 2015;**16**:202-209

[55] Little CJ, et al. Small-scale drivers: The importance of nutrient availability and snowmelt timing on performance of the alpine shrub *Salix herbacea*. Oecologia. 2016;**180**(4):1015-1024

[56] Sedlacek J, et al. What role do plant-soil interactions play in the habitat suitability and potential range expansion of the alpine dwarf shrub *Salix herbacea*? Basic and Applied Ecology. 2014;**15**(4):305-315

[57] Sedlacek J, et al. Evolutionary potential in the alpine: Trait heritabilities and performance variation of the dwarf willow *Salix herbacea* from different elevations and microhabitats. Ecology and Evolution. 2016;**6**(12):3940-3952

[58] Sedlacek J, et al. The response of the alpine dwarf shrub *Salix herbacea* to altered snowmelt timing: Lessons from a multi-site transplant experiment. PLoS One. 2015;**10**(4):e0122395

[59] Madriñán S, Cortés AJ, Richardson JE. Páramo is the world's fastest evolving and coolest biodiversity hotspot. Frontiers in Genetics. 2013;**4**(192)

[60] Blair MW, et al. A high-throughput SNP marker system for parental polymorphism screening, and diversity analysis in common bean (*Phaseolus vulgaris* L.). Theoretical and Applied Genetics. 2013;**126**(2):535-548

[61] Kelleher CT, et al. SNP discovery, gene diversity, and linkage disequilibrium in wild populations of *Populus tremuloides*. Tree Genetics & Genomes. 2012;**8**(4):821-829

[62] Galeano CH, et al. Gene-based single nucleotide polymorphism markers for genetic and association mapping in common bean. BMC Genetics. 2012;**13**(1):48

[63] Cortés AJ, Chavarro MC, Blair MW. SNP marker diversity in common bean (*Phaseolus vulgaris* L.). Theoretical and Applied Genetics. 2011;**123**(5):827-845

[64] Desta ZA, Ortiz R. Genomic selection: Genome-wide prediction in plant improvement. Trends in Plant Science. 2014;**19**(9):592-601

[65] Lo Bianco M, Grillo O, Sarigu M, Cremonini R, Venora G. Characterisation of Italian bean landraces (*Phaseolus vulgaris* L.) using seed image analysis and texture descriptors. Australian Journal of Crop Science. 2015;**9**(11):1022-1034

[66] Cortés AJ, Blair MW. Genotyping by sequencing and genome – environment associations in wild common bean predict widespread divergent adaptation to drought. Frontiers in Plant Science. 2018;**9**(128)

[67] Blair MW, Cortés AJ, Farmer A, Assefa T, Penmetsa RV, Cook D, et al. Uneven recombination rate and linkage disequilibrium across a reference SNP map for common bean. PLOS ONE. 2018; DOI: pone.0189597

[68] Cortés AJ. Prevalence in MSM is enhanced by role versatility. In: Mazari A, editor. Big Data Analytics in HIV/AIDS Research. Hershey: IGI Global; 2018

[69] Cortés AJ. On how role versatility boosts an STI. Journal of Theoretical Biology. 2017; **440**:66-69

Chapter 8

Critical Aspects on the Use of Microsatellite Markers for Assessing Genetic Identity of Crop Plant Varieties and Authenticity of their Food Derivatives

Fabio Palumbo and Gianni Barcaccia

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.70756

Abstract

A total of 90 original articles concerning the varietal characterization and identification by means of SSR analysis of the five most economically relevant crops in Italy (i.e., *Olea europaea* L., *Solanum lycopersicum* L., *Vitis vinifera* L., *Triticum* spp. and *Malus* × *domestica* Borkh.) have been selected and reviewed. Since the genetic traceability of processed products may result more complex, wine and olive oil have been considered too. Specifically, this chapter deals with three main aspects: (i) the criteria adopted for the selection of the most appropriate number, type, and distribution of SSR marker loci to be employed for varietal genotyping, (ii) the use of genetic statistics and parameters for the evaluation of the discriminant ability and applicability of SSR marker loci, and (iii) how to make different experimental works on the same species that are standardized, reliable, and comparable. What emerges from the studies reviewed here is a lack of wider consensus among the authors regarding the strategy to design and to adopt for genotyping plant varieties with SSR markers. This finding highlights the urgent need to establish a common procedure, especially for characterizing and preserving landraces, and for supporting its rediscovery and valorization locally.

Keywords: DNA genotyping, plant varieties, genetic traceability, food labeling

1. The Italian agriculture scenery and the utility of SSR markers to develop a reference method for genotyping plant varieties

The Food and Agriculture Organization (FAO) indices of agricultural production describe the relative level of the aggregate volume of agricultural production for each year in comparison with the base period 2004–2006 [1]. According to the most recent data available in The Food and Agriculture Organization Corporate Statistical Database, the gross value of the total Italian agricultural production was equal to $ 41.9 billion, about € 32.7 billion [2]. It is worth

noting that 20 products contribute to over 50% of gross production value (GPV), as shown in **Table 1**.

On average, each species is characterized by dozens or hundreds of cultivars and, as defined in Article 2 of the International Code of Nomenclature for Cultivated Plants, a "cultivar is an assemblage of plants that has been selected for a particular character or combination of characters, that is distinct, uniform, and stable in those characters, and that when propagated by appropriate means, retains those characters" [3]. If some cultivars are virtually ubiquitous, some others are associated with specific geographical contexts and often provide the basis for the establishment of protected designation of origin (PDO) and protected geographical indication (PGI) products (**Table 1**).

It is not a coincidence that Italy, with its 268 brand products, including 106 PGI, 160 PDO, and 2 traditional specialty guaranteed (TSG) labels, is the European leader in terms of certified productions and that 20% of them arise from the 20 crops listed in **Table 1**. As a whole, the Italian certified products reach around 500 units, including two important derivatives such as olive oil

Crop plants	**Value [2] of agriculture production USD (10^6)**	**Registered cultivars**	**PDO and PGI [4]**
Olives (table and oil)	5064.24	644 [5]	3
Tomatoes	4753.00	445 [6]	3
Grapes (table and wine)	2770.60	638 [7]	3
Wheat (durum, common, spelt)	2558.23	489 [8]	0
Maize	2363.83	1739 [8]	0
Apples	1129.69	75 [9]	5
Oranges	990.80	n.a.	3
Potatoes	751.58	56 [8]	3
Rice, paddy	750.28	194 [8]	3
Peaches and nectarines	570.69	311 [9]	4
Pumpkins	532.35	8 [6]	0
Pears	521.17	32 [9]	2
Mandarins, Clementines	431.27	n.a.	2
Artichokes	386.11	14 [6]	4
Carrots and Turnips	355.33	8 [6]	2
Cauliflowers and Broccoli	314.07	41 [6]	0
Beans	311.63	39 [6]	6
Lemons	267.16	n.a.	6
Onions	264.10	71 [6]	2
Hazelnuts (with shell)	247.36	25 [9]	3
Total	**25,333.49**	**4829**	**54**

Table 1. GPV, registered cultivars, PDO, and PGI products for the 20 most economically important crops in Italy.

and wine (**Table 2**). It is worth noting that the wine GPV (**Table 2**) is four times higher than the grape GPV and slightly less than half of the total GPV shown in **Table 1**, a demonstration that shows producing food derivatives could be more profitable than selling raw products.

One of the main problems that needs to be addressed is the lack of a uniform, complete, and updated register of cultivars. For the cultivars of some, species like cereals or vegetables are already available as official registers provided by the Ministry of Agricultural, Food and Forestry Policies (MIPAAF, National register of agricultural varieties and National register of horticultural varieties). Concerning fruit trees, on the contrary, there is not a register yet, although Article 7 of the Italian Legislative Decree no. 124/2010 has established a "National Register of fruit trees varieties" [11]. For this reason, the inventory of cultivars of some fruit species is still ongoing and there is a total lack of official data for some of them (see for instance orange, lemon, and mandarin, **Table 1**). Moreover, for species of particular interest, there exist registers apart (see for example, *Olea europaea* L. and *Vitis vinifera* L.).

In the past, cultivars have been extensively characterized by morphological traits, including plant, leaf, fruit, and seed characteristics. Since objectivity is crucial to perform an accurate morphological typing, it is constraining to use exclusively morphological descriptors for plant cultivars, especially because most of the morphological traits are influenced by environmental factors. Several cases of misidentification, owing to classifications carried out only employing morphological traits, are reported in the scientific literature for a wide range of vegetal crops [12–14] and fruit trees [15–18]. Moreover, the uneven distribution, simultaneous cultivation of local varieties, ambiguous names, continuous interchange of plant materials among varieties and/or farmers of different regions and countries, possibility of the cultivation of varietal clones, and uncertainty of varietal certification in nurseries have complicated the identification of genotypes [19–21]. At the same time, cultivar and clone identity is also very important for protecting plant breeders' rights not only for commercial seeds but also for processed materials and food derivatives, especially for the final consumers' safeguard. Another important aspect to highlight is the need to ensure that each specific variety grown by farmers and its food product bought by consumers is the one declared on the label. This is especially true if the product is sold in a processed or transformed form (thus difficult to recognize phenotypically) and/or if the product is subjected to a form of certification (PDO or PGI). In a modern market, it is crucial being able to identify agricultural products and foodstuffs by means of reliable traceability systems, including genetic molecular markers.

The method of DNA genotyping based on microsatellite markers represents an efficient, reliable, and suitable technique that is able to complement the information provided by morphological

	Value [2] of agriculture production USD (10^6)	**PDO and PGI [10]**
Wine	11603.83	403
Oil, olive, virgin	2126.78	41
Total	**13,730.61**	**444**

Table 2. GPV, PDO, and PGI products for wine and olive oil in Italy.

traits and that has been extensively used for the characterization of plant varieties [22–24] and the certification of food products [25–27].

Microsatellites (or simple sequence repeats (SSRs)) are PCR-based molecular markers valued for their abundant and uniform genome coverage, high levels of polymorphism information content as a consequence of their marked mutation rates, and other valuable qualities such codominant inheritance of DNA amplicons/alleles and request of little amount of DNA for the amplifications [28]. A unique pair of primers defines each SSR marker locus; as a consequence, the molecular information exchange among laboratories is easy and allows individuals to be uniquely genotyped in a reproducible way [29].

SSR markers have been shown repeatedly as being one of the most powerful marker methodologies for genetic studies in many crop species. In fact, since they are multiallelic chromosome-specific and well distributed in the genome, microsatellite markers have already been used for mapping genes with Mendelian inheritance [30], for identifying quantitative trait loci (QTLs, [31]) and for molecular marker-assisted selection [32]. In many species, microsatellite markers have also been used for ascertaining the genetic purity of seed lots [33], as well as to assess the capability to protect the intellectual property of plant varieties [34]. These markers are also largely used for assessing the genetic diversity and relationships among populations and lines, and for identifying crop varieties.

The advantages of SSRs over single-nucleotide polymorphisms (SNPs), another co-dominant marker system increasingly exploited in breeding programs, include relative ease of transfer between closely related species [35, 36] and high allelic diversity [37, 38]. On the contrary, SSRs when compared to SNPs have some limits: the development phase is quite long and expensive for multilocus assays and the throughput is relatively low because of drawbacks for automation and output data management. Recently, progresses in the development of multilocus assays have been made in several directions, suggesting that SSR markers still remain as relevant molecular tools at least for specific applications and genetic studies [39]. In fact, PCR-based SSR genotyping has rapidly evolved in plants, and methods for the simultaneous amplification of multiple marker loci coupled to semi-automated detection systems have been developed [40]. The identification and selection of SSR markers have become cheaper and faster due to the emergence of next-generation sequencing technology means. Moreover, the possibility to multiplexing specific combinations of microsatellite markers has become much easier and the availability of capillary electrophoresis equipment relying on automated laser-induced fluorescence DNA technology has facilitated the adoption and exploitation of this methodology in applied breeding programs [41–43].

Genotypic characterization through SSR loci analysis represents a molecular tool applicable to all species and able to support the phenotypic observation in order to characterize and describe a cultivated variety as well as to define its uniformity, distinctiveness, and stability (DUS testing). At the same time, SSR markers are largely used for the genetic identification of varieties and the authentication and traceability of their foodstuffs [44–46].

The main goal of this work is to provide an updated and detailed description of the applications of SSR markers for varietal characterization and identification, reviewing the state of the art of genotyping in the most economically relevant Italian crop plants and food products: *Olea*

europaea L., *Solanum lycopersicum* L., *Vitis vinifera* L., *Triticum* spp., and *Malus* × *domestica* Borkh., wine and olive oil. In this respect, the chapter aims to assess the real achievements of different genotyping analyses, to evaluate the strengths and limitations according to applied research studies, and to emphasize the striking lack of data related to the applications of SSR technology. Through the careful investigation and evaluation of a large number of scientific papers, our review highlights some critical aspects on the use of microsatellite markers and formulates recommendations for standardizing the strategies and methods for ascertaining the genetic identity of plant varieties and for achieving the genetic traceability of their food derivatives. Here, we focus on three main aspects: (i) how to choose and use SSR markers, (ii) which parameters/indices calculate for the genetic characterization of plant materials, and (iii) assess a standardized way to make SSR data from different works on the same species comparable.

2. Applications of SSR markers for the genetic characterization of crop plant varieties

Some of the most economically important crops in Italy have been chosen for this study, and the search has been focused on their varietal characterization through SSR analysis. In particular, olive (*Olea europaea* L.), grape (*Vitis vinifera* L.), and apple (*Malus* × *domestica* Borkh.) were reviewed among the fruit trees, whereas wheat (*Triticum* spp.) and tomato (*Solanum lycopersicum* L.) were selected as representative of cereals and vegetables, respectively. A large number of commercial cultivars are available for each of these species, and the annual Italian GPV for these crops is about 18 billion Euro [2]. Moreover, scientific articles dealing with the genetic identification in wines and olive oils were also evaluated because these two derivatives contribute to the annual Italian GPV for another 15 billion Euro [2].

Although passport data, morphological, and agronomical descriptors have been collected, data are not informative enough to assess the numerous cases of misidentification, mislabeling, homonymies, and synonymies as well as voluntary or accidental frauds [47]. With regard to this, several research groups characterized and identified cultivars using SSR markers (**Table 3**).

Crops	**References**
Olive (*Olea europaea* L.)	[23, 25, 48–63]
Tomato (*Solanum lycopersicum* L.)	[27, 44, 64–70]
Grape (*Vitis vinifera* L.)	[15, 19, 24, 71–89]
Wheat (*Triticum* spp.)	[22, 26, 90–98]
Apple (*Malus* × *domestica* Borkh.)	[99–111]
Derivatives	
Wine	[45, 112–116]
Olive oil	[46, 58, 117–125]

Table 3. Crops and derivatives reviewed.

Article searches were performed using the three most popular sources of scientific information: Scopus, Web of Science, and Google Scholar, while PubMed was excluded from the queried datasets because it focuses mainly on medicine and biomedical sciences and also because Google scholar already includes its index [126]. A total of 90 articles based on SSR genotyping analysis were selected from the international literature in the last 15 years, covering all the plant species/food products taken as reference list. Only articles dating from 2000 to now were reviewed assuming that researches published earlier would have lost their steering effects on the activities of plant DNA genotyping, given that the development of new and large marker datasets, and technologically advanced and automated protocols has been very fast in the last 15 years.

3. What number and how to select a panel of SSR marker loci according to their linkage map position and polymorphism information content

More than 800 SSR markers have been developed in apple (*Malus* × *domestica* Borkh., 2n = 2 × = 34), and nearly all of them have been mapped on a consensus map produced starting from five different genetic maps [127]. These markers are distributed across all 17 linkage groups, with an average of 49 microsatellites per linkage group. Moreover, the genome database for Rosaceae [128] is a long-standing community database resource providing hundreds of microsatellite loci, in most cases accompanied by a wealth of information about map position, repeat motifs, primers, PCR conditions, amplicon length, and publication source. A discriminatory set of markers should ensure the uniform distribution across the genome of the microsatellite loci to represent adequately each linkage group and, thus, the genome in its entirety [91]. In fact, assessing the genetic diversity by focusing only on restricted regions of the genome may threaten to distort results. Nevertheless, neglecting the most ambitious study on *Malus* × *domestica* Borkh. carried out by Patocchi et al. [105] using an extremely high number of SSR markers (82), the number of selected and analyzed genomic loci varies from 4 to 19 with an average value of 12 ± 6 SSR markers, less than a microsatellite locus per linkage group. Extending this reasoning to the other crops reviewed, the emerging output is often the same: for all the plant species, very detailed genetic maps are available [129–132] as well as dedicated databases for SSR markers (**Table 4**).

Olea europaea L. (2n = 2 × = 46) includes 23 chromosome pairs and the average number of microsatellite markers used in the reviewed articles is 11 ± 5, much less than a microsatellite locus per linkage group. The same is also true for *Vitis vinifera* L. (2n = 2 × = 38) in which the average number of microsatellite markers explored for genotyping cultivars is 15 ± 11 in spite of the 19 chromosome pairs of this species. Even the varietal identification of their respective derivatives (olive oil and wine) has been accomplished by exploring, on average, 8 ± 3 and 10 ± 4 SSR markers, respectively. On the contrary, in wheat, the varieties of both *Triticum durum* Desf. (2n = 4 × = 28) and *Triticum aestivum* L. (2n = 6 × = 42) have been characterized by means of genotyping with SSR markers analyzing, on average, 18 ± 3 and 21 ± 6 microsatellite loci respectively, that is more than one microsatellite per linkage group. This latter choice is perhaps associated with the high complexity and large size of the *Triticum aestivum* L. genome, approximately equal to 17 Gb/1C [148]. In fact, for a correct representation of the

entire genome, not only the number of homologous chromosomes but also their size (i.e., total amount of DNA) should be considered when choosing the optimal panel of microsatellite loci to be investigated. Finally, in tomato (*Solanum lycopersicum* L., 2n = 2 × = 24), the average number of SSR markers employed for genotyping varieties is 14 ± 7 (**Table 4**).

Only few studies [65, 74, 96, 106] evaluated the position within linkage groups of the microsatellites selected: the choice often falls on SSR markers with unknown or not specified position or mapped on few chromosomes, thus resulting in a poor representation of the entire genome. In this regard, the results from Cipriani et al. [74] and van Treuren et al. [106] represent a good model for the choice of molecular markers to investigate the genetic diversity in germplasm collections and to solve synonymy/homonymy cases as well as paternity and kinship issues. The former group selected microsatellite sequences from scaffolds anchored to the 19 linkage groups of *Vitis vinifera* L. with the aim of analyzing 38 well-distributed SSR

Species	Genome size (Gb)	Ploidy	SSR available (SSR database)	SSR employed (mean ± st. dev)	No. of reference cultivars	No. of reference SSRs
Olea europaea L.	1.42–2.28 [133]	2n = 2× = 46	12 (OLEA Database) [134]	11 ± 5	21 [53], 17 [52]	11 [53], 8 [52]
Solanum Lycopersicum L.	0.90–0.95 [132]	2n = 2 × = 24	146,602 (Tomato microsatellite database) [135]	14 ± 7	n.a.	n.a.
			66,823 (Tomato genomic resources database) [136]			
			21,100 (Tomato: Kazusa Marker Database) [137]			
Vitis vinifera L.	0.48 [129]	2n = 2× = 38	56 (Grape microsatellite collection) [138]	15 ± 11	49 [139]	6 [139], 38 [74]
			443 (Italian Vitis Database) [140]			
			6 (The European Vitis Database) [141]			
Triticum spp.	12.3–13.00 (*T. durum* Desf) [142]	2n = 4× = 28	588 (Wheat microsatellite consortium) [143]	18 ± 3	n.a.	46 [144]
	16.50–17.00 (*T. aestivum* L.) [142]	2n = 6× = 42		21 ± 6		
Malus × *domestica* Borkh.	0.75 [145]	2n = 2x = 34	664 (HiDRAS SSR database) [146]	12 ± 6	7 [147]	12 [147], 15 [108]
			2449 (Genome database for Rosaceae) [128]			

Table 4. Information on the five species analyzed in this book chapter, including genome size, ploidy, available SSR database and number of microsatellite regions included, average number of SSR employed in the articles reviewed, number of cultivars, and microsatellite used as reference.

markers, ideally two loci for each linkage group, whereas the latter group also considered the specific map position of genetic and genetic association with traits of agricultural interest.

Two important issues must be pointed out. The number of SSRs to employ should be also evaluated according to the type of analysis. For example, the EU-Project Genres CT96 No81 [139] selected six highly discriminating microsatellites, thus less than one marker per linkage group, that could be sufficient to differentiate among hundreds of grape cultivars. The same microsatellite set could be very inadequate to discriminate among clones. Moreover, it is worth noting that, in some cases, increasing the number of marker loci does not necessarily mean improving the resolution of cultivar characterization and identification. For example, Baric et al. [107] reported that extending the set of microsatellite markers to 48, from an initial analysis based on 14 SSR loci, it was impossible to improve the genetic discrimination among the 28 accessions of *Malus* × *domestica* Borkh. analyzed.

Connected to the distribution and position of the microsatellite loci within a genome, there is also the possibility to choose between genomic SSR (gSSR) and EST-derived SSR (EST-SSR). Generally, EST-SSR markers are less polymorphic than genomic SSR ones, as reported for *Triticum* spp. [93, 95] and *Solanum lycopersicum* L. [68], being the formers found in selectively more constrained regions of the genome. Of particular interest is the comparison of Leigh et al. [93] between sets of 20 EST-SSR and 12 genomic SSR markers in terms of discrimination ability among 66 varieties of *Triticum* spp. The results indicate that the panel of EST-derived SSR markers used is slightly less efficient at discriminating between hexaploid *Triticum aestivum* L. varieties compared with the second panel of genomic SSR markers. EST-SSR markers also have the disadvantage that amplicon sizes can differ from expectations, as a consequence of the undetected presence of introns in flanking regions [39]. Nevertheless, these findings support the possibility that EST-SSR markers could in the near future complement and outnumber the genomic SSR markers. In fact, EST-SSR markers should have some important advantages over genomic SSR markers. In particular, they are easily obtained by bioinformatic querying of EST databases while the development phase of genomic SSR markers is quite long and expensive; EST-SSR markers could be functionally more informative than genomic SSR markers because being associated with the transcribed regions of the genome, thus reflecting the genetic diversity inside or adjacent to the genes [149]. Moreover, the rate at which SSR flanking regions evolve is lower in expressed than nonexpressed sequences and the primers designed on these sequences are more likely to be conserved across species, thus resulting in high levels of SSR transferability [150]. A suitable combination of EST-SSR and genomic-SSR markers could be optimal for distinctiveness, uniformity, and stability testing applications for crop plant varieties [93]. Overall, the vast majority of studies are based on genomic SSR markers, and only three articles out of 90 take into account the possibility of employing EST-SSR markers.

In terms of location, nuclear SSR (nSSR) markers are largely used and more exploited than plastidial and mitochondrial SSR (cpSSR and mtSSR, respectively) markers. First, the development phase of extranuclear SSR markers is complicated: high purity chloroplast or mitochondrial DNA is typically very hard to extract due to nuclear DNA contaminations [151]. Moreover, Wolfe et al. [152] have shown that comparing nuclear, chloroplast, and mitochondrial genomes, the frequency of chloroplast genome gene silencing and replacement was half that of the nuclear

genome, and three times that of the mitochondrial genome, indicating that the evolution of mitochondrial genome has been slower and implicating lower levels of polymorphism. Nevertheless, the use of markers belonging to mitochondrial or chloroplast sequences may be useful due to their haploid nature, relative abundance, and stability in comparison with nuclear sequences. For instance, Borgo et al. [153] suggested that the circular form increases stability and resistance against heat disintegration. Boccacci et al. [113] analyzed musts and wine samples using a set of nine nSSR and seven cpSSR markers in order to identify cultivars. Findings from these studies confirm a low level of polymorphism for the extranuclear markers due to their lower frequency of mutation. Also Baleiras-Couto and Eiras-Dias [45] and Pérez-Jiménez et al. [125] have exploited this kind of SSR markers, with similar results.

The choice of the number of SSR loci usually depends on their polymorphism degree. With some exceptions for which this information is not available, the average number of marker alleles per SSR locus is equal to 7.1 for *Olea europaea* L., 3.5 for *Solanum lycopersicum* L., 8.2 for *Vitis vinifera* L., 6.9 for *Triticum* spp., 9.4 for *Malus* × *domestica* Borkh., 6.5 for olive oil, and 5.2 for wine. Both EST-SSR and cpSSR were found to be less polymorphic, with a low average number of alleles per locus, than genomic SSR markers [45, 68, 93, 113, 125]. The polymorphism degree may depend on several factors, including the SSR motif length and the SSR localization on coding or not-coding regions.

In order to estimate the level of genetic diversity detected by each microsatellite, marker frequencies are widely used to estimate the polymorphism information content (PIC, **Table 5**) values, according to the methods of Botstein et al. [154]. The authors reported the following formula for the calculation of the PIC value of an *n*-marker allele:

$$\mathrm{PIC} = 1 - \sum_{i=1}^{n} p_i^2 - \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} 2p_i^2 p_j^2, \tag{1}$$

where p_i and p_j are the population frequencies of the i^{th} and j^{th} marker alleles, respectively. A PIC > 0.5 is considered as being a highly informative marker, while 0.5 > PIC > 0.25 is an informative marker and PIC is 0.25, a slightly informative marker. As reported by Nagy et al. [155], PIC can be defined as the probability that the marker genotype of a given offspring will allow deduction, in the absence of crossing-over, of which of the two marker alleles of the affected parents it received. In other words, this parameter is a modification of the heterozygosity measure that subtracts from the H value an additional probability that an individual in a linkage analysis does not contribute information to the study. On this aspect, there is no full agreement among the authors. Some studies on olive oil [58, 122] and *Malus* × *domestica* Borkh. [101, 103], referring to Anderson et al. [156], contend that the occurrence of rare marker alleles has less impact than common marker alleles on the PIC estimates and consider that this index can be assimilated to the expected heterozygosity (H_e), calculated by the following simplified formula:

$$\mathrm{PIC} = 1 - \left(\sum_{i=1}^{n} p_i^2 \right), \tag{2}$$

where p_i is the population frequency of the *i*th marker allele.

Index	Full name	Formula	Definition	No. of papers account for it
PIC*	Polymorphism Information Content	$1 - \sum_{i=1}^{n} p_1^2 - \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} 2p_i^2 p_j^2$	Probability that the marker genotype of a given offspring will allow deduction, in the absence of crossing-over, of which of the two marker alleles of the affected parents it received [155]	36
PD**	Power of Discrimination	$1 - \sum_{i=1}^{n} p_i^2$	Probability that two randomly sampled accessions would be differentiated by their marker allele profiles [157]	14
C**	Confusion probability	$\prod_{i=1}^{n} (1 - PD_i)$	Probability that any two individuals are identical in their genotypes at all SSR loci by chance alone [157]	3
PI*	Probability of Identity	$\sum_i (p_i)^4 + \sum_i \sum_j (2p_i p_j)^2$	Probability that two individuals drawn at random from a population will have the same genotype at one marker locus [122]	21
PI_t*	Total probability of identity	$\prod_{i=1}^{n} PI_i$	Probability of two individuals sharing the same marker genotype by chance [122]	2

*pi and pj are the frequencies of the ith and jth marker alleles.
**pi is the frequency of the ith marker genotype.

Table 5. Summary information on the main parameters assessed by the 90 papers reviewed.

In addition to the PIC value, calculated taking into account allelic frequencies, there are several indexes focusing on genotype frequencies. For example, as reported by Aranzana et al. [157], other two important indexes that should be evaluated are the power of discrimination (usually PD)—or diversity index (D), as reported by Zulini et al. [71] and Martínez et al. [24]—and the confusion probability (C). The first one provides an estimate of the probability that two randomly sampled accessions of the study would be differentiated by their marker allele profiles:

$$\mathrm{PD} = 1 - \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{p}_i^2, \tag{3}$$

where p_i is the frequency of the *i*th marker genotype. As already described for the PIC, among the authors, there are different interpretations and procedures to calculate the PD index. Pasqualone et al. [25] in their study on *Olea europaea* L. genotyping reported that "the power of discrimination, sometimes referred to as polymorphism information content, or diversity index, was calculated […]," assuming in this way that PD and PIC correspond to the same parameter.

The confusion probability (C) index, also defined as the combined power of discrimination of overall loci [23], is the probability that any two cultivars are identical in their genotypes at all SSR loci by chance alone and it depends on PD. It can be estimated as follows:

$$\mathrm{C} = \prod_{\mathrm{i}=1}^{\mathrm{n}} (1 - \mathrm{PD}_i), \tag{4}$$

where PD_i is the power of discrimination value of the *i*th locus. Notwithstanding its informativeness, only three articles of the 90 reviewed take into account this value (**Table 5**). Martínez et al. [24] in their attempt to assess the genetic diversity of *Vitis vinifera* L. varieties calculated the power of discrimination index as follows:

$$\mathrm{PD} = 1 - \mathrm{C} \quad \text{being} \quad \mathrm{C} = \sum_{i=1}^{n} p_i^2, \tag{5}$$

where p_i is the frequency of different marker genotypes for a given locus. In this case, C is the probability of coincidence, corresponding to the probability that two varieties match by chance at one locus.

About 21 articles, mainly focused on the species *Vitis vinifera* L. and oil from *Olea europaea* L., report also the probability of identity (PI) index of each single SSR marker locus either in addition or in substitution of PD value (**Table 5**). This index can be estimated as follows:

$$\mathrm{PI} = \sum_{i} \left(p_i\right)^4 + \sum_{i} \sum_{j} \left(2p_i p_j\right)^2, \tag{6}$$

where p_i and p_j are the frequencies of *i*th and *j*th marker alleles, respectively. It represents the probability that two individuals drawn at random from a population will have the same genotype at one marker locus. For example, Vietina et al. [122] and Corrado et al. [58] in their studies, regarding the genetic traceability of monovarietal olive oils, refer to this value in order to determine the efficacy of the SSR marker pool to discriminate among the cultivars. Martínez et al. [24] adopted the following formula to calculate the same value:

$$\mathrm{PI} = \sum_{i} \left(p_i\right)^4 - \sum_{i} \sum_{j} \left(2p_i p_j\right)^2. \tag{7}$$

Equally interesting is the total probability of identity (PI_t) that represents a compound probability defined as the probability of two cultivars sharing the same marker genotype by chance and calculated as follows:

$$\mathrm{PI_t} = \prod_{i=1}^{n} \mathrm{PI_i}, \tag{8}$$

where PI_i is the probability of identity value of the *i*th marker locus.

Finally, Qanbari et al. [158] reported that PD and PI are complementary parameters:

$$\mathrm{PD} = 1 - \mathrm{PI}. \tag{9}$$

The use of standardized parameters is essential to make SSR data comparable across species and laboratories, and it can be especially beneficial for the preliminary evaluation of the discriminant ability and applicability of SSR marker loci.

4. The choice of the best microsatellite motifs and the problem of the null alleles

Microsatellite repeat units typically vary from one to six bases. Shortest motifs (mono- or dinucleotide repeats) usually have a high number of alleles [74], and they allow packing more loci on a given separation system, resulting in larger multiplexes. However, this kind of SSR motifs can be difficult to assay accurately. It is very common to observe a stuttering in terms of multiple bands or peaks, a phenomenon commonly caused by slippage of the DNA polymerase, but the main problem arises when there is a difference of one or two base-pairs between marker alleles: in case of homozygous loci, the electrophoretic analysis results in one main band or peak, but with heterozygous loci very often one of the two marker alleles is masked by the stutter. SSR markers containing trinucleotide or higher order repeats usually eliminate this technical problem because target sequences appear to be significantly less prone to slippage [52]. Nevertheless, microsatellite loci with long motifs are known to be less polymorphic and, in some cases, due to lack of stutter bands or peaks, which is not always possible to distinguish SSR amplicons from other aspecific PCR products and it may lead to an overestimation of the level of polymorphism of these loci [159].

Among the 90 studies we surveyed, only 25 of them specify the length of the SSR motifs employed and very few justifies the choice. Cipriani et al. [80] performed two distinct molecular analyses on the same set of cultivars, using the genetic profiles obtained from the two sets of microsatellites, the dinucleotide repeats from one side, and the tri-, tetra-, and pentanucleotide repeats from the other, with the aim of comparing their performance in the discrimination of the genotypes analyzed. Both microsatellite data sets produced identical consensus tree topology, but the authors underlined that dinucleotide SSR markers scored a higher number of alleles per locus, and consequently, a potentially higher power for identifying and distinguishing closely related genotypes. On the other hand, the microsatellite dataset based on tri-, tetra-, and pentanucleotide SSR markers proved to have the advantage of ease in scorability, while maintaining a very high power of discrimination for successful genotyping of the *Vitis vinifera* L. cultivars.

Microsatellites have also been classified according to the type of repeat sequence as perfect or imperfect, according to the occurrence of simple or uneven repeats, respectively [160]. The preference should be given to perfect motifs because using imperfect ones, there is no more equivalency between fragment length and amplicon sequence, and hence several sequences can correspond to a given length variant [39]. This is the reason why only four studies employed imperfect SSRs among the 25 ones specifying the motifs.

The occurrence of null alleles is something to avoid when using SSR markers for genotyping plant materials. A microsatellite null allele is any marker allele at a genomic locus that consistently fails to amplify by the polymerase chain reaction, resulting in the lack of detectable amplicons. Lack of amplified fragments could preclude the detection of heterozygous loci, which would be computed as homozygotes. In the same way, null alleles at homozygous loci are characterized by a complete lack of amplification with the consequent production of missing data. On the whole, null alleles may interfere with the genetic identification of cultivars, by wrongly reducing the genetic diversity among accessions [149]. In the 90 studies surveyed, only 38 of them estimated the probability of null alleles, mainly using the formula of Brookfield [161]:

$$r = \frac{H_e - H_o}{1 + H_e} \tag{10}$$

being H_e the expected heterozygosity and H_o the observed heterozygosity.

5. Comparisons across studies of SSR-based genotyping: Reference marker sets and reference plant varieties

In most cases, it is impossible to make valid comparisons across studies on the same species since different sets of SSR loci are used in different laboratories [162]. For some species, the choice of microsatellites begins to be fairly uniform (**Table 4**). For instance, almost all of the studies aimed to genotype *Olea europaea* L. cultivars make use of SSR markers belonging to four main datasets developed by Sefc et al. [163], Carriero et al. [164], Cipriani et al. [165], and de La Rosa et al. [166]. Based on these studies, two informal universal sets of SSR markers were proposed for genotyping *Olea europaea* L. cultivars by Doveri et al. [52] and Baldoni et al. [53]. Cipriani et al. [74] suggested a list of 38 markers with excellent quality of peaks, high power of discrimination, and uniform genome distribution (1–3 markers/chromosome) for genotyping *Vitis vinifera* L. cultivars. Li et al. [144] assembled a reference kit of SSR markers for genetic analysis in *Triticum* spp. that comprises 46 microsatellites. Moriya et al. [108] developed a set of SSR markers for genotyping *Malus* × *domestica* Borkh. cultivars, which includes 15 microsatellites. Not only independent research works, but also some international programs and projects attempted to pursue this goal. The European Cooperative Programme for Plant Genetic Resources (ECPGR) has recommended a new set of 12 SSR marker loci distributed in different linkage groups of the *Malus* × *domestica* Borkh. genome, organized in three multiplexes and designed for a four-dye system [147]. Comparable considerations have been presented within two projects focused on the grapevine genetic resources conservation and characterization (EU-project GENRES CT96 No 81, [139]) and on the Traceability of Origin and Authenticity of Olive Oil (Oliv-Track, [167]). It is worth noting that, to the best of our knowledge, for *Solanum Lycopersicum* L., no SSR set of reference has been proposed yet.

Unfortunately, by establishing a reference set of microsatellite markers to use in each analysis for a given species, it is not sufficient to ensure the comparability among different studies and the reproducibility among different laboratories. Some tests have been carried out in order to investigate the reproducibility of SSR data produced by different laboratories under varying local conditions. Four different laboratories performed independent marker analyses on a common set of 21 DNA samples of *Olea europaea* L. cultivars and with the same set of SSR markers, using different DNA polymerase enzymes, PCR cycling conditions, amplicon separation, and visualization methods [53]. The results are not encouraging. Many cases of allele drop out and discrepancies in allele length, up to five nucleotides for identical microsatellite loci, were recorded. This finding is probably attributable to a combination of different equipments, different sequencers, and different internal ladders, which may have affected the relative mobility estimates leading to noncomparable electropherograms. Similar results have been achieved from ten laboratories distributed in seven countries that analyzed the same 46 *Vitis vinifera*, L. cultivars at the same 6 SSR loci [72].

One of the main discoveries is that the specific microsatellite sequence dramatically influences the efficiency of analysis. Marmiroli et al. [168] showed that the repeatability of results among different laboratories was good enough for some microsatellites but rather low for others, confirming that the choice of SSR loci and of their primers is crucial for an efficient analysis.

Despite all the precautions and the establishment of a reference set of SSR markers, some residual variation in laboratory equipment and procedures cannot be completely avoided, and representative reference material with many different alleles should be adopted by all laboratories involved in a genotyping program for a given species [162]. For this purpose, 21 out of 90 studies included reference cultivars, promoting new ones or exploiting cultivars already used as reference in previous works. Independent researches and international institutions are trying to find an agreement filling lists of reference accessions in order to prevent that each group uses its own reference cultivars and to standardize all works performed on these species. For example, the ECPGR has chosen eight *Malus* × *domestica* Borkh. cultivars as reference set for this species [147]. Baldoni et al. [53] and Doveri et al. [52] proposed two different lists of reference cultivars for *Olea europaea* L. (**Table 4**).

Even if this approach is fully applicable also to the crop derivatives here taken into account (olive oil and wine), there are some additional aspects that must be considered when talking about processed products. First, sometimes, it is very difficult to make SSR marker analyses on food products and beverages because of the low DNA quantity and the lack of DNA integrity. For example, Baleiras-Couto and Eiras-Dias [45] reported their difficulties to investigate wines after about eight months of fermentation, as well as Recupero et al. [115] highlighted technical problems during the isolation of genomic DNA from Nebbiolo wine. Nevertheless, both of them managed to characterize must. For olive oil, Martins-Lopes et al. [119] as well as Vietina et al. [122], took advantage from extraction methods able to give good yield of genomic DNA and PCR amplificability. It is therefore evident how an optimized DNA extraction method is also a crucial step to carry out a reliable study on the applicability of molecular markers for identifying the varietal origin or assessing the varietal composition of crop plant derivatives.

It is not trivial considering the match between genetic profiles of crop plants and their derivatives. In this regard, there are some contrasting points of view. In the review of Agrimonti et al. [169], it is reported that several authors (e.g., [46, 118, 120]) have noticed a satisfying conformity between olive oil and leaf profiles with SSR markers. On the contrary, Doveri et al. [117] have proposed a cautionary note about the use of SSR markers, stressing the nonperfect concordance between the molecular genetic profiles of the olive oil and the original leaf sample. Furthermore, it is necessary to underline the extreme difficulty in characterizing multivarietal derivatives through SSR analysis. Most of the Italian PDO wines and olive oils are produced blending two or more cultivars in percentages strictly defined in the production regulation. In these cases, each SSR locus is represented by the combination of the marker alleles of each variety. For examples, Baleiras-Couto and Eiras-Dias [45], after having analyzed with six SSR markers in different divarietal musts at different percentages, reported results that confirm the complexity and difficulty of assessing multiple genotypes.

6. Conclusions

The genetic characterization of plant varieties by means of multilocus genotyping through SSR markers in the main crop species is still not based on standardized protocols making the acquisition of reproducible and transferable datasets difficult. What emerges from the analysis of the literature is a lack of wider consensus among the authors regarding the strategy to design and to adopt for genotyping plant varieties with SSR markers. This finding highlights the urgent need to establish a common procedure.

Some conclusions of general validity can be drawn on the basis of the articles here reviewed. First of all, it is quite difficult to define exactly the ideal number of microsatellite loci to assay. Usually, the number of SSR markers depends on the type and goal of the analysis. If the purpose is merely to distinguish among two or more cultivars (i.e., individual genotypes), it is possible to adopt an "as simple as possible strategy." For example, a novel approach called the cultivar identification diagram (CID) strategy has been recently developed. This method was designed so that, at each step, a polymorphic marker generated from each PCR analysis directly allows the separation of cultivar samples [109]. In this specific study, eight is considered the minimum number of SSR markers necessary to distinguish 60 cultivars in *Malus* × *domestica* Borkh.. Supposedly, the number of SSR markers could depend on the number of cultivars to distinguish, on their relationship and on the polymorphic degree of each marker locus. In this regard, we suggest AMaCAID [170] and UPIC [171], two very interesting tools that able the investigation of the minimum number of markers required to distinguish a specific number of accessions and, thus, the identification of the best marker combination that maximizes the genetic information.

When the purpose is to genetically characterize a cultivar in order to fulfill the requirements of a varietal register that could include hundreds or thousands of different varieties, the selection of SSR markers should be oriented to an exhaustive representation of the genome as whole. This is the reason why different authors consider one or two microsatellite for each linkage group for the minimum number required to reconstruct a reliable and selectable genotype for a given plant accession. For instance, Cipriani et al. [74] implemented an efficient method for *Vitis vinifera* L. fingerprinting using a set of 38 microsatellite marker loci scattered throughout the genome. In particular, two SSR loci were carefully chosen, on average, for each linkage group, selecting the best ones in terms of polymorphism information content (PIC) and power of discrimination (PD, **Figure 1**).

It is worth noting that despite some international programs and projects attempted to establish reference SSR set, there is still a lack of wider consensus. For instance, in 2003, the partners of the EU-project Genres CT96 No81 [139] agreed on the utilization of six highly polymorphic SSR-markers for the identification of *Vitis vinifera* L. cultivars, but, since then, several studies continue to be performing using a higher number of markers [74, 76, 78, 84, 86]. As reported by Cipriani et al. [74], grape varieties selected in Western Europe, which account for most of the worldwide production of wine, likely have extensive coancestry that is a common origin from the hybridization of a few ancestors. Because of this, using too few markers for fingerprinting could hamper the discrimination of sibling varieties. For this

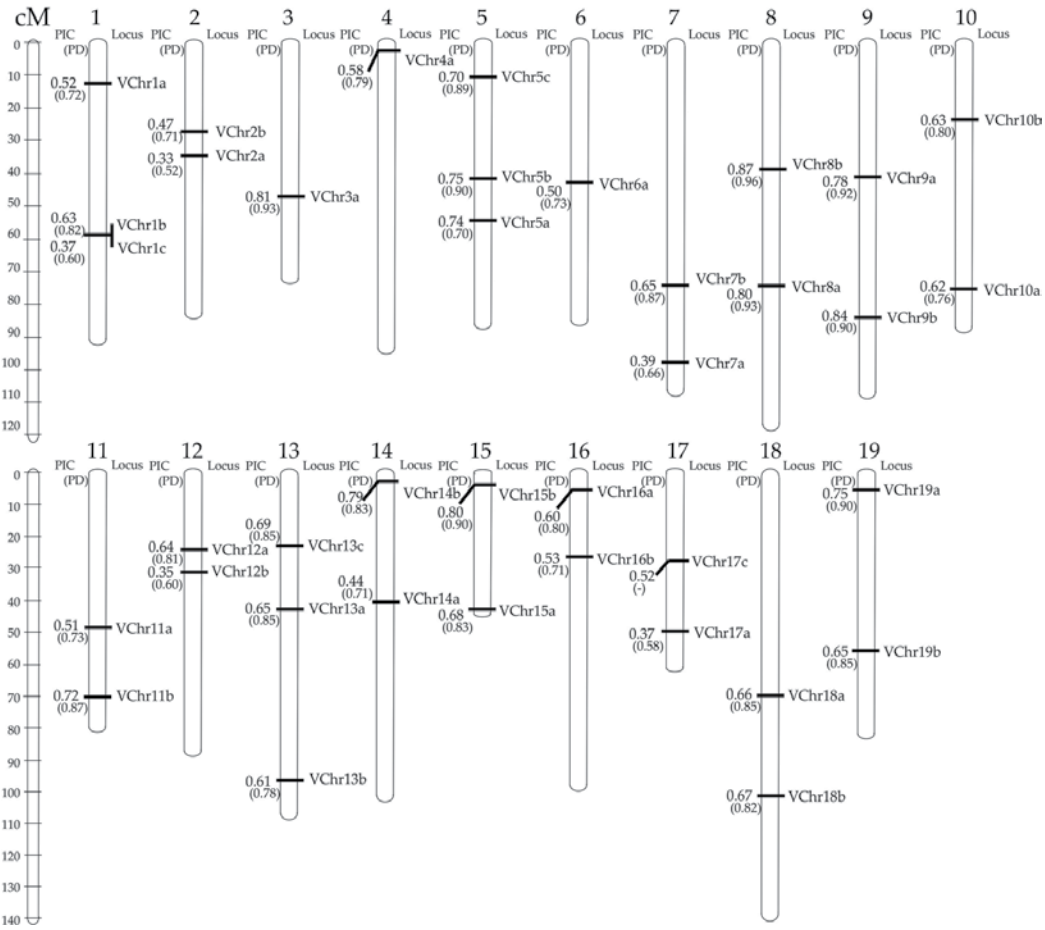

Figure 1. Schematic representation of the 19 basic linkage groups of *Vitis vinifera* L. with indication of the 38 mapped marker loci chosen on the basis of their discriminant informativeness. In addition to the marker name, each locus reports the individual power of discrimination value (PD) and the polymorphism information content (PIC). Figure modified from Cipriani et al. [74].

reason, they recommend using at least 19 markers (among the 38 markers employed in their work). In general, for the selection of the panel of SSR markers, the following criteria should be followed. Based on previous works, the SSR marker loci with the highest number of marker alleles and the highest PIC and PD scores should have the priority. In addition, the position of the SSR markers across the genome, as mapped in different linkage groups and associated with adjacent chromosome blocks, is crucial in order to get a representative multilocus marker genotype. In fact, microsatellites retrieved from noncoding regions (genomic SSR markers) meet this requirement more precisely than those derived from expressed regions (EST-SSR markers). Nevertheless, the application of EST-SSR markers cannot be excluded when phylogenetic relationships have to be investigated. It is well known that SSR markers belonging to coding regions may be functionally more informative than those

http://dx.doi.org/10.5772/intechopen.70756

deriving from noncoding ones, because they are associated with transcribed regions of the genome and thus reflecting the genetic diversity within genes or adjacent to genes [149]. Moreover, the association with trait loci with Mendelian inheritance is particularly requested in case of needs for marker-assisted selection (MAS).

About the localization of target microsatellites in the cellular genomes, nuclear SSR (nSSR) markers seem to be more polymorphic than plastidial and mitochondrial ones (cpSSR and mtSSR markers) and because of their co-dominance, the former are the only markers useful for assessing the genetic value of breeding stocks, even if the abundance and the haploid nature of the latter ones make them particularly suitable for phylogenetic and genetic diversity studies.

As far as the microsatellite repeat is concerned, the most recommended motifs are dinucleotide and trinucleotide repeats, whereas mononucleotide repeats need caution because of technical drawbacks, which can be experienced in the allele discrimination. SSR markers with tetranucleotide or more repeats display a polymorphism inversely proportional to the complexity of the motif. The so-called perfect SSR markers are preferred because of their ease of scorability. It is also worth emphasizing that the choice of SSR markers is also dependent on the occurrence of null alleles for a given locus and the informativeness in terms of allele diversity indexes. First of all, any rate of null alleles can underestimate heterozygosity and affect the reliability of the analysis. Second, the calculation of some informative indexes cannot be underrated: it represents a crucial step of the planning of any analysis. What emerges from the 90 studies here reviewed is a lack of wider consensus among the authors regarding the best informative index to calculate and this makes the comparison difficult also among studies performed on the same cultivars and with the same markers. The power of discrimination (PD), the confusion probability (C), the polymorphism information content (PIC), the probability of identity (PI), the total probability of identity (PI_t), and the probability of null allele (r) are all parameters able to describe exhaustively the efficiency of the set of SSR markers used in a given species.

In conclusion, there is the urgent need to establish a common procedure for SSR genotyping with a universal set of marker loci to be analyzed in each species. In parallel, the reference varieties must be defined in each species in order to maximize not only the reproducibility but also the portability of marker data, being aware that the residual variation in laboratory procedures and equipment cannot be completely avoided.

Author details

Fabio Palumbo and Gianni Barcaccia*

*Address all correspondence to: gianni.barcaccia@unipd.it

Laboratory of Genomics and Plant Breeding, DAFNAE – University of Padua, Campus of Agripolis, Legnaro, Italy

References

[1] Food and Agriculture Organization of the United Nations Statistics Division [Internet]. 2016 [cited 2016 May 30]. Available from: http://faostat3.fao.org/mes/methodology_list/E

[2] Food and Agriculture Organization of the United Nations. Value of Agricultural Production [Internet]. 2016 [cited 2016 May 30]. Available from: http://faostat3.fao.org/download/Q/QV/E

[3] Science IS for H. International Code of Nomenclature for Cultivated Plants. 9th 2016 p. 190

[4] List of Italian PDO and PGI [Internet]. [cited 2016 May 30]. Available from: https://www.politicheagricole.it/flex/cm/pages/ServeBLOB.php/L/IT/IDPagina/2090

[5] Sector Plan, Collaboration among MIPAAF, CRA, INEA and ISMEA [Internet]. [cited 2016 May 30]. Available from: http://www.pianidisettore.it/flex/FixedPages/Common/RegistroVarietale.php/L/IT

[6] National Register of Horticultural Varieties: Ministry of Agricultural, Food and Forestry Policies [Internet]. [cited 30-05-2016]. Available from: http://www.politicheagricole.it/flex/cm/pages/ServeBLOB.php/L/IT/IDPagina/7051

[7] National Register of Vitis Varieties. Ministry of Agricultural, Food and Forestry Policies [Internet]. [cited 2016 May 30]. Available from: http://catalogoviti.politicheagricole.it/catalogo.php

[8] National Register of Agrarian Varieties. Ministry of Agricultural, Food and Forestry Policies [Internet]. [cited 2016 May 30]. Available from: http://www.politicheagricole.it/flex/cm/pages/ServeBLOB.php/L/IT/IDPagina/7051

[9] List of Orientation among Fruit Trees: Ongoing Project Between MIPAAF and Regions [Internet]. [cited 2016 May 30]. Available from: http://plantgest.imagelinenetwork.com

[10] List of Italian DOC and DOCG Wines, Considered by the UE as PDO Products [Internet]. [cited 2016 May 30]. Available from: https://www.politicheagricole.it/flex/cm/pages/ServeBLOB.php/L/IT/IDPagina/4625

[11] Legislative Decree: Attuazione Della Direttiva 2008/90 Relativa Alla Commercializzazione Dei Materiali di Moltiplicazione Delle Piante da Frutto Destinate alla Produzione di Frutti. Article 7. Italy

[12] Lopez-Vizcón C, Ortega F. Detection of mislabelling in the fresh potato retail market employing microsatellite markers. Food Control. 2012 Aug [cited 2014 Dec 12];**26**(2):575-579. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0956713512000941

[13] Mahajan V, Jakse J, Havey MJ, Lawande KE. Genetic fingerprinting of onion cultivars using SSR markers. Indian Journal of Horticulture. 2009;**66**(1):62-68

[14] Mũnoz-Falcón JE, Vilanova S, Plazas M, Prohens J. Diversity, relationships, and genetic fingerprinting of the Listada de Gandía eggplant landrace using genomic SSRS and EST-SSRS. Scientia Horticulturae (Amsterdam). 2011;**129**(2):238-246

[15] De Mattia F, Imazio S, Grassi F, Lovicu G, Tardaguila J, Failla O, et al. Genetic characterization of sardinia grapevine cultivars by SSR markers analysis. Journal International des Sciences de la Vigne et du Vin. 2007;**41**(4):175-184

[16] Rao R, Bencivenni M, La Mura M, Araujo-Burgos T, Corrado G. Molecular characterisation of Vesuvian apricot cultivars: Implications for the certification and authentication of protected plant material. The Journal of Horticultural Science and Biotechnology. 2010;**85**(1):42-47

[17] Motilal L, Butler D. Verification of identities in global cacao germplasm collections. Genetic Resources and Crop Evolution. 2003;**50**(8):799-807

[18] Dossett M, Bassil NV, Finn CE. SSR Fingerprinting of Black Raspberry Cultivars Shows Discrepancies in Identification. 2012. p. 49-54

[19] Benjak A, Ercisli S, Vokurka A, Maletić E, Pejić I. Genetic relationships among grapevine cultivars native to Croatia, Greece and Turkey. VITIS-Journal of Grapevine Research. 2005;**44**(2):73-77

[20] Zhou L, Matsumoto T, Tan H-W, Meinhardt LW, Mischke S, Wang B, et al. Developing single nucleotide polymorphism markers for the identification of pineapple (*Ananas comosus*) germplasm. Horticulture Research. 2015;**2**(October):15056. Available from: http://www.nature.com/articles/hortres201556

[21] Chandra A, Grisham MP, Pan Y. Allelic divergence and cultivar-specific SSR alleles revealed by capillary electrophoresis using fluorescence-labeled SSR markers in sugarcane. Genome. 2014;**57**(6):363-372. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25247737

[22] Akkaya MS, Buyukunal-Bal EB. Assessment of genetic variation of bread wheat varieties using microsatellite markers. Euphytica. 2004;**135**(2):179-185. Available from: http://link.springer.com/10.1023/B:EUPH.0000014908.02499.41

[23] Rekik I, Salimonti A, Kamoun NG, Muzzalupo I, Lepais O, Gerber S, et al. Characterization and identification of tunisian olive tree varieties by microsatellite markers. Hortscience. 2008;**43**(5):1371-1376

[24] Martínez LE, Cavagnaro PF, Masuelli RW, Zúñiga M. SSR-based assessment of genetic diversity in South American *Vitis vinifera* varieties. Plant Science. 2006 Jun;**170**(6):1036-1044. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0168945205004292 [cited 2014 Dec 12]

[25] Pasqualone A, Di Rienzo V, Nasti R, Blanco A, Gomes T, Montemurro C. Traceability of Italian Protected Designation of Origin (PDO) table olives by means of microsatellite

molecular markers. Journal of Agricultural and Food Chemistry. 2013 Mar 27;**61**(12): 3068-3073. Available from http://www.ncbi.nlm.nih.gov/pubmed/23461435

[26] Pasqualone A, Alba V, Mangini G, Blanco A, Montemurro C. Durum wheat cultivar traceability in PDO Altamura bread by analysis of DNA microsatellites. European Food Research and Technology. 2010;**230**(5):723-729

[27] Caramante M, Corrado G, Monti LM, Rao R. Simple sequence repeats are able to trace tomato cultivars in tomato food chains. Food Control. 2011 Mar;**22**(3–4):549-554. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0956713510003348 [cited 2014 Dec 12]

[28] Wadl PA, Wang X, Moulton JK, Hokanson SC, Skinner JA, Rinehart TA, et al. Transfer of *Cornus florida* and *C. kousa* simple sequence repeats to selected cornus (Cornaceae) Species. Journal of the American Society for Horticultural Science. 2010;**135**(3):279-288

[29] Lacis G, Rashal I, Ruisa S, Trajkovski V, Iezzoni AF. Assessment of genetic diversity of Latvian and Swedish sweet cherry (*Prunus avium* L.) genetic resources collections by using SSR (microsatellite) markers. Scientia Horticulturae (Amsterdam). 2009;**121**(4):451-457

[30] Korinsak S, Sriprakhon S, Sirithanya P, Jairin J, Vanavichit A, Toojinda T. Identification of microsatellite markers (SSR) linked to a new bacterial blight resistance gene xa33 (t) in rice cultivar "Ba7". Maejo International Journal of Science and Technology. 2009;**3**(2):235-247

[31] Fan S, Bielenberg DG, Zhebentyayeva TN, Reighard GL, Okie WR, Holland D, et al. Mapping quantitative trait loci associated with chilling requirement, heat requirement and bloom date in peach (*Prunus persica*). The New Phytologist. 2010;**185**(4):917-930

[32] Ashkani S, Rafii MY, Rusli I, Sariah M, Abdullah SNA, Rahim HA, et al. SSRs for marker-assisted selection for blast resistance in rice (*Oryza sativa* L.). Plant Molecular Biology Reporter. 2012;**30**(1):79-86

[33] Cheta Kumar MR, Vishwanath K, Shivakumar N, Rajendra Prasad S, Rahda BR. Utilization of SSR markers for seed purity testing in popular rice hybrids (*Oryza sativa* L.). Annals of Plant Protection Sciences. 2012;**1**(1):1-5

[34] Ibaňez J, Van Eeuwijk FA. Microsatellite profiles as a basis for intellectual property protection in grape. Acta Horticulturae. 2003;**603**:41-47

[35] Fan L, Zhang MY, Liu QZ, Li LT, Song Y, Wang LF, et al. Transferability of newly developed pear SSR markers to other Rosaceae species. Plant Molecular Biology Reporter. 2013;**31**(6):1271-1282

[36] Satya P, Paswan PK, Ghosh S, Majumdar S, Ali N. Confamiliar transferability of simple sequence repeat (SSR) markers from cotton (*Gossypium hirsutum* L.) and jute (*Corchorus olitorius* L.) to twenty two Malvaceous species. 3 Biotech. 2016;**6**(1):65. Available from: http://link.springer.com/10.1007/s13205-016-0392-z

[37] Emanuelli F, Lorenzi S, Grzeskowiak L, Catalano V, Stefanini M, Troggio M, et al. Genetic diversity and population structure assessed by SSR and SNP markers in a large germplasm collection of grape. BMC Plant Biology. 2013;**13**(1):39. Available from: http://www.biomedcentral.com/1471-2229/13/39

[38] Filippi CV, Aguirre N, Rivas JG, Zubrzycki J, Puebla A, Cordes D, et al. Population structure and genetic diversity characterization of a sunflower association mapping population using SSR and SNP markers. BMC Plant Biology. 2015;**15**(1):52. Available from: http://www.biomedcentral.com/1471-2229/15/52

[39] Guichoux E, Lagache L, Wagner S, Chaumeil P, Léger P, Lepais O, et al. Current trends in microsatellite genotyping. Molecular Ecology Resources. 2011;**11**(4):591-611

[40] Masi P, Spagnoletti Zeuli PL, Donini P. Development and analysis of multiplex microsatellite markers sets in common bean (*Phaseolus vulgaris* L.). Molecular Breeding. 2003;**11**(4):303-313

[41] Ganal MW, Röder MS. Microsatellite and SNP markers in wheat breeding. Genomics. 2007;**2**:1-24

[42] Gonzaga ZJ. Evaluation of SSR and SNP Markers for molecular breeding in rice. Plant Breeding and Biotechnology. 2015;**3**(2):139-152. Available from: http://www.plantbreedbio.org/journal/view.html?uid=185&page=&sort=&scale=10&all_k=&s_t=&s_a=&s_k=&s_v=3&s_n=2&spage=&pn=search&year=&vmd=Full

[43] Sánchez-Pérez R, Ruiz D, Dicenta F, Egea J, Martínez-Gómez P. Application of simple sequence repeat (SSR) markers in apricot breeding: Molecular characterization, protection, and genetic relationships. Scientia Horticulturae (Amsterdam). 2005;**103**(3):305-315

[44] Sardaro MLS, Marmiroli M, Maestri E, Marmiroli N. Genetic characterization of Italian tomato varieties and their traceability in tomato food products-Sardaro-2012-Food Science & Nutrition-Wiley Online Library. Food Science & Nutrition. 2013 Jan [cited 2014 Dec 1];**1**(1):54-62. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3951568&tool=pmcentrez&rendertype=abstract

[45] Baleiras-Couto MM, Eiras-Dias JE. Detection and identification of grape varieties in must and wine using nuclear and chloroplast microsatellite markers. Analytica Chimica Acta. 2006 Mar;**563**(1–2):283-291. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0003267005016922 [cited 2014 Nov 13]

[46] Pasqualone A, Montemurro C, Caponio F, Blanco A. Identification of virgin olive oil from different cultivars by analysis of DNA microsatellites. Journal of Agricultural and Food Chemistry. 2004 Mar 10;**52**(5):1068-1071. Available from http://www.ncbi.nlm.nih.gov/pubmed/14995099

[47] Laurens F, Durel CE, Lascostes M. Molecular characterization of French local apple cultivars using SSRs. Acta Horticulturae. 2004;**663**:639-642

[48] Ercisli S, Ipek A, Barut E. SSR marker-based DNA fingerprinting and cultivar identification of olives (*Olea europaea*). Biochemical Genetics. 2011 Oct;**49**(9–10):555-561. Available from [cited 2014 Dec 12]http://www.ncbi.nlm.nih.gov/pubmed/21476017

[49] Montemurro C, Simeone R, Pasqualone A, Ferrara E, Blanco A. Genetic relationships and cultivar identification among 112 olive accessions using AFLP and SSR markers. The Journal of Horticultural Science and Biotechnology. 2005;**80**(1):105-110

[50] Díaz A, De la Rosa R, Martín A, Rallo P. Development, characterization and inheritance of new microsatellites in olive (*Olea europaea* L.) and evaluation of their usefulness in cultivar identification and genetic relationship studies. Tree Genetics & Genomes. 2006 Mar 28;**2**(3):165-175. Available from: http://link.springer.com/10.1007/s11295-006-0041-5 [cited 2014 Dec 12]

[51] Taamalli W, Geuna F, Bassi D, Daoud D, Zarrouk M. SSR marker based DNA fingerprinting of Tunisian olive (*Olea europaea* L.) varieties. Journal of Agronomy. 2008;**7**:176-181

[52] Doveri S, Sabino Gil F, Díaz A, Reale S, Busconi M, da Câmara Machado A, et al. Standardization of a set of microsatellite markers for use in cultivar identification studies in olive (*Olea europaea* L.). Scientia Horticulturae (Amsterdam). 2008 May;**116**(4):367-373. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0304423808000496; [cited 2014 Dec 11]

[53] Baldoni L, Cultrera NG, Mariotti R, Ricciolini C, Arcioni S, Vendramin GG, et al. A consensus list of microsatellite markers for olive genotyping. Molecular Breeding. 13-05-2009;**24**(3):213-231. Available from: http://link.springer.com/10.1007/s11032-009-9285-8 [cited 2014 Dec 12]

[54] Muzzalupo I, Stefanizzi F, Perri E. Evaluation of olives cultivated in southern italy by simple sequence repeat markers. Hortscience. 2009;**44**(3):582-588

[55] Muzzalupo I, Stefanizzi F, Salimonti A, Falabella R, Perri E. Microsatellite markers for identification of a group of Italian olive accessions. Science in Agriculture. 2009;**66**(5): 685-690

[56] Bracci T, Sebastiani L, Busconi M, Fogher C, Belaj A, Trujillo I. SSR markers reveal the uniqueness of olive cultivars from the Italian region of Liguria. Scientia Horticulturae (Amsterdam). 2009 Sep;**122**(2):209-215. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0304423809002155 [cited 2014 Dec 9]

[57] Alba V, Montemurro C, Sabetta W, Pasqualone A, Blanco A. SSR-based identification key of cultivars of *Olea europaea* L. diffused in Southern-Italy. Scientia Horticulturae (Amsterdam). 2009 Dec;**123**(1):11-16. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0304423809003550 [cited 2014 Dec 12]

[58] Corrado G, Imperato A, la Mura M, Perri E, Rao R. Genetic diversity among olive varieties of Southern Italy and the traceability of olive oil using SSR markers. The Journal of Horticultural Science and Biotechnology 2011;**86**(5):461–466.

[59] Ipek A, Barut E, Gulen H, Ipek M. Assessment of inter- and intra-cultivar variations in olive using SSR markers. Science in Agriculture. 2012;**69**(5):327-335 Available from: http://www.scopus.com/inward/record.url?eid=2-s2.0-84867483027&partnerID=40&md5=c6c81882560aca1bd542c465b04eb2f1

[60] Muzzalupo I, Salimonti A, Stefanizzi F, Falabella R, Perri E. Microsatellite Markers for Characterization and Identification of Olive (*Olea europaea*) Cultivars in South Italy. VI International Symposium on Olive Growing. 2012;**949**:67-70 Available from: https://

apps.webofknowledge.com/full_record.do?product=UA&search_mode=GeneralSearch&qid=5&SID=V1H3TbxBXvPOXeKSsLj&page=1&doc=1

[61] Las Casas G, Scollo F, Distefano G, Continella A, Gentile A, La Malfa S. Molecular characterization of olive (*Olea europaea* L.) Sicilian cultivars using SSR markers. Biochemical Systematics and Ecology. 2014 Dec;**57**:15-19. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0305197814002026 [cited 2014 Dec 4]

[62] Abdessemed S, Muzzalupo I, Benbouza H. Assessment of genetic diversity among Algerian olive (*Olea europaea* L.) cultivars using SSR marker. Scientia Horticulturae (Amsterdam). 2015;**192**:10-20. Available from: http://dx.doi.org/10.1016/j.scienta.2015.05.015

[63] Sakar E, Unver H, Ercisli S. Genetic diversity among historical olive (*Olea europaea* L.) genotypes from Southern anatolia based on SSR markers. Biochemical Genetics. 2016; **54**(6):842-853. Available from: http://link.springer.com/10.1007/s10528-016-9761-x

[64] Caramante M, Rao R, Monti LM, Corrado G. Discrimination of "San Marzano" accessions: A comparison of minisatellite, CAPS and SSR markers in relation to morphological traits. Scientia Horticulturae (Amsterdam). 2009 May;**120**(4):560-564. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0304423808005347 [cited 2014 Dec 12]

[65] Meng F, Xu X, Huang F, Li J. Analysis of genetic diversity in cultivated and wild tomato varieties in Chinese market by RAPD and SSR. Agricultural Sciences in China. 2010 Oct;**9**(10):1430-1437. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1671292709602340 [cited 2014 Dec 12]

[66] El-awady MAM, El-tarras AAE, Hassan MM. Genetic diversity and DNA fingerprint study in tomato (*Solanum lycopersicum* L) cultivars grown in Egypt using simple sequence repeats (SSR) markers. African Journal of Biotechnology. 2012;**11**(96):16233-16240

[67] Todorovska E, Ivanova A, Ganeva D, Pevicharova G, Molle E, Bojinov B, et al. Assessment of genetic variation in Bulgarian tomato (*Solanum lycopersicum* L.) genotypes, using fluorescent SSR genotyping platform. Biotechnology and Biotechnological Equipment. 2014 Jun 4;**28**(1):68-76. Available from: http://www.tandfonline.com/doi/abs/10.1080/13102818.2014.901683 [cited 2014 Dec 12]

[68] Korir NK, Diao W, Tao R, Li X, Kayesh E, Li A, et al. Genetic diversity and relationships among different tomato varieties revealed by EST-SSR markers. Genetics and Molecular Research. 2014 Jan;**13**(1):43-53. Available from http://www.ncbi.nlm.nih.gov/pubmed/24446286

[69] Miskoska-Milevska E, Popovski Z, Dimitrievska B, Bandzo K. DNA microsatellite analysis for tomato genetic differentiation. Genetika. 2015;**47**(3):1123-1130. Available from: http://www.doiserbia.nb.rs/Article.aspx?ID=0534-00121503123M

[70] Mercati F, Longo C, Poma D, Araniti F, Lupini A, Mammano MM, et al. Genetic variation of an Italian long shelf-life tomato (*Solanum lycopersicon* L.) collection by using SSR and morphological fruit traits. Genetic Resources and Crop Evolution. 2014;**62** (5):721-732. Available from: http://link.springer.com/10.1007/s10722-014-0191-5

[71] Zulini L, Russo M, Peterlunger E. Genotyping wine and table grape cultivars from Apulia (Southern Italy) using microsatellite markers. Vitis. 2002;**41**(4):183-187

[72] This P, Jung A, Boccacci P, Borrego J, Botta R, Costantini L, et al. Development of a standard set of microsatellite reference alleles for identification of grape cultivars. Theoretical and Applied Genetics. 2004 Nov;**109**(7):1448-1458. Available from [cited 2014 Dec 9]http://www.ncbi.nlm.nih.gov/pubmed/15565426

[73] Hvarleva T, Hadjinicoli A, Atanassov I, Atanassov A, Ioannou N. Genotyping *Vitis vinifera* L. cultivars of Cyprus by microsatellite analysis. Vitis. 2005;**44**(2):93-97

[74] Cipriani G, Marrazzo MT, Di Gaspero G, Pfeiffer A, Morgante M, Testolin R. A set of microsatellite markers with long core repeat optimized for grape (*Vitis* spp.) genoty ping. BMC Plant Biology. 2008 Jan;**8**:127. Available from: http://www.pubmedcentral. nih.gov/articlerender.fcgi?artid=2625351&tool=pmcentrez&rendertype=abstract; [cited 2014 Dec 12]

[75] Salmaso M, Forestan C, Varotto S, Lucchin M. Biodiversity of local Italian grapevine cultivars detected by SSR markers. ISHS Acta Hort: Proc. IXth Intl. Conf. on Grape Genetics and Breeding. 2009;**827**:137-142

[76] Marinoni DT, Raimondi S, Ruffa P, Lacombe T, Schneider A. Identification of grape cultivars from liguria (north-western Italy). VITIS-Journal of Grapevine Research. 2009;**48**(4):175-183

[77] Jahnke G, Májer J, Lakatos A, Molnár JG, Deák E, Stefanovits-Bányai É, et al. Isoenzyme and microsatellite analysis of *Vitis vinifera* L. varieties from the Hungarian grape germplasm. Scientia Horticulturae (Amsterdam). 2009 Apr;**120**(2):213-221. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0304423808004470 [cited 2014 Dec 12]

[78] Zoghlami N, Riahi L, Laucou V, Lacombe T, Mliki A, Ghorbel A, et al. Origin and genetic diversity of Tunisian grapes as revealed by microsatellite markers. Scientia Horticulturae (Amsterdam). 2009 May;**120**(4):479-486. Available from: http://linkinghub. elsevier.com/retrieve/pii/S030442380800527X [cited 2014 Dec 12]

[79] Veloso MM, Almandanim MC, Baleiras-Couto M, Pereira HS, Carneiro LC, Fevereiro P, et al. Microsatellite database of grapevine (*Vitis vinifera* L.) cultivars used for wine production in Portugal. Ciência e Técnica Vitivinícola. 2010;**25**(2):53-61

[80] Cipriani G, Marrazzo MT, Peterlunger E. Molecular characterization of the autochthonous grape cultivars of the region Friuli Venezia Giulia - North-Eastern Italy. VITIS-Journal of Grapevine Research. 2010;**49**(1):29-38

[81] Cipriani G, Spadotto A, Jurman I, Di Gaspero G, Crespan M, Meneghetti S, et al. The SSR-based molecular profile of 1005 grapevine (*Vitis vinifera* L.) accessions uncovers new synonymy and parentages, and reveals a large admixture amongst varieties of different geographic origin. Theoretical and Applied Genetics. 2010 Nov;**121**(8):1569-1585. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20689905; [cited 2014 Dec 12]

[82] Jahnke G, Májer J, Varga P, Szőke B, et al. Scientia Horticulturae (Amsterdam). 2011 May;**129**(1):32-37. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0304423811001129 [cited 2014 Dec 12]

[83] Moreno-Sanz P, Loureiro MD, Suárez B. Microsatellite characterization of grapevine (*Vitis vinifera* L.) genetic diversity in Asturias (Northern Spain). Scientia Horticulturae (Amsterdam). 2011 Jun;**129**(3):433-440. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0304423811002019 [cited 2014 Dec 12]

[84] Díaz-Losada E, Tato Salgado A, Ramos-Cabrer AM, Díaz-Hernández B, Pereira-Lorenzo S. Genetic and geographical structure in grapevines from northwestern Spain. The Annals of Applied Biology. 2012 Jul 10;**161**(1):24-35. Available from: http://doi.wiley.com/10.1111/j.1744-7348.2012.00548.x [cited 2014 Dec 12]

[85] Guo DL, Zhang Q, Zhang GH. Characterization of grape cultivars from China using microsatellite markers. Czech Journal of Genetics and Plant Breeding. 2013;**49**(4):164-170

[86] Basheer-Salimia R, Lorenzi S, Batarseh F, Moreno-Sanz P, Emanuelli F, Grando MS. Molecular identification and genetic relationships of Palestinian grapevine cultivars. Molecular Biotechnology. 2014 Jun;**56**(6):546-556. Available from http://www.ncbi.nlm.nih.gov/pubmed/24469973; [cited 2014 Dec 12]

[87] Galbacs Z, Molnar S, Halasz G, Kozma P, Hoffmann S, Kovacs L, et al. Identification of grapevine cultivars using microsatelite-based DNA barcodes. Vitis. 2009;**48**(1):17-24

[88] Zarouri B, Vargas AM, Gaforio L, Aller M, de Andrés MT, Cabezas JA. Whole-genome genotyping of grape using a panel of microsatellite multiplex PCRs. Tree Genetics & Genomes. 2015;**11**:17

[89] Lei W, Juan Z, Linde L, Li Z, Lijuan W, Dechang H. Genetic diversity of grape germplasm as revealed by microsatellite (SSR) markers. African Journal of Biotechnology. 2015;**14**(12):990-998. Available from: http://academicjournals.org/journal/AJB/article-abstract/DA5410651677

[90] Dograr N, Akin-Yalin S, Akkaya AM. Discriminating durum wheat cultivars using highly polymorphic simple sequence repeat DNA markers. Plant Breeding. 1999;**119**:884-886

[91] Prasad M, Varshney RK, Roy JK, Balyan HS, Gupta PK. The use of microsatellites for detecting DNA polymorphism, genotype identification and genetic diversity in wheat. TAG Theoretical and Applied Genetics. 2000 Feb;**100**(3–4):584-592. Available from: http://link.springer.com/10.1007/s001220050077

[92] Röder MS, Wendehake K, Korzun V, Bredemeijer G, Laborie D, Bertrand L, et al. Construction and analysis of a microsatellite-based database of European wheat varieties. Theoretical and Applied Genetics. 2002 Dec;**106**(1):67-73. Available from http://www.ncbi.nlm.nih.gov/pubmed/12582872; [cited 2014 Dec 12]

[93] Leigh F, Lea V, Law J, Wolters P, Powell W, Donini P. Assessment of EST- and genomic microsatellite markers for variety discrimination and genetic diversity studies in wheat. Euphytica. 2003;**133**(3):359-366

[94] Perry DJ. Identification of Canadian durum wheat varieties using a single PCR. Theoretical and Applied Genetics. 2004;**109**(1):55-61

[95] Fujita Y, Fukuoka H, Yano H. Identification of wheat cultivars using EST–SSR markers. Breeding Science. 2009;**59**(2):159-167. Available from: http://joi.jlc.jst.go.jp/JST.JSTAGE/jsbbs/59.159?from=CrossRef

[96] Schuster I, Vieira ESN, da Silva GJ, Franco FDA, Marchioro VS. Genetic variability in Brazilian wheat cultivars assessed by microsatellite markers. Genetics and Molecular Biology 2009;**32**(3):557-563

[97] Salem KFM, Röder MS, Börner A. Assessing genetic diversity of Egyptian hexaploid wheat (*Triticum aestivum* L.) using microsatellite markers. Genetic Resources and Crop Evolution. 2015;**62**(3):377-385

[98] Henkrar F, El-Haddoury J, Ouabbou H, Nsarellah N, Iraqi D, Bendaou N, et al. Genetic diversity reduction in improved durum wheat cultivars of Morocco as revealed by microsatellite markers. Science in Agriculture. 2016;**73**(2):134-141. Available from: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-90162016000200134&lng=en&nrm=iso&tlng=en

[99] Goulão L, Oliveira C. Molecular characterisation of cultivars of apple (Malus × domestica Borkh.) using microsatellite (SSR and ISSR) markers. Euphytica. 2001;**122**:81-89. Available from: http://link.springer.com/article/10.1023/A:1012691814643

[100] Kitahara K, Matsumoto S, Yamamoto T, Soejima J, Abe K. Molecular characterization of apple cultivars in Japan. Science and Technology. 2005;**130**(6):885-892

[101] Galli Z, Halász G, Kiss E, Heszky L, Dobránszki J. Molecular identification of commercial apple cultivars with microsatellite markers. Hortscience. 2005;**40**(7):1974-1977

[102] Melchiade D, Foroni I, Corrado G, Santangelo I, Rao R. Authentication of the "Annurca" apple in agro-food chain by amplification of microsatellite loci. Food Biotechnology. 2007 Mar 6;**21**(1):33-43. Available from: http://www.tandfonline.com/doi/abs/10.1080/08905430701191114 [cited 2014 Dec 12]

[103] Garkava-Gustavsson L, Kolodinska Brantestam A, Sehic J, Nybom H. Molecular characterisation of indigenous Swedish apple cultivars based on SSR and S-allele analysis. Hereditas. 2008;**145**(3):99-112

[104] Pereira-Lorenzo S, Ramos-Cabrer AM, González-Díaz AJ, Díaz-Hernández MB. Genetic assessment of local apple cultivars from La Palma, Spain, using simple sequence repeats (SSRs). Scientia Horticulturae (Amsterdam). 2008 Jun;**117**(2):160-166. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0304423808001039 [cited 2014 Dec 12]

[105] Patocchi A, Fernández-Fernández F, Evans K, Gobbin D, Rezzonico F, Boudichevskaia A, et al. Development and test of 21 multiplex PCRs composed of SSRs spanning most of the apple genome. Tree Genetics & Genomes. 2008 Aug 29;**5**(1):211-223. Available from: http://link.springer.com/10.1007/s11295-008-0176-7 [cited 2014 Dec 12]

[106] van Treuren R, Kemp H, Ernsting G, Jongejans B, Houtman H, Visser L. Microsatellite genotyping of apple (Malus × domestica Borkh.) genetic resources in the Netherlands: Application in collection management and variety identification. Genetic Resources and Crop Evolution. 2010 Jan 28;**57**(6):853-865. Available from: http://link.springer.com/10.1007/s10722-009-9525-0; [cited 2014 Nov 30]

[107] Baric S, Wagner J, Storti A, Via JD. Application of an extended set of microsatellite DNA markers for the analysis of presumed synonym cultivars of apple. Acta Horticulturae. 2011;**918**:303-308

[108] Moriya S, Iwanami H, Okada K, Yamamoto T, Abe K. A practical method for apple cultivar identification and parent-offspring analysis using simple sequence repeat markers. Euphytica. 2010 Nov 5;**177**(1):135-150. Available from: http://link.springer.com/10.1007/s10681-010-0295-8 [cited 2014 Dec 12]

[109] Liu GS, Zhang YG, Tao R, Fang JG, Dai HY. Identification of apple cultivars on the basis of simple sequence repeat markers. Genetics and Molecular Research. 2014 Jan;**13**(3):7377-7387. Available from http://www.ncbi.nlm.nih.gov/pubmed/25222236

[110] Omasheva ME, Chekalin SV, Galiakparov NN. Evaluation of molecular genetic diversity of wild apple Malus sieversii populations from Zailiysky Alatau by microsatellite markers. Russian Journal of Genetics. 2015;**51**(7):647-652. Available from: http://link.springer.com/10.1134/S1022795415070108

[111] Lassois L, Denancé C, Ravon E, Guyader A, Guisnel R, Hibrand-Saint-Oyant L, et al. Genetic diversity, population structure, parentage analysis, and construction of core collections in the French apple germplasm based on SSR markers. Plant Molecular Biology Reporter. 2016;**34**(4):827-844. Available from: http://dx.doi.org/10.1007/s11105-015-0966-7

[112] Bigliazzi J, Scali M, Paolucci E, Cresti M, Vignani R. DNA extracted with optimized protocols can be genotyped to reconstruct the varietal composition of monovarietal wines. American Journal of Enology and Viticulture. 2012 Dec 1;**63**(4):568-573. Available from: http://www.ajevonline.org/cgi/doi/10.5344/ajev.2012.12014 [cited 2014 Dec 4]

[113] Boccacci P, Akkak A, Torello Marinoni D, Gerbi V, Schneider A. Genetic traceability of Asti Spumante and Moscato d'Asti musts and wines using nuclear and chloroplast microsatellite markers. European Food Research and Technology. 2012 Jul 7;**235**(3):439-446. Available from: http://link.springer.com/10.1007/s00217-012-1770-3 [cited 2014 Dec 12]

[114] Pereira L, Martins-Lopes P, Batista C, Zanol GC, Clímaco P, Brazão J, et al. Molecular markers for assessing must varietal origin. Food Analytical Methods. 2012 Feb 11;**5**(6):

1252-1259. Available from: http://link.springer.com/10.1007/s12161-012-9369-7 [cited 2014 Nov 30]

[115] Recupero M, Garino C, De Paolis A, Cereti E, Coïsson J-D, Travaglia F, et al. A method to check and discover Adulteration of nebbiolo-based monovarietal musts: Detection of Barbera and Dolcetto cv via SSR analysis coupled with Lab-On-Chip® microcapillary electrophoresis. Food Analytical Methods. 2012 Sep 16;**6**(3):952-962. Available from: http://link.springer.com/10.1007/s12161-012-9506-3 [cited 2014 Dec 12]

[116] Muccillo L, Gambuti A, Frusciante L, Iorizzo M, Moio L, Raieta K, et al. Biochemical features of native red wines and genetic diversity of the corresponding grape varieties from Campania region. Food Chemistry. 2014 Jan 15;**143**:506-513. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24054274 [cited 2014 Dec 12]

[117] Doveri S, O'Sullivan DM, Lee D. Non-concordance between genetic profiles of olive oil and fruit: A cautionary note to the use of DNA markers for provenance testing. Journal of Agricultural and Food Chemistry. 2006;**54**(24):9221-9226

[118] Pasqualone A, Montemurro C, Summo C, Sabetta W, Caponio F, Blanco A. Effectiveness of microsatellite DNA markers in checking the identity of protected designation of origin extra virgin olive oil. Journal of Agricultural and Food Chemistry. 2007;**55**(10):3857-3862

[119] Martins-Lopes P, Gomes S, Santos E, Guedes-Pinto H. DNA markers for Portuguese olive oil fingerprinting. Journal of Agricultural and Food Chemistry. 2008;**56**(24):11786-11791

[120] Alba V, Sabetta W, Blanco A, Pasqualone A, Montemurro C. Microsatellite markers to identify specific alleles in DNA extracted from monovarietal virgin olive oils. European Food Research and Technology. 2009 Apr 11;**229**(3):375-382. Available from: http://link.springer.com/10.1007/s00217-009-1062-8 [cited 2014 Dec 12]

[121] Rabiei Z, Tahmasebi Enferadi S, Saidi A, Patui S, Paolo Vannozzi G. Simple sequence repeats amplification: A tool to survey the genetic background of olive oils. Iranian Journal of Biotechnology. 2010;**8**:24-31

[122] Vietina M, Agrimonti C, Marmiroli M, Bonas U, Marmiroli N. Applicability of SSR markers to the traceability of monovarietal olive oils. Journal of the Science of Food and Agriculture. 2011 Jun;**91**(8):1381-1391. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21384371 [cited 2014 Dec 12]

[123] Rotondi A, Beghè D, Fabbri A, Ganino T. Olive oil traceability by means of chemical and sensory analyses: A comparison with SSR biomolecular profiles. Food Chemistry. 2011 Dec;**129**(4):1825-1831. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0308814611008259 [cited 2014 Dec 11]

[124] Ben-Ayed R, Grati-Kamoun N, Sans-Grout C, Moreau F, Rebai A. Characterization and authenticity of virgin olive oil (*Olea europaea* L.) cultivars by microsatellite markers. European Food Research and Technology. 2011 Nov 30;**234**(2):263-271. Available from: http://link.springer.com/10.1007/s00217-011-1631-5 [cited 2014 Dec 12]

[125] Pérez-Jiménez M, Besnard G, Dorado G, Hernandez P. Varietal tracing of virgin olive oils based on plastid DNA variation profiling. PLoS One. 2013 Jan;**8**(8):e70507. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3737381&tool=pmcentrez&rendertype=abstract; [cited 2014 Dec 12]

[126] Falagas ME, Pitsouni EI, Malietzis GA, Pappas G. Comparison of PubMed, Scopus, Web of Science, and Google Scholar: Strengths and weaknesses. The FASEB Journal. 2008;**22**(2): 338-342. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17884971

[127] Khan MA, Han Y, Zhao YF, Troggio M, Korban SS. A multi-population consensus genetic map reveals inconsistent marker order among maps likely attributed to structural variations in the apple genome. PLoS One. 2012 Jan;**7**(11):e47864. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3489900&tool=pmcentrez&rendertype=abstract; [cited 2014 Dec 12]

[128] Genome Database for Rosaceae [Internet]. [cited 2016 May 30]. Available from: http://www.rosaceae.org/search/markers

[129] Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A, et al. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature. 2007;**449**. Available from: http://dx.doi.org/10.1038/nature06148

[130] Paux E, Sourdille P, Salse J, Saintenac C, Choulet F, Leroy P, et al. A physical map of the 1Gb bread wheat chromosome 3B. Science (80-). 2008;**322**(2008):101-104

[131] Muleo R, Morgante M, Velasco R, Cavallini A, Perrotta G, Baldoni L. Olive Tree Genomic. In: Olive Germplasm - The Olive Cultivation, Table Olive and Olive Oil Industry in Italy. Rijeka, Croatia: InTech. 2012

[132] Sato S, Tabata S, Hirakawa H, Asamizu E, Shirasawa K, Isobe S, et al. The tomato genome sequence provides insights into fleshy fruit evolution. Nature. 2012;**485**(7400):635-641. Available from: http://dx.doi.org/10.1038/nature11119%5Cnhttp://www.nature.com/nature/journal/v485/n7400/full/nature11119.html?WT.ec_id=NATURE-20120531

[133] Cruz F, Julca I, Gómez-Garrido J, Loska D, Marcet-Houben M, Cano E, et al. Genome sequence of the olive tree, *Olea europaea*. Gigascience. 2016;**5**(1):29. Available from: http://gigascience.biomedcentral.com/articles/10.1186/s13742-016-0134-5

[134] Olea Database [Internet]. [cited 2016 May 30]. Available from: http://www.oleadb.it/ssr/ssr_dca_search.php

[135] Tomato Microsatellite Database [Internet]. [cited 2016 May 30]. Available from: http://webapp.cabgrid.res.in/tomsatdb/tutorial.html

[136] Tomato Genomic Resources Database [Internet]. [cited 2016 May 30]. Available from: http://59.163.192.91/tomato2/ssr.html

[137] Tomato-Kazusa Marker Database [Internet]. [cited 2016 May 30]. Available from: http://marker.kazusa.or.jp/tomato

[138] Grape Microsatellite Collection [Internet]. [cited 2016 May 30]. Available from: http://meteo.iasma.it/genetica/gmc.html

[139] This P, Dettweiler E. EU-Project Genres CT96 No81 : European Vitis Database and Results Regarding the Use of a Common Set of Microsatellite Markers. 2003.

[140] Italian Vitis Database [Internet]. [cited 2016 May 30]. Available from: http://www.vitisdb.it/descriptors/microsatellites

[141] European Vitis Database [Internet]. [cited 2016 May 30]. Available from: http://www.eu-vitis.de/index.php

[142] Bennett MD, Leitch IJ. Nuclear-DNA amounts in angiosperms. Annals of Botany. 1995;**76**:113-176

[143] Wheat Microsatellite Consortium [Internet]. [cited 2016 May 30]. Available from: http://wheat.pw.usda.gov/ggpages/SSR/WMC

[144] Li GY, Dreisigacker S, Warburton ML, Xia XC, He ZH, Sun Q. Development of a fingerprinting database and assembling an SSR reference kit for genetic diversity analysis of wheat. Acta Agronomica Sinica. 2006;**32**(12):1771-1778

[145] Han Y, Korban SS. An overview of the apple genome through BAC end sequence analysis. Plant Molecular Biology. 2008;**67**(6):581-588

[146] HiDRAS SSR Database [Internet]. [cited 2016 May 30]. Available from: http://www.hidras.unimi.it/HiDRAS-SSRdb/pages/index.php

[147] Lateur M, Ordidge M, Engels J, Lipman E. Report of a Working Group on Malus/Pyrus. Report of a Working Group on Malus/Pyrus. Fourth Meeting, 7-9 March 2012, Weggis, Switzerland. 2012

[148] Tanaka TS, Kobayashi FU, Joshi GIRIPR, Onuki RI, Sakai HI, Nasuda SH, et al. Next-Generation Survey Sequencing and the Molecular Organization of Wheat Chromosome 6B. DNA Research. 2013 Oct:103-114

[149] Varshney RK, Graner A, Sorrells ME. Genic microsatellite markers in plants: Features and applications. Trends in Biotechnology. 2005;**23**(1):48-55

[150] Hu J, Wang L, Li J. Comparison of genomic SSR and EST-SSR markers for estimating genetic diversity in cucumber. Biologia Plantarum. 2011;**55**(3):577-580

[151] Nunzia Scotti TC, Marechal-Drouard L. Mitochondrial DNA and RNA Isolation from Small Amounts of Potato Tissue. Plant Molecular Biology Reporter. 2001;**19** (March):67

[152] Wolfe KH, Li W-H, Sharp PM. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proceedings of the National Academy of Sciences of the United States of America. 1987;**84**(24):9054-9058. Available from: http://www.jstor.org/stable/30764

[153] Borgo R, Souty-Grosset C, Bouchon D, Gomot L. PCR-RFLP analysis of mitochondrial DNA for identification of snail meat species. Journal of Food Science. 1996;**61**(1):1-4

[154] Botstein D, White RL, Skolnick M, Davis RW. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. American Journal of Human Genetics. 1980;**32**(3):314-331 Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1686077&tool=pmcentrez&rendertype=abstract

[155] Nagy S, Poczai P, Cernák I, Gorji AM, Hegedus G, Taller J. PICcalc: An online program to calculate polymorphic information content for molecular genetic studies. Biochemical Genetics. 2012;**50**(9–10):670-672

[156] Anderson JA, Churchill GA, Autrique JE, Tanksley SD, Sorrells ME. Optimizing parental selection for genetic linkage maps. Genome Dynamics. 1993;**36**(i):181-186

[157] Aranzana MJ, Carbó J, Arús P. Microsatellite variability in peach (*Prunus persica* (L.) Batsch): Cultivar identification, marker mutation, pedigree inferences and population structure. Theoretical and Applied Genetics. 2003 May;**106**:8, 1341-1352. Available from http://www.ncbi.nlm.nih.gov/pubmed/12750778; [cited 2014 Dec 12]

[158] Qanbari S, Eskandari Nasab MP, Osfoori R, Nazari AH. Power of microsatellite markers for analysis of genetic variation and parentage verification in sheep. Pakistan Journal of Biological Sciences. 2007;**10**:1632-1638

[159] Silfverberg-Dilworth E, Matasci CL, Van De Weg WE, Van Kaauwen MPW, Walser M, Kodde LP, et al. Microsatellite markers spanning the apple (*Malus* x *domestica* Borkh.) genome. Tree Genetics & Genomes. 2006;**2**(4):202-224

[160] Urquhart A, Kimpton CP, Downes TJ, Gill P. Variation in short tandem repeat sequences–A survey of twelve microsatellite loci for use as forensic identification markers. International Journal of Legal Medicine. 1994;**107**:13-20

[161] Brookfield JF. A simple new method for estimating null allele frequency from heterozygote deficiency. Molecular Ecology. 1996;**5**(3):453-455

[162] Sehic J, Garkava-Gustavsson L, Nybom H. More harmonization needed for DNA-based identification of apple germplasm. Acta Horticulturae. 2013;(976):277-284

[163] Sefc KM, Lopes MS, Mendonça D, Rodrigues Dos Santos M, Laimer Da Camara Machado M, Machado ADC. Identification of microsatellite loci in olive (*Olea europaea* L.) and their characterization in Italian and Iberian olive trees. Molecular Ecology. 2000;**9**:1433-1449 Available from: http://www.ncbi.nlm.nih.gov/pubmed/10972783

[164] Carriero F, Fontanazza G, Cellini F, Giorio G. Identification of simple sequence repeats (SSRs) in olive (*Olea europaea* L.). Theoretical and Applied Genetics. 2002;**104**(2–3):301-307

[165] Cipriani G, Marrazzo MT, Marconi R, Cimato A, Testolin R. Microsatellite markers isolated in olive (*Olea europaea* L.) are suitable for individual fingerprinting and reveal polymorphism within ancient cultivars. Theoretical and Applied Genetics. 2002;**104**(2–3):223-228

[166] de la Rosa R, James CM TK. Isolation and characterization of polymorphic microsatellites in olive (*Olea europaea* L.) and their transferability to other genera in the Oleaceae. Molecular Ecology Notes. 2002;**2**:265-267

[167] Olive-Track. Traceability of Origin and Authenticity of Olive Oil by Combined Genomic and Metabolomic Approaches [Internet]. [cited 2016 May 30]. Available from: http://www.dsa.unipr.it/foodhealth/oliv-track/index.html

[168] Marmiroli N, Maestri E, Pafundo S, Vietina M, Biotechnology E. Molecular traceability of olive oil: From plant genomics to food genomics. Advances in Olive Resources. Kerala, India: Transworld Research Network. 2009

[169] Agrimonti C, Vietina M, Pafundo S, Marmiroli N. The use of food genomics to ensure the traceability of olive oil. Trends in Food Science and Technology. 2011 May;**22**(5):237-244. Available from: http://linkinghub.elsevier.com/retrieve/pii/S092422441100029X [cited 2014 Dec 11]

[170] Caroli S, Santoni S, Ronfort J. AMaCAID: A useful tool for accurate marker choice for accession identification and discrimination. Molecular Ecology Resources. 2011;**11**(4): 733-738

[171] Arias RS, Ballard LL, Scheffler BE. UPIC: Perl scripts to determine the number of SSR markers to run. Bioinformation. 2009;**3**(8):352-360. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19707300%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2720665

Chapter 9

Wild Soybeans: An Opportunistic Resource for Soybean Improvement

Muhammad Amjad Nawaz, Seung Hwan Yang and Gyuhwa Chung

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.74973

Abstract

Reduced genetic diversity in cultivated soybean coupled with changing dietary expectations, climate change, and increase in population demands expansion of current gene pool. Wild soybeans are an opportunistic resource and a rational choice to discover novel genes and gene families for alternative crop production systems and to improve soybean. Multiple agronomic traits, lineage-specific genes, and domestication-related traits have been studied in wild soybeans in contrast to cultivated soybeans, and it has been proved that wild soybeans are an essential genomic resource containing unique and useful genetic resources that have been lost during domestication to expand the gene pool in order to improve soybean. Wild soybean is very often a plant of disturbed habitats of Southeast Asia. The vulnerability of these habitats to agriculture systems and urban expansion causes a reduction in the area of distribution and hence the diversity. To capture the wild soybean genetic diversity in its main distribution areas, a unique and comprehensive germplasm collection, characterization, and conservation platform is direly needed. Chung's Wild Legume Germplasm Collection is preserving and maintaining a representative wild soybean germplasm collection guided by the principles of conservation genetics. These wild legumes and particularly wild soybean is a promising genetic resource for soybean breeders.

Keywords: CWLGC, genetic diversity, genetic resource, *Glycine soja*, rediscovery of landraces, wild legumes, germplasm conservation

1. Introduction: rediscovery of crop wild relatives

Challenged by limited land and water resources and a concomitant increase in population, changing dietary expectations and climatic change are demanding escalated food supplies

[1]. The nutritional value of grain legumes is far better than cereals even if their production is low, making them a unique and essential component of balanced diet [2]. Grain legumes have suffered a reduction in genetic diversity largely due to plant breeding activities aimed at artificial selection of desirable traits. The new varieties, as well as land races in farmer's field, have desirable characteristics which have become genetically diverged from their ancestors or wild progenitors [3].

In order to cope with the global warming led climatic variations and limited water supplies, there is a constant need of crop improvement; the crop potential has been reduced due to domestication, genetic bottlenecks, and artificial selection [2]. To explore more genes and gene families for alternative production systems, crop wild relatives are a rational choice mainly due to limited or no breeding barriers [4]. The wild progenitors of crops are sometimes easily available, but this is not the case for all species as some of the wild species have gone extinct, or in other cases, multiple progenitors contributed to the genome of the domesticated plants, e.g., wheat. In some cases, some species are indirectly expanding the genomes of the domesticated crops as they may be related species of wild progenitors or wild cousins [5]. Wild crop relatives are mostly adapted to larger climatic variations and are evolved to withstand biotic and abiotic stresses [6]. Therefore, for crop improvement, we have two possibilities, namely, genetic modification or introduction of genetic materials through breeding with crop wild relatives. Of course the use of genetic engineering to create genetically modified plants is relatively quick and efficient but the acceptance of genetically modified plants among the consumers is still controversial. On the other hand, the desirable traits including resistance to biotic and abiotic stresses, nutritional values can be incorporated into the current agricultural crop by using conventional and new breeding technologies. This practice is sometimes quite challenging mainly due to linkage drag; however, recent advances in genetics and genomic approaches have expanded our understanding of evolution, linkage, and heredity of complex traits [4–6] (**Table 1**).

2. Wild soybean

The *Glycine* genus comprises of two subgenera, namely, *Glycine* Willd and *Soja* (Moench) F. J. Hermann. Among 28 species classified under two subgenera, only two annual species *G. soja* Sieb & Zucc (wild) and *G. max* (cultivated) are consumed as food or feed either directly or indirectly [13]. On the basis of cytological, proteomic, and genomic evidence, the wild species *G. soja* is considered as the progenitor of cultivated species *G. max*. Both herbaceous annual species are Asian inborn; mainly distributed in Southeast and Far East Asia including China, Korea, Japan, Taiwan, and Russia [14]. Wild soybean harbors treasured genetic resource and extraordinarily important gene pool; genes and gene families responsible for higher oil and protein contents, resistance to drought and high temperatures, disease resistance, and insect pest resistance [15].

Wild soybeans grow on roadsides, riversides villages, lakeshores, wastelands, and fertile valleys. Apart from numerous phenotypic distinctions among both species, their annual growing habit with similar ploidy level and ability to produce fertile offspring without genetic isolation

Crop	Wild relatives	Reference
Rice	*O. glaberrima Steud.*	[7]
Oryza sativa L.	*O. nivara Sharma et Shastry*	
	O. rufipogon Griff.	
	O. breviligulata	
	O. barthii	
	O. meridionalis	
	O. glumaepatula	
Wheat	*T. boeoticum*	[8]
Triticum aestivum L.	*T. monococcum*	
	T. dicoccoides	
	T. araraticum	
	T. dicoccum	
	T. palaeocolchicum	
	T. timopheevii	
	T. durum	
	T. turgium	
	T. persicum (=*T. carthlicum* Nevski)	
	T. polonicum	
	T. turanicum (=*T. orientale* Perc.)	
	T. parvicoccum	
	T. spelta	
	T. macha	
	T. vavilovii	
	T. aestivum	
	T. compactum (= *T. aestivum* grex aestivo-compactum Schiem.)	
	T. sphaerococcum	
Corn	*Z. luxurians*	[9]
Zea mays L.	*Z. diploperennis*	
	Z. Mexicana	
	Zea nicaraguensis	
	Zea perennis	
Soybean	*G. soja* Sieb. & Zucc.	[10]
Glycine max L.	See **Table 2** for wild cousins of soybean	
Common bean	*P. coccineus*	[11]
Phaseolus vulgaris L.	*P. dumosus*	
	P. maculatifolius	
	P. costaricensis	
	P. acutifolius	
	P. lunatus	
	P. parvifolius	
	P. costaricensis	
	P. albescens	
	P. filiformis	

Crop	Wild relatives	Reference
Chickpea	*C. reticulatum*	[12]
Cicer arietinum L.	*C. echinospermum*	
	C. pinnatifidum	
	C. judaicum	
	C. bijugum	
	C. cuneatum	
	C. chorassanicum	
	C. yamashitae	

Table 1. List of wild relatives of three major cereals and three major legumes.

results in a flow of certain characteristics from wild to cultivated populations [16]. Wild soybeans exhibit distinct geographical patterns as well as interspecific horizontal mechanisms of flow of genetic information to cultivated species mainly because of sharing the same gene pool and close proximity [10] (**Table 2**).

Many parts of China and South Korea which were previously regarded as habitat for wild soybean are now being used for agricultural, commercial purposes (roads, buildings or dams) or are now part of the sea. Destruction of natural soybean habitats due to land clearance for agricultural or industrial purposes has led to decreased wild germplasm resources [19]. Furthermore, reduction in genetic diversity has been witnessed due to the domestication of soybean during past three decades. Progenitor wild species exhibit discrete geographic patterns with greater genetic diversity but the selection and allele frequency changes during domestication has curtailed genetic variability. Various studies have reported a reduction of genetic diversity up to 50% in domesticated/improved cultivars as compared to wild progenitors [3, 15]. Artificial selection and domestication mainly focused dominant selection of desirable traits such as oil content, seed size, and seed coat luster, imposing selection pressure on particular traits and ignoring other important traits. Another factor involved in diminishing genetic diversity is habitat fragmentation [20] (**Figure 1**).

2.1. Wild soybean genome: rediscovering the lost diversity

Whole genome sequencing of wild soybean genome started a new era for soybean functional and comparative genomics and has substantially increased our understanding about soybean domestication history, bottlenecks, lost diversity and has created a way forward towards its potential use in expanding the gene pool of soybean. The wild and cultivated soybeans have significant and useful genomic differences which highlight the phenotypic differences and as well as the domestication-related traits. Kim et al. [21] aligned 915.4 Mb genomic sequence of wild soybean with soybean reference genome excluding the gaps and found that wild soybean genome covered 97.65% of soybean genome with a difference of 35.2 Mb (3.76% of 937.5 Mb). The difference region consisted of 0.267% substitution bases, 0.043% insertion/deletions (indels), and 3.45% of large deleted sequences. Single nucleotide polymorphisms (SNPs) and

Genus	Subgenus	Species	Authority	Reference
Glycine	*Soja*	*G. max*	Merr.	[17, 18]
		G. soja	Sieb. & Zucc.	
	Glycine	*G. canescens*	F. J. Hermann	
		G. clandestine	Wendl.	
		G. latrobeana	Benth.	
		G. tabacina	Benth.	
		G. tomentella	Hayata	
		G. falcate	Benth.	
		G. latifolia	(Benth) Newell & Hymowitz	
		G. argyrea	Tindale	
		G. cyrtoloba	Tindale	
		G. curvata	Tindale	
		G. arenaria	Tindale	
		G. microphylla	(Benth.) Tindale	
		G. albicans	Tindale & Craven	
		G. hirticaulis	Tindale & Craven	
		G. lactovirens	Tindale & Craven	
		G. dolichocarpa	Tateishi & Ohashi	
		G. pindanica	Tindale & Craven	
		G. stenophita	B. Pfeil & Tindale	
		G. peratosa	B. E. Pfeil & Tindale	
		G. rubiginosa	Tindale & B. E. Pfeil	
		G. aphyonota	B. Pfeil	
		G. pullenii	B. Pfeil, Tindale & Craven	
		G. gracei	B. E. Pfeil & Craven	
		G. montis-douglas	B. E. Pfeil, Tindale & Craven	
		G. pescadrensis	Hayata	
		G. syndetika	B. E. Pfeil & Craven	

Table 2. Members of genus ***Glycine***.

insertions/deletions (indels) in precisely aligned areas differed by 0.31% between cultivated and wild soybean. The complex genome rearrangement is mainly caused by indels, inversions, and translocations (up to thousands of base pairs); along with SNPs and indels [22]. The wild soybean genome has greater allelic diversity than that of soybean. Resequencing of 17 wild and 14 cultivated soybean genomes to an average of ×5 depth and >90% coverage identified higher allelic diversity [15].

Based on whole-genome SNP analysis using the parameter $\theta\pi$, a higher level of genetic diversity was found in wild soybeans (2.97×10^{-3}) as compared to cultivated soybeans (1.89×10^{-3}).

Figure 1. Geographic distribution of wild soybeans.

Similar findings were also reported when 302 wild and cultivated soybeans were whole-genome sequenced to an average depth of > x11; the genetic diversity (π) decreased from 2.94×10^{-3} in *G. soja* to 1.40×10^{-3} in landraces and to 1.05×10^{-3} in improved cultivars suggesting that nearly half of the annotated resistance-related sequences were lost during the domestication [3]. Noticeably, total number of SNPs in wild (5,924,662) and cultivated soybeans (4,127,942) was comparable (35 and 5%, respectively) and the ratio of nonsynonymous SNP to synonymous SNPs was higher in cultivated (1.38) than wild soybeans (1.36). It was observed that the number of fixed loci were lower in the wild (463,409) as compared to cultivated soybeans (2,148,585). These findings suggest that there should be low-frequency alleles in domesticated soybean owing to domestication bottleneck. However, contrary to this, low-frequency alleles were abundant in wild soybean suggesting that the wild soybean habitat has reduced and the cultivated soybean population has expanded [15].

2.2. Expanding the gene pool for soybean improvement

Wild soybeans are a potential genetic resource for the improvement of cultivated soybean and aid greatly in exploring alternative production systems. Wild soybeans, as in case of another wild relative of cultivated crop species, contain higher genetic diversity as they had a long time opportunity to evolve and withstand under varied environmental conditions without inference by humans [4, 15]. Wild soybeans are interfertile with cultivated soybeans and represent an easily accessible or primary gene pool for soybean improvement [10]. However, the global climate change and increase in human population have developed a scenario of securing, conserving, characterizing, and using wild soybeans as a resource for soybean improvement. Loss of genetic diversity during the process of soybean domestication and presence of a domestication bottleneck, i.e., domestication syndrome has led towards changes in growth habits, loss of germination inhibition and mechanisms of seed dispersal [4]. This domestication has also enabled the crop plants to withstand and adapt to modern agriculture and farming system, which is very encouraging. However, loss of diversity in cultivated soybeans calls for revisiting natural diversity reservoirs, i.e., wild soybeans in search of potential genes/alleles

for higher yield. Multiples sequences that are unique to wild soybeans have been discovered but a report from Korea by Chung et al. [23] also witnessed gene loss events in wild soybeans. However, this discrepancy might be due to diversity of wild collections [4]. Multiple agronomic traits, lineage-specific genes, and domestication-related traits have been studied in wild soybeans in contrast to cultivated soybeans, and it has been proved that wild soybeans are an essential genomic resource containing unique and useful genetic resources that have been lost during domestication to expand the gene pool in order to improve soybean [3, 15, 24, 25]. One recent example is the salt-resistant gene GmCH1X identified in wild soybean. The salt-resistant gene originally did not have a Ty1/copia retrotransposon insertion into its exon 3 in wild soybean and controls 80% salt tolerance in wild soybean (W05) as compared to its counterpart C08 (cultivated soybean) which had retrotransposon insertion possibly due to recent round of whole genome duplication [24], strongly implying that wild soybeans' genetic diversity must be explored.

2.3. Traits from the wild

Recent development in high-throughput sequencing technologies is clearly promoting a revolution in the comparative genomic sequencing of major crops. A rapid growth in the number of sequenced genomes of crops and their wild relatives has established that wild species tend to have higher genetic diversities, making the wild relatives promising natural resources of novel genes/alleles for crop improvement. Many studies have provided the details on wild soybean specific genes/alleles controlling major abiotic and biotic stress tolerance-related traits. Contrastingly, cultivated soybeans also have unique genes/alleles which have been possibly lost during the evolution of wild soybeans. However, the results of each comparative study must be based on the genetic diversity present within the subject population [4, 21, 26].

2.3.1. Domestication-related traits

Identification of genes for domestication-related traits is an important task to maintain diversity in crops for improvement. Such a knowledge provides essential understanding of how and what genetic signatures have brought necessary changes in plant phenotype and physiology during the process of domestication. In soybean, the domestication-related traits are the increased size of inflorescence, grain yield, seed size, seed color, hilum color, pubescence form, apical dominance, stem determinacy, and plant height. Many of the domestication QTLs have been identified, such as twining habit (Ch. 02 and 18), hard seededness (Ch. 02 and 06), determinate habit (Ch. 17), maximum internode length (Ch. 06, 18 and 19), flowering time (Ch. 06 and 16), pod dehiscence (Ch. 16), seed weight (Ch. 17), stem determinacy (Ch. 17), oil content (Ch. 03, 11, 12, 13, 15, 17), flower color (Ch. 13), seed coat color (Ch. 08), pubescence form (Ch. 01, 12, 18, 19, 20), and plant height (Ch. 18) [3, 15].

2.3.2. Other traits

Many useful QTLs/genes have been obtained by characterizing wild soybeans or using genetic populations resulting from crosses between wild and cultivated soybeans. These

genes/QTLs have been characterized to understand the stress resistance mechanisms and various biochemical pathways related to plant development, yield, and local breeding traits. (1) Multiple genes/alleles responsible for flower color, i.e., pinkish-white and white flowers have been identified from wild soybeans. Flavonoid 3′5′-hydroxylase (F3′5′H) and dihydroflavonol-4-reductase (DFR) are responsible for anthocyanin production. Different loci control the anthocyanin content and decide the fate of flower color. Different loci, i.e., W1, W3, W4, w1-s1, w1-s2, w1-Ip, and w1-p2 have been reported to control white color and pinkish-white [27, 28]. (2) Other studies on seed antioxidant, phenolics, and flavonoid contents have identified GmMATE1, 2, 4 genes [29]. Astringent taste in soy products is caused by group A saponins. *Glyma15g39090* has been successfully characterized as sg-5 gene in natural wild soybean mutant line CWS5095. The gene oxygenates the C-21 position of soyasapogenol B or other intermediate which results in the production of saponin A [30, 31]. Another gene Sg-1 of this pathway has also been characterized from Korean wild soybean natural mutant which controls Ab series of saponins [32]. (3) Soybean cyst nematode is a global threat to soybean and host plant resistance is an ideal way of managing the damages. Wild soybean was used to discover SNPs and candidate genes significantly associated with soybean cyst nematode resistance, i.e., 10 SNPs and genes related to disease resistance-related proteins with leucine-rich region. Two genes, namely, a mitogen-activated protein kinase gene (Ch. 18) and a MYB transcription factor (Ch. 19) were found to be strong candidates [33]. (4) Many genes and transcription factors have also been identified from wild soybeans which play important role in drought stress tolerance [34]. (5) A salt-tolerant gene, i.e., GmCHX1 has been identified from wild soybean through whole-genome de novo sequencing approach. (6) A phosphatase 2C-1 (PP2C-1) allele has been reported from wild soybean to be involved in seed weight and seed size. Apart from these genes many QTLs related to (7) linolenic acid production, (8) yield, height and maturity, (9) soybean cyst nematode, (10) seed yield, seed weight, seed filling period, maturity, plant height, and lodging, (11) salt tolerance, (12) sclerotinia stem rot, (13) root traits, (14) oil and local breeding, (15) shoot fresh weight, (16) seed antioxidant, phenolics and flavonoids, and many other traits have been mapped and identified from wild soybeans or populations developed by crossing wild and cultivated soybeans [24, 29, 34–41]. Taking the advantage of higher genetic diversity and identified genes and QTLs for important traits from wild soybean will gear up the soybean yield improvement in changing climatic conditions and modern dietary demands. Joint ventures guided by the principles of plant breeding, genetics, genomics, and modern biotechnology are underway in many parts of the world to improve soybean in terms of resistance to biotic and abiotic stresses, adaptation to low water and higher temperature conditions, as well as intensive agricultural systems.

3. Wild soybean germplasm conservation

Wild soybean (the presumed progenitor of soybean) is very often a plant of disturbed habitats of Southeast Asia. Such habitats are mostly on the roadsides, intensive agricultural lands, and areas with higher human disturbances in terms of land use. The adaptation to these disturbed areas actually predisposes the wild soybeans to agricultural systems; this is one of the reasons for its domestication in East Asia [2, 19]. The vulnerability of these habitats to agriculture

systems and urban expansion causes reduction in area of distribution and hence the diversity. As discussed earlier, wild soybean is an efficient resource for identification and characterization of furnished and important genes for soybean improvement [10]. Economically, genebanks and wild genetic resources have been unambiguously reported to have led towards higher economic return by increasing soybean productivity [5]. Wild soybean germplasm preservation is underway in many countries mainly China, Korea, Japan, and the United States of America [10, 19]. Surely, the collections are growing by following the principles of conservation genetics; however, the complete representative collections are yet to be achieved as there remain many unexplored uninhabited natural habitats of wild soybeans which might carry many useful genes, alleles, or mutations. The undiscovered variations are greatly in demand by plant breeders to increase soybean production for an ever-increasing population [18]. Wild soybean germplasm should be collected mainly to (a) to understand the taxonomy and phylogenetic relationships, (b) to understand the biosystematics of certain yield-related pathways, (c) characterize and conserve germplasm, and (d) make it available for soybean breeders across the globe [10, 18–19]. Currently, there are many gene banks which are working on wild soybean germplasm conservation. In Southeast Asia, China holds the largest wild soybean germplasm collection of 6172 accessions under Chinese Crop Germplasm Information system [42], followed by Chung's Wild Legume Germplasm Collection (CWLGC) holding 6012 accessions and National Institute of Agrobiological Sciences Genebank in Japan which holds 1131 accessions [43]. Outside East Asia, the largest collection of wild soybean germplasm is USDA Soybean Germplasm Collection holding nearly 21,810 accessions belonging to 21 species of genus Glycine [44]. Almost 1179 accessions belong to 20 wild relatives of cultivated soybean including the wild soybean. N. I. Vavilov Institute of Plant Genetic Resources (VIR) holds ~350 accessions [45]. All of these germplasm collections are either focused on cultivated soybean or limited to a particular country/region. There is a dire need of germplasm center particularly focused on the collection, characterization, and dissemination of wild soybean accessions from the main distribution area of the species i.e., Southeast Asia. Out of the abovementioned germplasm collections, CWLGC is primarily focused on wild soybean germplasm collections from China, Korea, Japan, and Far East Russia near Chinese border.

3.1. Chung's wild legume germplasm collection

Chung's wild legume germplasm collection strives to develop a comprehensive conservation program resourcefully and efficiently to conserve and promote the genetic diversity within wild legumes with main focus on wild legumes. Guided with the principles of conservation genetics, CWLGC focuses on (a) direct collection, (b) acquirement, (c) conservation, (d) evaluation and characterization, and (e) documenting and distribution of wild soybean germplasm. CWLGC was established in 1983 by Professor Gyuhwa Chung at Department of Biotechnology, Chonnam National University, Yeosu campus, Republic of Korea. CWLGC holds 10,314 different legume genera. However, particular emphasis is on the germplasm collection, multiplication, evaluation, and utilization of *G. soja*, *Amphicarpaea edgeworthii*, *Vigna vexillata*, *Rhynchosia volubilis*, and *Phaseolus nipponensis*. The CWLGC efforts incredibly focused on wild soybeans, native to main centers of diversity, particularly East Asia. Wild soybean seeds are the main collection at CWLGC, and the seeds are considered as the currency of germplasm which is safeguarded for global food availability and security as well as

Figure 2. Wild soybean germplasm conservation at CWLGC.

Figure 3. A glimpse of diverse legume germplasm seed collection at CWLGC.

to preserve natural genetic diversity of wild soybeans and other legumes. This diverse collection of pinhead to large-sized seeds is important than ever as plant breeders, geneticists, conservationists, and biotechnologists need to cope with the changing climate, reduction in arable land area, occurrence of natural disasters, environmental degradation, and rising expectations in nutritional standards (**Figures 2** and **3**).

Conflict of interest

The authors declare that there exists no conflict of interest.

Author details

Muhammad Amjad Nawaz, Seung Hwan Yang and Gyuhwa Chung*

*Address all correspondence to: chung@chonnam.ac.kr

Department of Biotechnology, Chonnam National University, Yeosu, Republic of Korea

References

[1] Considine MJ, Siddique KHM, Foyer CH. Nature's pulse power: legumes, food security and climate change. Journal of Experimental Botany. 2017;**68**:1815-1818. DOI: 10.1093/jxb/erx099

[2] Foyer CH, Lam HM, Nguyen HT, Siddique KHM, Varshney RK, Colmer TD, Cowling W, Bramley H, Mori TA, Hodgson JM, Cooper JW, Miller AJ, Kunert K, Vorster J, Cullis C, Ozga JA, Qahlqvist ML, Liang Y, Shou H, Shi K, Yu J, Fodor N, Kaiser BN, Wong FL, Valliyodan B, Considine MJ. Neglecting legumes has compromised human health and sustainable food production. Nature Plants. 2016;**2**:16112. DOI: 10.1038 / nplants.2016.112

[3] Zhou Z, Jiang Y, Wang Z, Gou Z, Lyu J, Li W, Yu Y, Shu L, Zhao Y, Ma Y, Fang C, Shen Y, Liu T, Li C, Li Q, Wu M, Wang M, Wu Y, Dong Y, Wan W, Wang X, Ding Z, Gao Y, Xiang H, Zhu B, Lee SH, Wang W, Tian X. Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean. Nature Biotechnology. 2015;**33**:408-415. DOI: 10.1038/nbt.3096

[4] Munoz N, Liu A, Kan L, Li MW, Lam HM. Potential uses of wild germplasms of grain legumes for crop improvement. International Journal of Molecular Sciences. 2017;**18**:328. DOI: 10.3390/ijms18020328

[5] Brozynska M, Furtado A, Henry RJ. Genomics of crop wild relatives: Expanding the gene pool for crop improvement. Plant Biotechnology Journal. 2015;**14**:1070-1085

[6] Zhang H, Mittal N, Leamy LJ, Barazani O, Song BH. Back into the wild—Apply untapped genetic diversity of wild relatives for crop improvement. Evolutionary Applications. 2017;**10**:5-24. DOI: 10.1111/eva.12434

[7] Sanchez PL, Wing RA, Brar DS. The wild relatives of rice: Genomes and genomics. In: Genetics and Genomics of Rice, Plant Genetics and Genomics: Crop and Models. New York: Springer Science + Business Media; 2014. pp. 9-25. DOI: 10.1007/978-1-4614-7903-1_2

[8] Arrigo N, Guadagnuolo R, Lappe S, Pasche S, Parisod C, Felber F. Gene flow between wheat and wild relatives: empirical evidence from *Aegilops geniculate*, *Ae. neglecta* and *Ae. truncialis*. Evolutionary Application. 2011;**4**:685-695. DOI: 10.1111/j.1752-4571.2011.00191.x

[9] Maazou ARS, Qiu J, Mu J, Liu Z. Utilization of wild relatives for maize (*Zea mays* L.) improvement. African Journal of Plant Science. 2017;**11**:105-113. DOI: 10.5897/AJPS2017.1521

[10] Chung G, Singh RJ. Broadening the genetic base of soybean: A multidisciplinary approach. Critical Reviews in Plant Sciences. 2008;(5):295-341. DOI: 10.1080/07352680802333904

[11] Porch TG, Beaver JS, Debouck DG, Jackson SA, Kelly JD, Dempewolf H. Use of wild relatives and closely related species to adapt common bean to climate change. Agronomy. 2013;**3**:433-461. DOI: 10.3390/agronomy3020433

[12] Sing R, Sharma P, Varshney RK, Sharma SK, Sing SK. Chickpea improvement: Role of wild species and genetic markers. Biotechnology and Genetic Engineering Reviews. 2008;**25**:267-314. DOI: 10.5661/bger-25-267

[13] Guo J, Wang Y, Song C, Zhou J, Qiu L, Huang H, Wang Y. A single origin and moderate bottleneck during domestication of soybean (*Glycine max*): Implications from microsatellites and nucleotide sequences. Annals of Botany. 2010;**106**:505-514. DOI: 10.1093/aob/mcq125

[14] Phylogeographcial studies of *Glycine soja*: Implicating the refugium during the quaternary glacial period and large-scale expansion after the Last Glacial Maximum. Turkish Journal of Agriculture and Forestry. 2016;**40**:825-838. DOI: 10.3906/tar-1607-30

[15] Lam HM, Xu X, Liu X, Chen W, Yang G, Wong FL, Li MW, He W, Qin N, Wang B, Li J, Jian M, Wang J, Shao G, Wang J, SSM S, Zhang G. Resequencing of 31 wild and cultivated soybean genomes identifies patters of genetic diversity and selection. Nature Genetics. 2015;**42**:1053-1059. DOI: 10.1038/ng.715

[16] Ohara M, Shimamoto Y. Importance of genetic characterization and conservation of plant genetic resources: The breeding system and genetic diversity of wild soybean (*Glycine soja*). Plant Species Biology. 2002;**17**:51-58. DOI: 10.1046/j.1442-1984.2002.00073.x

[17] Sherman-Broyles S, Bombarely A, Powell AF, Doyle JL, Eqan AN, Coate JE, Doyle JJ. The wild side of a major crop: Soybean's perennial cousins from down under. American Journal of Botany. 2014;**101**:1651-1665. DOI: 10.3732/ajb.1400121

[18] Singh RJ, Nelson RL, Chung G. Soybean [*Glycine max* (L.) Merr.]. In: Singh RJ, editor. Genetic Resources, Chromosome Engineering, and Crop Improvement. Vol. 4. Boca Raton, Florida: CRC Press, (Taylor & Frances Group); 2007. pp. 13-50

[19] Nawaz MA, Yang SH, Rehman HM, Baloch FS, Lee JD, Park JH, Chung G. Genetic diversity and population structure of Korean wild soybean (*Glycine soja* Sieb. and Zucc.) inferred from microsatellite markers. Biochemical Systematics and Ecology. 2017;**71**:87-96. DOI: 10.1016/j.bse.2017.02.002

[20] Van K, Kim MY, Shin JH, Kim KD, Lee YH, Lee SH. Molecular evidence for soybean domestication. In: Tuberosa R, Graner A, Frison E, editors. Genomics of Plant Genetic Resources. Dordrecht: Springer; 2014. pp. 465-481

[21] Kim MY, Lee S, Van K, Kim TH, Jeong SC, Choi IY, Kim DS, Lee YS, Park D, Ma J, Kim WY, Kim BC, Park S, Lee KA, Kim DH, Kim KH, Shin JH, Jang YE, Kim KD, Liu WX,

Chaisan T, Kang YJ, Lee YH, Kim KH, Moon JK, Schmutz J, Jackson SA, Bhak J, Lee SH. Whole-genome sequencing and intensive analysis of the undomesticated soybean (*Glycine soja* Sieb. and Zucc.) genome. PNAS. 2010;**107**:22032-22037. DOI: 10.1073/pnas.1009526107

[22] Kim MY, Van K, Kang YJ, Kim KH, Lee SH. Tracing soybean domestication history: From nucleotide to genome. Breeding Science. 2012;**61**:445-452. DOI: 10.1270/jsbbs.61.445

[23] Chung WH, Jeong N, Kim H, Lee WK, Lee YG, Lee SH, Yoon W, Kim JH, Choi IY, Choi HK, Moon JK, Kim N, Jeong SC. Population structure and domestication revealed by high-depth resequencing of Korean cultivated and wild soybean genomes. DNA Research. 2014;**21**:153-167. DOI: 10.1093/dnares/dst047

[24] Qi X, Li MW, Xie M, Liu X, Ni M, Shao G, Song C, Yim AKY, Tao Y, Wong FL, Isobe S, Wong CF, Wong KS, Xu C, Li C, Wang Y, Guan R, Sun F, Fan G, Ziao Z, Zhou F, Phang TH, Liu X, Tong SW, Chang TF, Yiu SM, Tabata S, Wang J, Xu X, Lam HM. Identification of a novel salt tolerance gene in wild soybean by whole-genome sequencing. Nature Communications. 2014;**5**:4340. DOI: 10.1038/ncomms5340

[25] Li YH, Zhou G, Ma J, Jiang W, Jin LG, Zhang Z, Guo Y, Zhang J, Sui Y, Zheng L, Zhang SS, Zuo Q, Shi XH, Li YF, Zhang WK, Hu Y, Kong G, Hong HL, Tan B, Song J, Liu ZX, Wang Y, Ruan H, Yeung CKL, Liu J, Wang H, Zhang LJ, Guan RX, Wang KJ, Li WB, Chen SY, Chang RZ, Jiang Z, Jackson SA, Li R, Qiu LJ. De novo assembly of soybean wild relatives for pan-genome analysis of diversity and agronomic traits. Nature Biotechnology. 2014;**32**:1045-1052. DOI: 10.1038/nbt.2979

[26] Stupar RM. Into the wild: The soybean genome meets its undomesticated relative. PNAS. 2010;**107**:21947-21948. DOI: 10.1073/pnas.1016809108

[27] Sundaramoorthy J, Park GT, Chang JH, Lee JD, Kim JH, Seo HS, Chung G, Song JT. Identification and molecular analysis of four new alleles at the W1 locus associated with flower colour in soybean. PLoS One. 2016;**11**:e0159865. DOI: 10.1371/journal.pone.0159865

[28] Kim JH, Bae DN, Park SK, Jeong N, Lee K, Kang H, Kang ST, Moon JK, Park E, Jeong SC. Molecular genetic analysis of a novel recessive white flower gene in wild soybean. Crop Science. 2017;**57**:3027-3034. DOI: 10.2135/cropsci2017.03.0163

[29] Li MW, Munoz NB, Wong CF, Wong FL, Wong KS, Wong JWH, Qi X, Li KP, Ng MS, Lam HM. QTLs regulating the contents of antioxidants, phenolics and flavonoids in soybean seeds share a common genomic region. Frontiers in Plant Sciences. 2016;**7**:854. DOI: 10.3389/fpls.2016.00854

[30] Yano R, Takagi K, Takada Y, Mukaiyama K, Tsukamoto C, Sayama T, Kaga A, Anai T, Sawai S, Ohyama K, Saito K, Ishimoto M. Metabolic switching of astringent and beneficial triterpenoid saponins in soybean is achieved by a loss of function mutation in cytochrome P450 72A69. The Plant Journal. 2017;**89**:527-539

[31] Rehman HM, Nawaz MA, Shah ZH, Yang SH, Chung G. Functional characterization of naturally occurring wild soybean mutant (sg-5) lacking astringent saponins using whole genome sequencing approach. Plant Science. 2017; in press. DOI: 10.1016/j.plantsci.2017.11.014

[32] Park J, Kim JH, Krishnamurthy P, Tsukamoto C, Song JT, Chung G, Shannon G, Lee JD. Characterization of new allele of the saponin-synthesizing gene Sg-1 in soybean. Crop Science. 2016;**56**:1-7. DOI: 10.2135/cropsci2015.04.0250

[33] Luo X, Bai X, Sun X, Zhu D, Liu B, Ji W, Cai H, Cao L, Wu J, Hu M, Liu X, Tang L, Zhu Y. Expression of wild soybean WRKY20 in Arabidopsis enhances drought tolerance and regulates ABA signalling. Journal of Experimental Botany. 2013;**64**:2155-2169. DOI: 10.1093/jxb/ert073

[34] Li D, Pfeiffer TW, Cornelius PL. Soybean QTL for yield and yield components associated with *Glycine soja* alleles. Crop Science. 2007;**48**:571-581. DOI: 10.2135/cropsci2007.06.0361

[35] Pantalone VR, Rebetzke GJ, Burton JW, Wilson RF. Genetic regulation of linolenic acid concentration in wild soybean *Glycine soja* accessions. Journal of the American Oil Chemists' Society. 1997;**74**:159-163

[36] Zhang WK, Wang YJ, Luo GZ, Zhang JS, He CY, Wu XL, Gai JY, Chen SY. QTL mapping of ten agronomic traits on the soybean (*Glycine max* L. Merr.) genetic map and their association with EST markers. Theoretical and Applied Genetics. 2004;**108**:1131-1139

[37] Iquira E, Humira S, François B. Association mapping of QTLs for sclerotinia stem rot resistance in a collection of soybean plant introductions using a genotyping by sequencing (GBS) approach. BMC Plant Biology. 2015;**15**:5. DOI: 10.1186/s12870-014-0408-y

[38] Zhang H, Li C, Davis EL, Wang J, Griffin JD, Kofsky J, Song BH. Genome-wide association study of resistance to soybean cyst nematode (*Heterodera glycines*) HG Type 2.5.7 in wild soybean (*Glycine soja*). Frontiers in Plant Science. 2016;**17**:1214. DOI: 10.3389/fpls.2016.01214

[39] Manavalan LP, Prince SJ, Musket TA, Chaky J, Deshmukh R, Vuong TD, Song L, Cregan PB, Nelson JC, Shannon JG, Specht JE, Nguyen HT. Identification of novel QTL governing root architectural traits in an interspecific soybean population. PLoS One. 2015;**10**:e0120490. DOI: 10.1371/journal.pone.0120490

[40] Ha BK, Vuong TD, Velusamy V, Nguyen HT, Shannon JG, Lee JD. Genetic mapping of quantitative trait loci conditioning salt tolerance in wild soybean (*Glycine soja*) PI 483463. Euphytica. 2013;**193**:79-88

[41] Asekova S, Kulkarni KP, Patil G, Kim M, Song JT, Nguyen HT, Shannon JG, Lee JD. Genetic analysis of shoot fresh weight in a cross of wild (*G. soja*) and cultivated (*G. max*) soybean. Molecular Breeding. 2016;**36**:103

[42] Chinese Crop Germplasm Information system. http://www.cgris.net/cgris_english.html [Accessed: 2017-11-20]

[43] National Institute of Agrobiological Sciences Genebank. http://www.naro.affrc.go.jp/archive/nias/index_e.html. [Accessed: 2017-11-20]

[44] United States Department of Agriculture. Soybean Germplasm Collection. http://grscicoll.org/institutional-collection/usda-soybean-germplasm-collection. [Accessed: 2017-11-20]

[45] N. I. Vavilov Institute of Plant Genetic Resources. http://www.vir.nw.ru/. [Accessed: 2017-11-20]

Chapter 10

Genetic Variation of Landraces of Common Bean Varying for Seed Coat Glossiness and Disease Resistance: Valuable Resources for Conservation and Breeding

Enéas Ricardo Konzen and Siu Mui Tsai

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.73425

Abstract

In this chapter, we outline the significance of landraces of common bean (*Phaseolus vulgaris* L.) for unraveling novel morphological, biochemical and genetic variation that could be integrated to breeding programs, related to seed coat color and glossiness and disease resistance. Moreover, we emphasize how important the conservation of such genetic resources is in small-farming areas, the prevailing system for bean cultivation. A particular Brazilian landrace referred as Serro Azul by local farmers is highlighted to show new evidences of the genetic control of seed glossiness in common bean and how it implicates in the seed protection against diseases and insects. Moreover, new findings presented here give insights into a remarkable anthracnose resistance of one of the variants of Serro Azul, which also presents seed coat glossiness. The potential benefits for human health after consuming beans with glossy seed coat are also discussed. This is one among the various landraces that need better understanding for strengthening the knowledge of the genetic diversity of common bean. Such knowledge is important for conducting conservation actions and performing new crosses for providing genetic materials with desirable combinations for farmers, breeders and consumers.

Keywords: Serro Azul, anthracnose resistance, asper gene, proanthocyanidins, crop genetic resources

1. Introduction

The conservation of crop genetic resources is a fundamental step for further breeding of traits of interest. Common bean (*Phaseolus vulgaris* L.) is the major legume for human consumption

throughout the world [1, 2]. It is naturally distributed from northern Mexico to northern Argentina, with a marked genetic structure. From the classic to the most recent reports, two major gene pools have been recognized for the species, the Mesoamerican and the Andean [3, 4]. Domestication has been independently performed within each gene pool [5, 6], selecting specific genes for growth habit, seed size, color and yield, a phenomenon described as the "domestication syndrome" [7]. Therefore, an enormous panel of diversity can be observed from wild to domesticated genotypes of common bean, making it as a model for understanding crop evolution [8].

The introduction of common bean to other areas than its natural habitat, such as in Brazil [9], led to new combinations of seed and flower colors, shapes and sizes, growth habits, cycle, photoperiod and yield [10]. It has also shaped the interaction between beans and the environment, leading to new diseases and pests.

In Brazil, common bean is a staple food among most citizens along with rice. Usually, its cultivation is performed in small farming systems along with other crops and animal production systems, providing self-sufficiency for farmers in various regions of the country. Several cycles of selection by local farmers in specific environments have generated new landraces, which are not yet known and available from core collections such as the ones from Centro Internacional de Agricultura Tropical (CIAT, Colombia), United States Department of Agriculture (USDA, USA) and Empresa Nacional de Pesquisa Agropecuária (EMBRAPA, Brazil). Landraces have singular aspects that might assist breeding programs for disease resistance, abiotic stress tolerance, improvement of nutrition facts, among several other desirable aspects of common bean grains.

A particular landrace has been discovered with local farmers from the municipality of Cunha, Sao Paulo state, Southeast Brazil. The farmers referred this landrace as "Serro Azul" and have been cultivating it along with other varieties [11]. Serro Azul shows considerable morphological variability, revealing different types of seeds. It shows high genetic variation for seed colors and patterns [12], disease and insect resistance and nodulation ability [11, 13], which are among the main aspect studies in breeding programs of common bean.

We drive the topic of this chapter highlighting Serro Azul as a case study of how the conservation of landraces might be important for further breeding strategies. First, we describe the importance of landraces of common bean in the discovery of new allele combinations for seed color and pattern genes, using the example of Serro Azul. Then, we briefly discuss the implications of the research of common bean landraces for nutrition and health. Moreover, we outline a significant number of original findings about our landrace in focus, which serve as examples of disease resistance as well as indications of insect resistance. Finally, we guide the reader through other perspectives of the importance of a better knowledge of landraces, their collection and conservation for future endeavors in breeding programs.

2. The Serro Azul landrace: parental genotypes and segregating populations

Serro Azul is a landrace that was so named by local farmers living in the countryside of the small municipality of Cunha, in São Paulo state, Brazil. From the search along the distinct

farms at that area, De Oliveira et al. [11] were able to collect seeds with three main patterns of colors such as gray background with black strips (Serro Azul Malhado—SAM), light brown with glossy seed coat (Serro Azul Brilhante—SAB) and dark gray with dull seed coat (Serro Azul Fosco—SAF). Interestingly, at one of the farms, from a fresh sample of SAF seeds collected (approximately 1 kg), only 10 seeds presented the SAB pattern (**Figure 1A** shows the phenotypes of SAB and SAF). Seeds with the SAB pattern were crossed to the SAF pattern and the reciprocal as well, and populations were advanced to F_4 generation, revealing consistent patterns of segregation for seed glossiness and color [12].

Serro Azul has been the object of a few studies, considering its importance to local farmers and the need of improving its productivity. It has been subjected to many diseases, such as anthracnose, and pests. De Oliveira et al. [11] and De Oliveira and Tsai [13] showed low yield from this variety from experiments performed at different farms in Cunha, SP, when no fertilizer was applied. This treatment represented the traditional way that farmers cultivated this landrace at their farms. However, when a sort of fertilizer treatments were applied in soil cultivated with this variety, along with one commercial standard at the time (IAC-Carioca 80SH), the yield was significantly improved from around 930 kg ha^{-1} (without fertilizer) to 1360 kg ha^{-1} (fertilizer application along with lime and foliar spray of molybdenum) [13].

3. Unraveling the genetic control of seed glossiness and its importance taking Serro Azul as an example

3.1. Genetic control of seed glossiness

The genetic control of seed coat color, pattern and glossiness in common bean has been a major issue for scientists in decades. Frequently, the gene nomenclature for loci related to such traits was confusing, leading to a series of meetings for establishing standard nomenclatures (Bean Improvement Cooperative meetings). As regards seed glossiness, Bassett [14] has clarified the differences between glossy and opaque seed coats through a series of genetic crosses, providing genetic stocks (pure lines for specific alleles of seed coat loci) that have been used since then for unraveling the genetic and biochemical aspects of the trait.

Landraces might be the source of additional alleles or allele combinations for studying genes related to color, pattern and glossiness of the seed coat. Moreover, these new sources might be interesting for being added to the breeding programs concerned with such traits. The study of Konzen and Tsai [12] examined the particular aspect of Serro Azul of segregating for seed glossiness and color patterning. The population developed from the cross SAF × SAB was analyzed from the parental genotypes to the F_4 generation for the segregation of these traits. The variation of glossiness was attributed to two alleles at the *Asp* locus, usually referred as the gene controlling glossiness in common bean [14]. SAB carries the dominant allele *Asp*, which confers the intense glossy aspect of the seed coat, while SAF has a dull seed coat [12] (**Figure 1A**). This gene is known to be located in chromosome 7 [15], where genes associated with anthracnose and angular leaf spot resistance [16, 17], common bacterial blight [17] and nodulation [18] are also located.

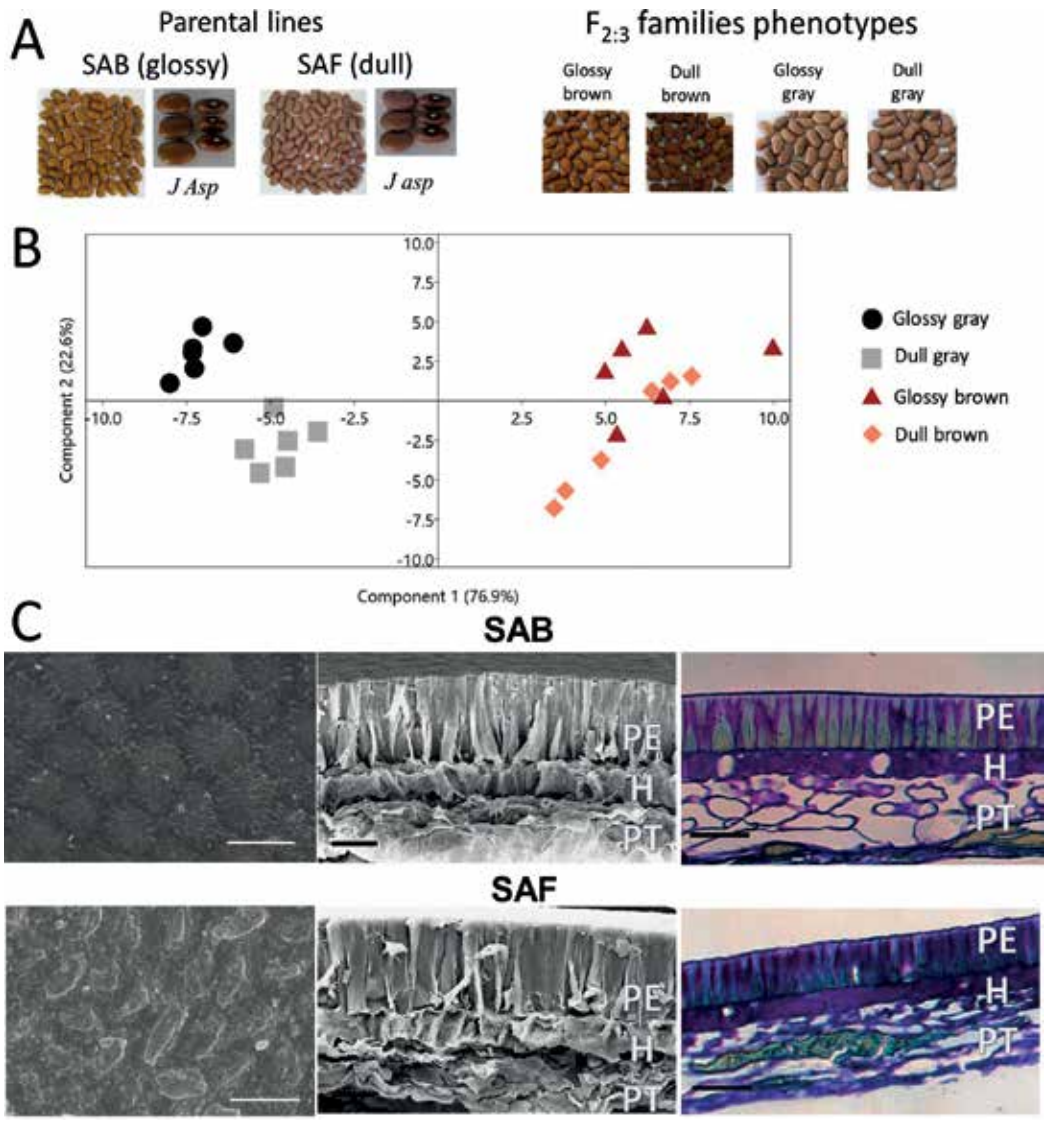

Figure 1. Morphological and anatomical aspects of seeds of the common bean landrace Serro Azul. (A) Parental genotypes (Serro Azul Brilhante—SAB and Serro Azul Fosco—SAF) and the four phenotypes of $F_{2:3}$ families of the cross SAF × SAB. The parent SAB has a glossy brown seed coat, while the SAF seed coat shows a dull gray phenotype. The $F_{2:3}$ families segregate for both color and glossiness. (B) Principal component analysis with the L*a*b* variables, obtained with a colorimeter for $F_{2:3}$ families of the cross SAF × SAB. a*: the amount of green or red; b*: the quantity of blue or yellow; L*: the quantity of brightness. (C) Scanning electron profiles of the seed coat surface (left image) and their transversal profile (central image) and histological sections (image on the right) of transversal sections of the seed coat of both SAB and SAF. PE: palisade epidermis; H: hypodermis; PT: parenchymatous tissue. The scale bars indicate 10 μm (left image), 20 μm (central image) and 100 μm (right image) (adapted from [12]).

However, the glossiness of the seed coat in Serro Azul is also related to the expression another gene located in chromosome 10 and referred as the *Joker* (*J*) locus [19]. The dominant allele *J* is responsible for an even distribution of the color shown by each seed, while *jj* genotypes exhibit irregular coloring, especially around the hilum ring [14]. Another peculiar aspect

related to the *J* allele is that it intensifies the glossy aspect of the seed, as seen in both SAB and SAF (**Figure 1A**). Therefore, SAF presents a slight shiny aspect on the seed coat due to the expression of *J* [12].

The segregation of the seed coat glossiness along with the color pattern is clearly observed in $F_{2:3}$ lines (**Figure 1A**) of the cross SAF × SAB and is only due to the *Asp* locus. *J* only contributed to intensify the shiny aspect and is present as the dominant allele in both SAB and SAF. The segregation of glossiness is according to the expected Mendelian proportions (3,1) in F_2 and F_3 generations of the cross between SAF (*asp asp*) and SAB (*Asp Asp*), being attributed to *Asp*. Measurements performed with a colorimeter allowed to validate the categories established (glossy and dull) (**Figure 1B**), based on the L*a*b* color system, in which L* is the main variable associated with glossiness, while a* and b* are measures of distinct colors [12]. In this study, in general, glossy seeds presented higher L* values than seeds with dull seed coat. However, there was some extent of interaction between the color and the brightness, especially with brown-colored seeds (**Figure 1B**).

Another important aspect of the seed coat glossiness of SAB was shown from microscopy analyses. It followed and confirmed previous findings of classic studies of the glossiness of other common bean genotypes. In general, dull genotypes (*asp asp*) have a rough textured surface of the seed coat, while glossy seed coats have an even surface, as shown from scanning electron micrographs [20] (**Figure 1C**, left image). Moreover, glossy seeds show a thicker palisade epidermis from the seed coat than the dull seed coats (**Figure 1C**, central and right figures) [12].

3.2. How important might seed glossiness be?

The seed glossiness has been frequently neglected in breeding programs due to consumer preferences. This is explained by the fact that glossy seeds tend to require higher cooking times than seeds with an opaque seed coat [21]. At first, it seems that since glossiness retards water absorption by the seeds [12, 22] (check **Figure 2**), they take longer to be cooked. However, some line of evidence showed no significant correlation between the cooking time

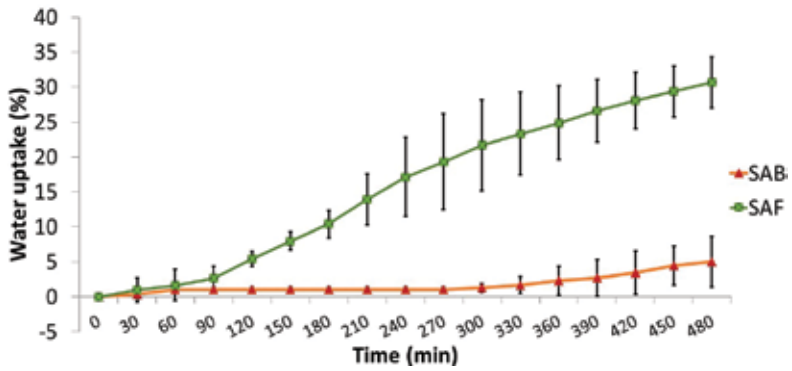

Figure 2. Water uptake on the course of 480 min of the variants Serro Azul Brilhante (SAB, with glossy seed coat) and Serro Azul Fosco (SAF, with dull seed coat). Three replicates of seeds were embedded in distilled water and paper dried every 30 min for weighing and determining the weight change (adapted from [12]).

and the water absorption rate [23]. Further examination of the genes involved in cooking time is necessary, though.

It is well known that the seed coat is the structure that protects the seeds from pathogens and insects, and the glossiness seems to have an important role in such protection. Moreover, seeds with glossiness might have enhanced antioxidant properties due to a higher concentration of specific secondary metabolites in the seed coat, therefore, having an impact in human health [12].

Usually, in the case of landraces, where local selection has been performed, it is more frequent to find common bean accessions that show glossy seed coat [24] (checking the list of genotypes) than in breeding programs. In the case of the landrace Serro Azul, both variants Serro Azul Brilhante (glossy) and Serro Azul Fosco (opaque) have been cultivated [11, 13]. Morphological and biochemical findings are hereafter discussed to show advantages of the seed glossiness for aspects related to human health.

3.3. Biochemical nature of seed coat glossiness and its implications for human health

The seed coat glossiness has been studied to be mainly conditioned by the *Asp* gene but also influenced by the *J* locus, especially with the dominant allele [14]. A number of studies have also been conducted to better understand the biochemical implications of the expression of such genes on the seed coat.

Classical work has suggested that *J* is essential for the synthesis of proanthocyanidins or condensed tannins [14]. Therefore, the recessive *jj* genotypes have been found to be absent in condensed tannins [25], contrary to the *J_* genotypes, that are able to synthesize such compounds. Based on the genetic maps that identified the RAPD marker as linked to *J*, a recent study has shown that *J* is linked to a region containing *MYB123* [26], similar to TT2 in *Arabidopsis thaliana* (AT5G35550.1) [27] and *Glycine max* [26], which acts as a key determinant in the proanthocyanidin accumulation of a developing seed [27].

Proanthocyanidins are oligomers or polymers formed by the condensations of flavan-3-ols units such as catechins and epicatechins [28, 29]. In common bean, condensed tannins are mainly composed of catechin monomers [30]. As secondary metabolites, they play important roles as antioxidants, anticarcinogenic and anti-inflammatory [28, 31, 32].

On the other hand, the *Asp* locus is said to be the main gene involved in seed glossiness. Some line of evidence has shown that *Asp* affects the accumulation of anthocyanins due to a structural change that it promotes on the seed coat. Therefore, genotypes with glossy seed coat (*Asp_*) accumulate more anthocyanins than dull seed coats (*asp asp*) [20]. Anthocyanins have been investigated for their roles in humans such as in anti-inflammatory, lipid peroxidation and membrane strengthening processes [33, 34], as well as in preventing cancer [35].

Therefore, although glossiness has been generally neglected by consumers and, as a result, by selection programs, it might have positive implication for human health. Moreover, the indication that glossiness is not necessarily associated with higher cooking time (since there is a lack of correlation between water absorption and cooking time) as shown by Garcia et al.

[23] needs to be further explored by the researchers. Landraces such as Serro Azul are one of the sources for rescuing the value of seed glossiness.

4. Serro Azul as a source of disease resistance

One of the most important aspects of a breeding program is to find genotypes that are tolerant or even resistant to diseases. The cultivation of common bean is majorly affected by diseases such as common bacterial blight caused by *Xanthomonas axonopodis* pv. *phaseoli*, the angular leaf spot caused by the fungus *Pseudocercospora griseola* and anthracnose by the fungus

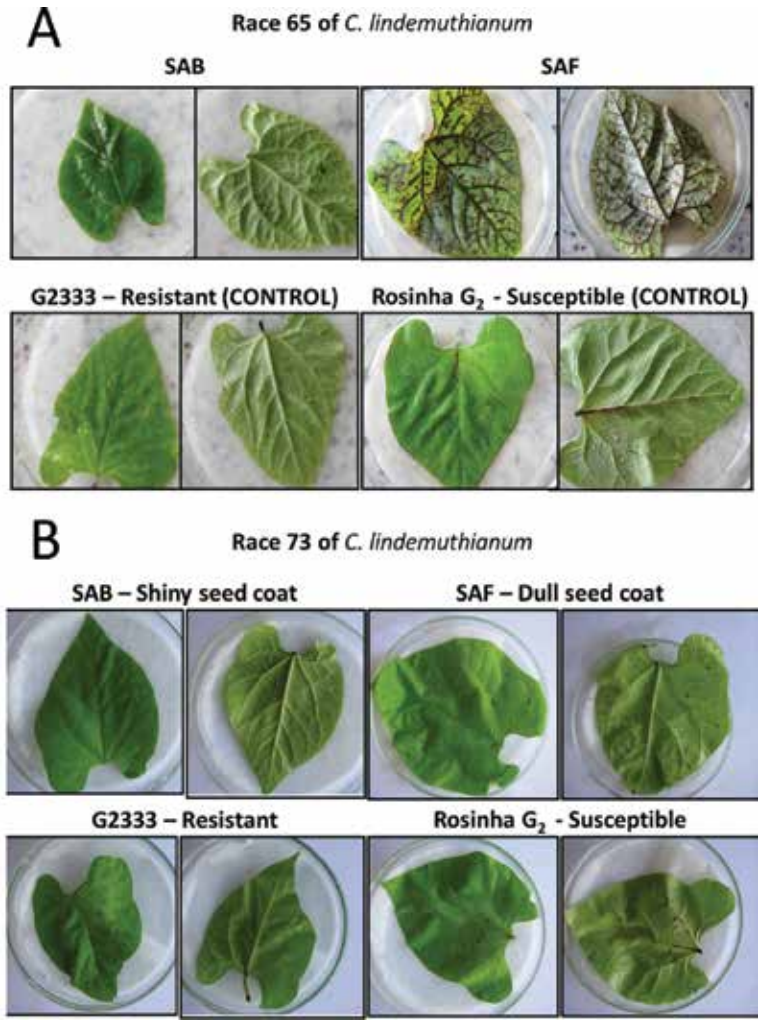

Figure 3. Screening for anthracnose resistance with races 65 and 73 (*Colletotrichum lindemuthianum*) on SAB and SAF plants. Controls: Rosinha G2 (susceptible) and G2333 (resistant).

Colletotrichum lindemuthianum. The genetic architecture and mechanisms of resistance to such diseases have been studied at genomic scales, identifying specific genes, pathways and QTL associated to each one (see [16, 36, 37]). As landraces are usually genetically structured and locally adapted, they might be the source of new alleles for disease resistance, which could be added to the disease resistance breeding programs.

Here, we present new findings obtained with experiments conducted with Serro Azul Brilhante and Serro Azul Fosco, as regards their variation for anthracnose resistance. The resistance degree to *C. lindemuthianum* was evaluated using the method of detached leaves [38], with modifications. Seeds of each parental line (SAB and SAF) and control standards from the literature (cultivars Rosinha G2 and G2333 are highly susceptible and resistant, respectively) were germinated and transferred to pots with substrate (Plantmax) in a greenhouse, and irrigated properly until the establishment of the first trifoliate (around 21 days). One leaf was collected from each plant and immediately placed in a suspension of 1.2×10^6 conidia mL^{-1} of *C. lindemuthianum* races 65 and 73 [39]. After 1 min, the leaves were placed in Petri dishes containing two layers of moistened filter paper. The plates were then incubated in a BOD type chamber and maintained in a photoperiod of 12 h at 21°C (± 2°C) for 7 days. For analysis, plant resistance was inferred according to the standards proposed by CIAT, using a scale from 1 (resistant—no symptoms) to 9 (susceptible—evident necrosis).

The results revealed an interesting difference between the parents SAF and SAB, used to constitute the segregating populations. The detached leaf method clearly showed that SAB was highly resistant to both races studied (65 and 73), while SAF showed high susceptibility to the *Colletotrichum* races, especially race 65 (**Figure 3**). The scores for disease resistance were significantly different among the two parents (SAB and SAF) for the race 65, but the average score with the race 73 also suggested higher degree susceptibility in SAF (**Table 1**). It is interesting to notice that the score for SAF was higher than the standard, which is used as a control for anthracnose susceptibility (Rosinha G2). In a similar manner, SAB was even more resistant than the standard resistant line G2333 (**Table 1**).

Genotype	**Mean score**	
	Race 65	**Race 73**
G2333 (resistant control)	2 ± 0.8	1 ± 0.5
Rosinha G2 (susceptible control)	5 ± 0.4	4 ± 0.2
SAB	1 ± 0.0	3 ± 0.6
SAF	7 ± 3.0	5 ± 1.1
Mean comparisons—Tukey's test		
G2333 × Rosinha G2	p = 0.012*	p = 0.001*
SAB × SAF	p = 0.0001**	p = 0.081 ns

*Significant at $P < 0.05$, **Significant at $P < 0.01$.

Table 1. Average scores of ***Colletotrichum lindemuthianum*** infection in leaves of the variants Serro Azul Brilhante (SAB) and Serro Azul Fosco (SAF), compared to G2333 (resistant control) and Rosinha G2 (susceptible control).

The evident difference between SAF and SAB needs further examination. It raises questions such as if the anthracnose resistance is somehow influenced by the glossiness of SAB. After all, genes related to anthracnose resistance are also located in chromosome 7 [40], but a specific study linking *Asp* to such genes is not yet available. However, it could be simply a new source of resistance originated from a mutation or to a combination of specific alleles conferring resistance. These are only speculations that need experimental clarification. The available genomic technology and the bioinformatic tools for constructing genetic maps with high resolution might be helpful in answering those questions.

5. Analysis and conservation of the genetic diversity of landraces

The genetic diversity of common bean landraces, cultivars and wild accessions has been investigated in multiple studies, mainly based on morphological and molecular markers in the last two decades (from 1995 to 2017). In the 1990s, studies have shown high genetic diversity based on morphological, enzyme and DNA-based markers (RAPD and microsatellites) [41–43]. After 2000, several studies involving amplified fragment length polymorphic (AFLP) [44, 45] and microsatellite [24, 46, 47] markers have been conducted. After 2010, with the advances of the sequencing technology, numerous papers have addressed the molecular diversity of common bean based on SNP markers [4, 48–51]. In general, most of the studies revealed that common beans are divided into two main gene pools, the Mesoamerican and Andean. In the case of Brazil, where common beans were introduced and have been cultivated mainly in small farming systems, varieties of both pools have been encountered. However, Burle et al. [24] investigated the genetic diversity of almost 300 landraces cultivated in Brazil and demonstrated that almost 80% of the genotypes have a Mesoamerican background, based on a population structure analysis with microsatellite polymorphisms.

In the case of Serro Azul, we also examined the molecular diversity of a sample of plants from both the variants (Serro Azul Brilhante and Serro Azul Fosco) by using the AFLP markers. Selective amplifications were done with four primer combinations (*EcoRI-A*/*MseI-CC*; *EcoRI-AC*/*MseI-CC*; *EcoRI-AT*/*MseI-CA*; *EcoRI-T*/*MseI-CA*). The AFLP gels revealed a considerable variation within and among SAB and SAF plants (**Figure 4**). An UPGMA tree was designed using Bionumerics software (**Figure 5**). Our data revealed a good separation among SAB and SAF samples, although some mixture was detected. In general, the SAB subgroup presented 75% of similarity with the SAF subgroup. Two plants of the SAF, however, were grouped within the SAB subgroup. From the analysis of the UPGMA tree, it is possible to infer that SAB plants were derived from SAF, which is consistent with the prior observation that SAB seeds were observed within 1 kg of SAF seeds harvested at a farm in Cunha, SP. Hereby, the high variability at the morphological levels can be verified at the molecular level, revealing a landrace with considerable genetic diversity to be explored. The development of the SAF × SAB population, for instance, revealed a high variation for seed size, color and water uptake [12]. Moreover, we showed that the parental lines have noticeable differences for disease resistance.

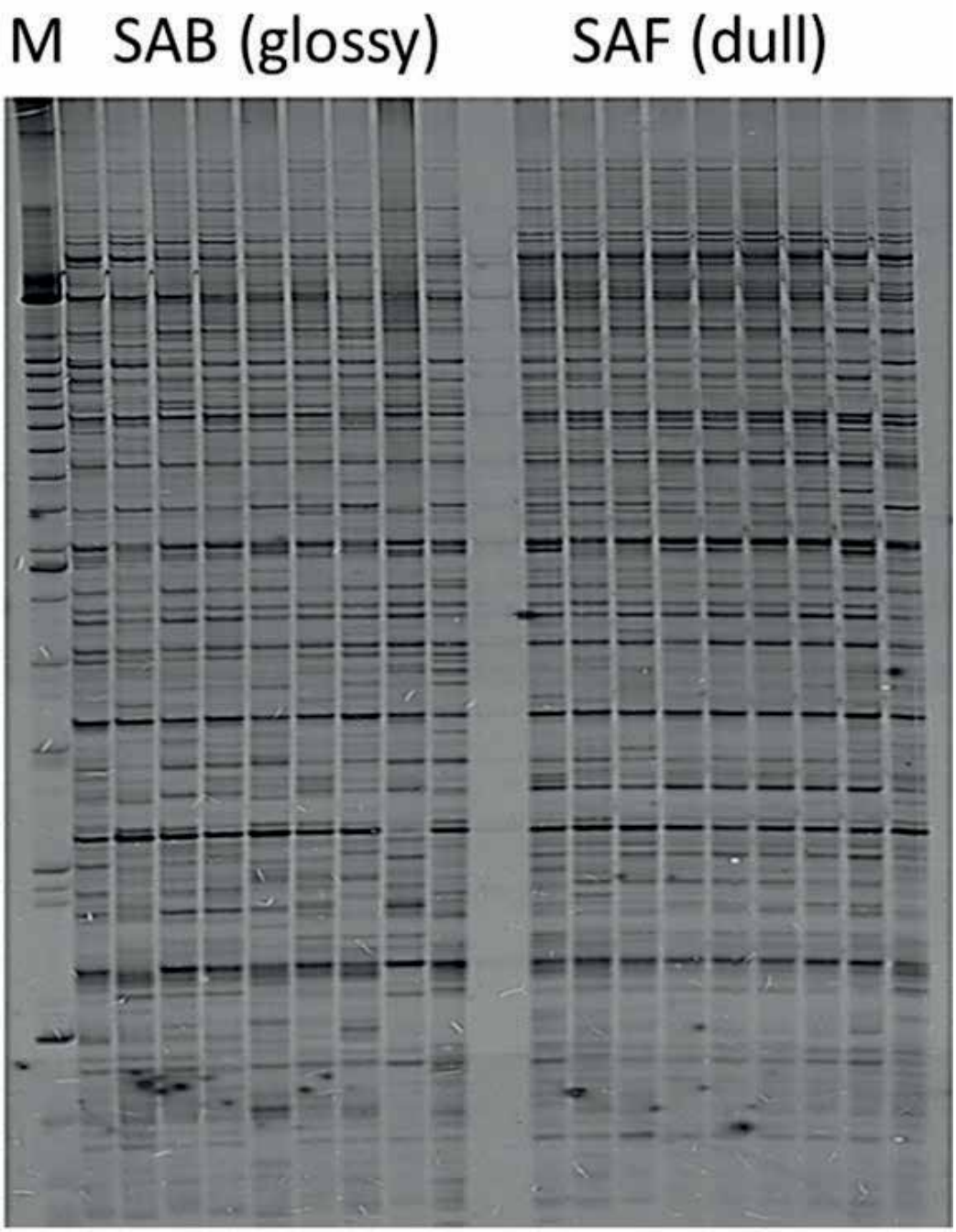

Figure 4. AFLP profile (primer combination *EcoRI-A/MseI-CC*) of individual plants from SAB (Serro Azul Brilhante) and SAF (Serro Azul Fosco). M stands for the ladder DNA of 100 bp.

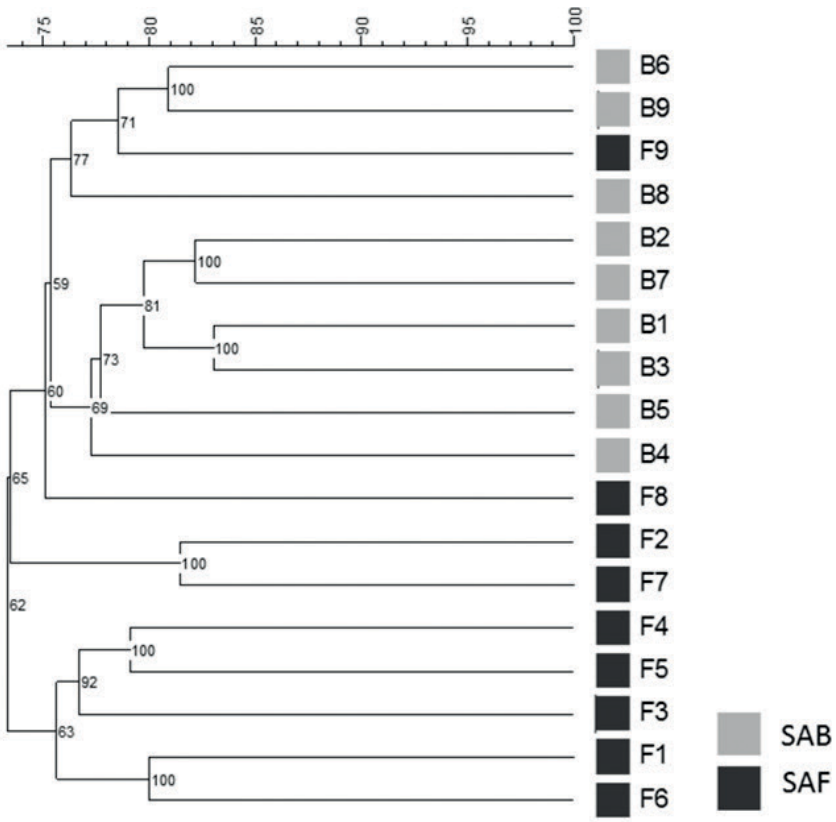

Figure 5. UPGMA tree based on Jaccard similarity analysis and AFLP profiles of the two variants of the landrace Serro Azul.

6. Perspectives of the conservation and use of bean landraces

The findings about Serro Azul provide interesting insights of the use and application of landraces in common bean breeding. A distinguishable morphologic diversity is noticeable within the landrace, which can be further explored to investigate genes responsible for color and glossiness [12]. SAB and SAF are consistently different at the molecular level as well, as revealed by the AFLP profiles. An examination of AFLP polymorphisms among F_4 lines of the cross SAF × SAB revealed a potential discrimination of color classes in the population by the molecular approach [12]. Furthermore, Serro Azul and the population developed might be used to investigate the genetic control for such incredible difference in anthracnose resistance between SAB and SAF. Another interesting observation comes from field observations where the SAF × SAB population was being tested. Usually, lines with similar features to the SAB parental line, especially the seed glossiness, presented very low incidence of bruchid attacks. On the other hand, SAF-derived lines were usually susceptible to the insects, leading to damages to the seeds.

The remarkable variability of Serro Azul and the interesting association with glossy seeds (SAB) with resistance to anthracnose and bruchids raises further research questions and opportunities for new crosses. As previously suggested, seed coat glossiness might after all have an important role in protecting seeds against biotic stresses, as SAB has shown. Conversely, we have not demonstrated that the disease resistance of SAB plants has association with glossiness, which needs more experiments. Either way, this is an important feature which might be explored in depth with the population derived from this landrace to appropriately answer this question. Landraces such as Serro Azul hold particularities that should not be disregarded, after all, local communities need those seeds for their supply, and they have traits of high interest to be explored by breeders, especially concerning the threat of anthracnose and insects to common bean cultivation.

From this example to a wider set of landraces, several traits of interest might be improved with the use of distinct allele combinations if not new alleles provided by such genetic materials. Burle et al. [24] analyzed 279 landraces of common bean from Brazil and discovered considerable genetic diversity among all the accessions evaluated. Those genotypes are distributed from colder to warmer and from wetter to drier areas in the country. The local adaptation implicated in such genotypes has implicated in potential sources of disease and insect-resistant accessions. Moreover, the climatic diversity provides the potential for adapting to distinct abiotic stresses. New sources for tolerance to drought, soil salinity, high and low temperatures are to be investigated from these collections. In fact, Burle et al. [10] continued the previous work and integrated phenotypic evaluations to the genetic analysis of the same 279 landraces. The authors screened these accessions based on 22 morphological traits, including resistance to rust and common bacterial blight, yield, flowering time, determinacy and growth habit, seed coat color and brilliancy, among others. The study provided valorous information for supporting initiatives toward conservation and management of the accessions. It also allowed to detect the particularities of landraces and how they can be explored in controlled crosses for designing new populations and cultivars.

7. Conclusions

In this chapter, we described the importance of the conservation and study of landraces of common bean. As an example, to go through the extent of morphological, biochemical and genetic aspects of landraces, we outlined the previous results as well as the new findings about the Brazilian landrace Serro Azul. This landrace has been produced by local farmers from Sao Paulo state and has remarkable features for exploring variation of seed and yield traits in common bean, constituting an additional and valuable genetic resource for germplasm collections. We also showed a genetic diversity analysis of Serro Azul by examining the molecular variability within subsamples of SAB and SAF, based on amplified fragment length polymorphic (AFLP) markers. Furthermore, a detached leaf assay for screening anthracnose resistance was employed for both variants within this landrace, revealing remarkable differences between the highly resistant Serro Azul Brilhante (glossy seed coat) and the susceptible Serro Azul Fosco (opaque seed coat). Together, these results demonstrated the importance of

studying genetic aspects related to traits such as color, glossiness, disease resistance and yield components. This is necessary to conserve such valuable resources as the ones maintained by small farming systems. It may as well be applicable to other landraces of common bean, in order to provide better understanding of the genetic resources available and how they can be explored in favor of the farmers and breeders.

Acknowledgements

We thank Monica Rossi Lanzoni and Prof. Adriana Pinheiro Martinelli for their assistance in the microscopy analyses. We also thank the assistance and materials provided by the Agronomic Institute of Campinas for performing the analyses with the races of *C. lindemuthianum* and the valorous technical assistance of Ana Luiza Ahern Beraldo in the experiments. The results presented in this chapter received funding from the National Council for Scientific and Technological Development (CNPq).

Conflict of interest

The authors declare no conflicts of interest.

Author details

Enéas Ricardo Konzen[1,2]* and Siu Mui Tsai[2]

*Address all correspondence to: erkonzen@gmail.com

1 State University of Santa Catarina, Lages, SC, Brazil

2 Center for Nuclear Energy in Agriculture, University of São Paulo, Piracicaba, SP, Brazil

References

[1] Broughton WJ, Hernández G, Blair M, Beebe S, Gepts P, Vanderleyden J. Beans (Phaseolus spp.)—Model food legumes. Plant and Soil. 2003;**252**(1):55-128

[2] Beebe SE, Rao IM, Devi MJ, Polania J. Common beans, biodiversity, and multiple stresses: Challenges of drought resistance in tropical soils. Crop and Pasture Science 2014;**65**(7):667-675. DOI: 10.1071/CP13303

[3] Gepts P, Osborn TC, Rashka K, Bliss FA. Phaseolin-protein variability in wild forms and landraces of the common bean (*Phaseolus vulgaris*): Evidence for multiple centers of domestication. Economic Botany. 1986;**40**(4):451-468

[4] Schmutz J, Mcclean PE, Mamidi S, Wu GA, Cannon SB, Grimwood J, et al. A reference genome for common bean and genome-wide analysis of dual domestications. Nature Publishing Group. 2014;**46**(7):707-713. DOI: 10.1038/ng.3008

[5] Kwak M, Gepts P. Structure of genetic diversity in the two major gene pools of common bean (*Phaseolus vulgaris* L., Fabaceae). Theoretical and Applied Genetics. 2009;**118**(5):979-992

[6] Bitocchi E, Bellucci E, Giardini A, Rau D, Rodriguez M, Biagetti E, et al. Molecular analysis of the parallel domestication of the common bean (*Phaseolus vulgaris*) in Mesoamerica and the Andes. New Phytologist. 2013;**197**(1):300-313. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23126683

[7] Koinange EMK, Koinange EMK, Singh SP, Singh SP, Gepts P, Gepts P. Genetic control of the domestication syndrome in common-bean. Crop Science. 1996;**36**:1037-1045

[8] Bitocchi E, Rau D, Bellucci E, Rodriguez M, Murgia ML, Gioia T, et al. Beans (Phaseolus ssp.) as a model for understanding crop evolution. Frontiers in Plant Science. 2017;**8**(May):1-21. Available from: http://journal.frontiersin.org/article/10.3389/fpls.2017.00722/full

[9] Freitas FDO. Evidências genético-arqueológicas sobre a origem do feijão comum no Brasil. Pesqui Agropecu Bras. 2006;**41**(7):1199-1203

[10] Burle ML, Fonseca JR, del Peloso MJ, Melo LC, Temple SR, Gepts P. Integrating phenotypic evaluations with a molecular diversity assessment of a Brazilian collection of common bean landraces. Crop Science. 2011;**51**(6):2668-2680

[11] De Oliveira WS, Meinhardt LW, Sessitsch A, Tsai SM. Analysis of Phaseolus–Rhizobium interactions in a subsistence farming system. Plant and Soil. 1998;**204**(1):107-115

[12] Konzen ER, Tsai SM. Seed coat shininess in *Phaseolus vulgaris*: Rescuing a neglected trait by its screening on commercial lines and landraces. 2014;**6**(8):1-18

[13] De Oliveira WS, Tsai SM. Bean cultivation in a subsistence farming system in the northeast of São Paulo state—Brazil. Science in Agriculture. 2001;**58**(2):265-269

[14] Bassett MJ. The margo (mar) seedcoat color gene is a synonym for the joker (j) locus in common bean. Journal of American Society for Horticultural Science. 1996;**121**(6):1028-1031

[15] Pérez-Vega E, Pañeda A, Rodríguez-Suárez C, Campa A, Giraldez R, Ferreira JJ. Mapping of QTLs for morpho-agronomic and seed quality traits in a RIL population of common bean (*Phaseolus vulgaris* L.). Theoretical and Applied Genetics. 2010;**120**(7):1367-1380

[16] Perseguini JMKC, Oblessuc PR, Rosa JRBF, Gomes KA, Chiorato AF, Carbonell SAM, et al. Genome-wide association studies of anthracnose and angular leaf spot resistance in common bean (*Phaseolus vulgaris* L.). PLOS One. 2016;**11**(3):1-19

[17] Wu J, Zhu J, Wang L, Wang S. Genome-wide association study identifies NBS-LRR-encoding genes related with anthracnose and common bacterial blight in the common bean. Frontiers in Plant Science. 2017;**8**(August):1-15. Available from: http://journal.frontiersin.org/article/10.3389/fpls.2017.01398/full

[18] Tsai SM, Nodari RO, Moon DH, Camargo LEA, Vencovsky R, Gepts P. QTL mapping for nodule number and common bacterial blight in *Phaseolus vulgaris* L. Plant and soil. 1998;**204**(1):135-145

[19] Mcclean PE, Lee RK, Otto C, Gepts P, Bassett MJ. Molecular and phenotypic pattern and color in common bean (*Phaseolus vulgaris* L.). The Journal of Heredity. 2002;**93**:148-152

[20] Beninger CW, Hosfield GL, Bassett MJ, Owens S. Chemical and morphological expression of the B and Asp seedcoat genes in *Phaseolus vulgaris*. Journal of the American Society for Horticultural Science. 2000;**125**(1):52-58. Available from: http://journal.ashspublications.org/cgi/content/abstract/125/1/52

[21] Elia FM, Hosfield GL, Kelly JD, Uebersax M. Genetic analysis and interrelationships between traits for cooking time, water absortion, and protein and tannin content of andean dry beans. Journal of the American Society for Horticultural Science. 1997;**122**(4):512-518

[22] Bushey SM, Hosfield GL, Owens S. The role of the epicuticular wax layer in water movement across the bean seed. Annual Report of the Bean Improvement Cooperative. 2001;**45**:12-13

[23] Garcia RAV, Rangel PN, Bassinello PZ, Brondani C, Melo LC, Sibov ST, Vianello-Brondani RP. QTL mapping for the cooking time of common beans. Euphytica. 2012;**186**(3): 779-792

[24] Burle ML, Fonseca JR, Kami JA, Gepts P. Microsatellite diversity and genetic structure among common bean (*Phaseolus vulgaris* L.) landraces in Brazil, a secondary center of diversity. Theoretical and Applied Genetics. 2010;**121**(5):801-813

[25] Beninger CW, Hosfield GL, Nair MG. Flavonol glycosides from the seed coat of a new Manteca-type dry bean (*Phaseolus vulgaris* L.). Journal of Agricultural and Food Chemistry. 1998;**46**(8):2906-2910. DOI: 10.1021/jf9801522

[26] Reinprecht Y, Yadegari Z, Perry GE, Siddiqua M, Wright LC, McClean PE, et al. In silico comparison of genomic regions containing genes coding for enzymes and transcription factors for the phenylpropanoid pathway in *Phaseolus vulgaris* L. and *Glycine max* L. Merr. Frontiers in Plant Science. 2013;**4**(September):1-25. Available from: http://journal.frontiersin.org/article/10.3389/fpls.2013.00317/abstract

[27] Nesi N, Jond C, Debeaujon I, Caboche M, Lepiniec L. The Arabidopsis TT2 gene encodes an R2R3 MYB domain protein that acts as a key determinant for proanthocyanidin accumulation in developing seed. The Plant Cell. 2001;**13**(9):2099-2114

[28] Zhao J, Pang Y, Dixon RA. The mysteries of proanthocyanidin transport and polymerization. Plant Physiology. 2010;**153**(2):437-443. Available from: http://www.plantphysiol.org/cgi/doi/10.1104/pp.110.155432

[29] Ropiak HM, Lachmann P, Ramsay A, Green RJ, Mueller-Harvey I. Identification of structural features of condensed tannins that affect protein aggregation. PLoS One. 2017;**12**(1):1-23

[30] Díaz AM, Caldas GV, Blair MW. Concentrations of condensed tannins and anthocyanins in common bean seed coats. Food Research International. 2010;**43**(2):595-601

[31] Beninger CW, Hosfield GL. Antioxidant activity of extracts, condensed tannin fractions, and pure flavonoids from *Phaseolus vulgaris* L. seed coat color genotypes. Journal of Agricultural and Food Chemistry. 2003;**51**(27):7879-7883. Available from: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=14690368

[32] Dzomba P, Togarepi E, Mupa M. African journal of agricultural research anthocyanin content and antioxidant activities of common bean species (*Phaseolus vulgaris* L.) grown in Mashonaland Central, Zimbabwe. 2013;**8**(25):3330-3333. Available from: http://www.academicjournals.org/AJAR

[33] Lila MA. Anthocyanins and human health: An in vitro investigative approach. Journal of Biomedicine and Biotechnology. 2004;**2004**(5):306-313. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1082894&tool=pmcentrez&rendertype=abstract

[34] Peña-Sanhueza D, Inostroza-Blancheteau C, Ribera-Fonseca A, Reyes-Díaz M. Anthocyanins in Berries and Their Potential Use in Human Health, Superfood and Functional Food - The Development of Superfoods and Their Roles as Medicine. In: Shiomi N, editor. Croatia: InTech; 2017. DOI: 10.5772/67104. Available from: https://www.intechopen.com/books/superfood-and-functional-food-the-development-of-superfoods-and-their-roles-as-medicine/anthocyanins-in-berries-and-their-potential-use-in-human-health

[35] Wang LS, Stoner G. Anthocyanins and their role in cancer prevention. Cancer Letters. 2008;**269**(2):281-290

[36] Akhavan A, Bahar M, Askarian H, Lak MR, Nazemi A, Zamani Z. Bean common bacterial blight: Pathogen epiphytic life and effect of irrigation practices. Springerplus [Internet]. 2013;**2**(1):41. Available from: http://springerplus.springeropen.com/articles/10.1186/2193-1801-2-41

[37] Oblessuc PR, Matiolli CC, Chiorato AF, Camargo LEA, Benchimol-Reis LL, Melotto M. Common bean reaction to angular leaf spot comprises transcriptional modulation of genes in the ALS10.1 QTL. Frontiers in Plant Science. 2015;**6**(March):1-11. Available from: http://www.frontiersin.org/Plant-Microbe_Interaction/10.3389/fpls.2015.00152/abstract

[38] Tu JC. A detached leaf technique for screening beans (*Phaseolus vulgaris* L.) *in vitro* against anthracnose (Colletotrichum lindemuthianum). Canadian journal of plant science. 1986;**66**(3):805-809

[39] Ribeiro T, De JA, Esteves F, Silva DA, Guilherme J, Gonçalves R, et al. Classification of Colletotrichum lindemuthianum races in differential cultivars of common bean. Frontiers in Plant Science. 2016;**38**(2):179-184. Available from: http://www.uem.br/acta

[40] Campa A, Giraldez R, Ferreira JJ. Genetic analysis of the resistance to eight anthracnose races in the common bean differential cultivar Kaboon. Phytopathology. 2011;**101**(6):757-764

Edited by Oscar Grillo

In recent years, all over the world, the attention paid to local and traditional productions is growing, especially in the agro-food sector. Maybe, it is not only due to the impact of globalization and the social and economic changes but also due to the increased consideration to health and nutritional aspects of food. Hence, for economic, social, historical, and nutritional reasons, this trend has led to the rediscovery and reuse of landraces of many different crops, responding to requests for more and more demanding market.

This volume collects examples of local crops and old landraces of different areas of the planet that testify the extreme importance of the relation existing among a land, the local productions, the historical traditions, the conservation of biodiversity, the health benefits, the environmental impact and the local economies, also including the significance to dedicate resources to scientific researches in local crops.

Published in London, UK

IntechOpen

ISBN 978-1-83881-396-3